GENDERED PATHS INTO STEM. DISPARITIES BETWEEN FEMALES AND MALES IN STEM OVER THE LIFE-SPAN

EDITED BY: Bernhard Ertl, Silke Luttenberger, M. Gail Jones, Rebecca Lazarides and Manuela Paechter

PUBLISHED IN: Frontiers in Psychology, Frontiers in Education and Frontiers in Sociology

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88963-439-2
DOI 10.3389/978-2-88963-439-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

GENDERED PATHS INTO STEM. DISPARITIES BETWEEN FEMALES AND MALES IN STEM OVER THE LIFE-SPAN

Topic Editors:
Bernhard Ertl, Universität der Bundeswehr München, Germany
Silke Luttenberger, University College of Teacher Education Styria, Austria
M. Gail Jones, North Carolina State University, United States
Rebecca Lazarides, University of Potsdam, Germany
Manuela Paechter, University of Graz, Austria

Citation: Ertl, B., Luttenberger, S., Jones, M. G., Lazarides, R., Paechter, M., eds. (2020). Gendered Paths into STEM. Disparities Between Females and Males in STEM Over the Life-Span. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88963-439-2

Table of Contents

06 ***Editorial: Gendered Paths into STEM. Disparities Between Females and Males in STEM Over the Life-Span***
Bernhard Ertl, Silke Luttenberger, Rebecca Lazarides, M. Gail Jones and Manuela Paechter

10 ***Do Gender-Related Stereotypes Affect Spatial Performance? Exploring When, How and to Whom Using a Chronometric Two-Choice Mental Rotation Task***
Carla Sanchis-Segura, Naiara Aguirre, Álvaro J. Cruz-Gómez, Noemí Solozano and Cristina Forn

27 ***Gender Stereotypes in a Children's Television Program: Effects on Girls' and Boys' Stereotype Endorsement, Math Performance, Motivational Dispositions, and Attitudes***
Eike Wille, Hanna Gaspard, Ulrich Trautwein, Kerstin Oschatz, Katharina Scheiter and Benjamin Nagengast

44 ***Sex Differences in Gains Among Hispanic Pre-kindergartners' Mental Rotation Skills***
Carla Abad, Rosalie Odean and Shannon M. Pruden

53 ***Changes in United States Latino/a High School Students' Science Motivational Beliefs: Within Group Differences Across Science Subjects, Gender, Immigrant Status, and Perceived Support***
Ta-yang Hsieh, Yangyang Liu and Sandra D. Simpkins

65 ***Mothers and Fathers—Who Matters for STEM Performance? Gender-Specific Associations Between STEM Performance, Parental Pressure, and Support During Adolescence***
Frances Hoferichter and Diana Raufelder

75 ***Self-Concept and Achievement in Math Among Australian Primary Students: Gender and Culture Issues***
Feifei Han

84 ***Contextual Choices in Online Physics Problems: Promising Insights Into Closing the Gender Gap***
Samuel R. Wheeler and Margaret R. Blanchard

98 ***STEM/Non-STEM Divide Structures Undergraduate Beliefs About Gender and Talent in Academia***
Kimberlyn A. Bailey, David Horacek, Steven Worthington, Ampalavanar Nanthakumar, Scott Preston and Carolina C. Ilie

107 ***Self-Concept Profiles in Lower Secondary Level – An Explanation for Gender Differences in Science Course Selection?***
Steffani Saß and Nele Kampa

121 ***The Interest Profiles and Interest Congruence of Male and Female Students in STEM and Non-STEM Fields***
Bernhard Ertl and Florian G. Hartmann

139 ***Pipped at the Post: Knowledge Gaps and Expected Low Parental IT Competence Ratings Affect Young Women's Awakening Interest in Professional Careers in Information Science***
Angela Schorr

157 ***Young Spanish People's Gendered Representations of People Working in STEM. A Qualitative Study***
Milagros Sáinz, José-Luis Martínez-Cantos, María Rodó-de-Zárate, María José Romano, Lidia Arroyo and Sergi Fàbregues

169 ***Traditional Gender Role Beliefs and Career Attainment in STEM: A Gendered Story?***
Anna-Lena Dicke, Nayssan Safavian and Jacquelynne S. Eccles

183 ***Do Teachers' Beliefs About Math Aptitude and Brilliance Explain Gender Differences in Children's Math Ability Self-Concept?***
Anke Heyder, Ricarda Steinmayr and Ursula Kessels

194 ***Beliefs in "Brilliance" and Belonging Uncertainty in Male and Female STEM Students***
Anne Deiglmayr, Elsbeth Stern and Renate Schubert

201 ***Gendered Paths Into STEM-Related and Language-Related Careers: Girls' and Boys' Motivational Beliefs and Career Plans in Math and Language Arts***
Rebecca Lazarides and Fani Lauermann

218 ***Self-Concept and Support Experienced in School as Key Variables for the Motivation of Women Enrolled in STEM Subjects With a Low and Moderate Proportion of Females***
Silke Luttenberger, Manuela Paechter and Bernhard Ertl

228 ***The Impacts of Gender and Subject on Experience of Competence and Autonomy in STEM***
Sabrina Sobieraj and Nicole C. Krämer

244 ***Adolescents' Motivational Profiles in Mathematics and Science: Associations With Achievement Striving, Career Aspirations and Psychological Wellbeing***
Helen M. G. Watt, Micaela Bucich and Liam Dacosta

267 ***Gendered Development of Motivational Belief Patterns in Mathematics Across a School Year and Career Plans in Math-Related Fields***
Julia Dietrich and Rebecca Lazarides

272 ***The Relations of Science Task Values, Self-Concept of Ability, and STEM Aspirations Among Finnish Students From First to Second Grade***
Janica Vinni-Laakso, Jiesi Guo, Kalle Juuti, Anni Loukomies, Jari Lavonen and Katariina Salmela-Aro

287 ***The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations***
Elena Makarova, Belinda Aeschlimann and Walter Herzog

298 ***Sources of Male and Female Students' Belonging Uncertainty in the Computer Sciences***
Elisabeth Höhne and Lysann Zander

311 ***Differential Contributions of Empathy to Math Achievement in Women and Men***
Nermine Ghazy, Eleanor Ratner and Miriam Rosenberg-Lee

324 ***Social or Economic Goals? The Professional Goal Orientation of Students Enrolled in STEM and Non-STEM Majors in University***
Ilka Wolter, Lisa Ehrtmann, Tina Seidel and Barbara Drechsel

335 ***The More Interest, the Less Effort Cost Perception and Effort Avoidance***
Juyeon Song, Sung-il Kim and Mimi Bong
348 ***Mathematics Self-Concept in New Zealand Elementary School Students: Evaluating Age-Related Decline***
Penelope W. St J. Watson, Christine M. Rubie-Davies and Kane Meissel
360 ***The Role of Social and Ability Belonging in Men's and Women's pSTEM Persistence***
Sarah Banchefsky, Karyn L. Lewis and Tiffany A. Ito
376 ***It Takes More Than One Swallow to Make a Summer: Measures to Foster Girls' and Women's Pathways Into STEM***
Silke Luttenberger, Petra Steinlechner, Bernhard Ertl and Manuela Paechter

frontiers in Psychology

EDITORIAL
published: 17 December 2019
doi: 10.3389/fpsyg.2019.02758

Editorial: Gendered Paths into STEM. Disparities Between Females and Males in STEM Over the Life-Span

Bernhard Ertl[1]*, *Silke Luttenberger*[2], *Rebecca Lazarides*[3], *M. Gail Jones*[4] and *Manuela Paechter*[5]

[1] Institute of Education, Universität der Bundeswehr München, Neubiberg, Germany, [2] Institute for Early Childhood and Primary Teacher Education, University College of Teacher Education Styria, Graz, Austria, [3] Department of Education, University of Potsdam, Potsdam, Germany, [4] Department of STEM Education, North Carolina State University, Raleigh, NC, United States, [5] Department of Psychology, University of Graz, Graz, Austria

Keywords: attributions, self-concept, motivation, gender-sensitive didactics, occupational choices

Editorial on the Research Topic

Gendered Paths into STEM. Disparities Between Females and Males in STEM Over the Life-Span

OPEN ACCESS

Edited and reviewed by:
Yenchun Jim Wu, National Taiwan Normal University, Taiwan

***Correspondence:**
Bernhard Ertl
bernhard.ertl@unibw.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *08 November 2019*
Accepted: *22 November 2019*
Published: *17 December 2019*

Citation:
Ertl B, Luttenberger S, Lazarides R, Jones MG and Paechter M (2019) Editorial: Gendered Paths into STEM. Disparities Between Females and Males in STEM Over the Life-Span. Front. Psychol. 10:2758. doi: 10.3389/fpsyg.2019.02758

Choosing a career path into STEM (Science, Technology, Engineering, and Mathematics) is a longitudinal process rather than an *ad-hoc* decision: experiences in childhood and school form individuals' interests, motivation, and ability beliefs—their expectations according to Eccles et al. (1983). These serve as basis for a decision against or toward STEM. However, while youth are considering careers, barriers can emerge, for example students may form stereotyped impressions of STEM as a "male" domain or develop perceptions that brilliance is a prerequisite for STEM attainments. Such assumptions downgrade expectations and often shape women's as well as minority students' self-evaluation of not being suited to a career in STEM.

Altogether, deciding for and following a specific career path is a developmental process (Gottfredson, 2005) of circumscription and compromise and female students often rule out STEM professions during this process. According to expectancy-value theories (EVT; e.g., Eccles et al., 1983), an individual evaluates during this process the balance between the personal expectations for success (resp. activity specific ability beliefs) and the subjective task value for achievement-related values, engagement, and persistence. This evaluation is influenced by the broader context of socializers and the milieu that frame the individual's perceptions and interpretations of experiences. Many papers in this Research Topic refer to this theoretical approach.

While EVT focus the interactions of the different factors during balancing expectancies and values, the Social Cognitive Career Theory (SCCT; Lent et al., 1994, 2018) proposes a step-wise model how personal and environmental variables interact to finally shape choices for performance domains and attainment. The model proposes that (1) person inputs (like predispositions and gender) as well as (2) background contextual affordances and societal characteristics (like cultural norms) shape (3) learning experiences that lead to individual attainments (10) which then may receive feedback from the environment. These learning experiences contextualize an individual's expectations regarding one's self-efficacy (4) and consequently also one's expectations about outcomes of one's actions and attainments (5). Task values such as utility values or interest (7) develop out of self-efficacy and outcome expectations and provide a basis for choice goals (8) and choice actions (9). However, contextual influences proximal to choice behavior (6) also influence interests and choices. Finally, (10) performance domains and attainments result from choice actions.

When connecting the model of the SCCT with Gottfredson's (2005) assumptions of developmental processes, it becomes clear that choice actions do not simply result from predispositions, aptitudes, learning experiences, self-assessments, and interests but that contextual factors moderate such processes. While EVT describe the complex interactions of this moderation process (see Eccles et al., 1983), SCCT rather focuses the steps from the individual's personal inputs toward choice actions moderated by expectancies and contextual factors. These are especially important for female students' career paths into STEM because cultural stereotypes of STEM as a "male" domain as well as interactions with teachers and significant others may influence women to steer away from STEM to more "female" domains or to not consider a STEM career at all (see Ertl et al., 2017).

Against this background, the present Research Topic investigates career decisions, to illustrate the complexity and difficulties of steering more females onto a STEM career path, as well as to summarize evidence about female students' career paths into STEM. The Research Topic comprises 30 articles by 94 authors from ten countries in Europe, America, Oceania, Asia, and Africa. We will structure the editorial according to the factors of SCCT that include expectancies as well as the steps toward a decision for STEM.

(1) **Person Inputs**
Person inputs may affect self-assessments, motivation, behavior, and thus attainments and can be seen as a starting point for a career path into STEM. If, for example, a person finds his or her aptitude in STEM lacking and conceives her or his talents being outside of STEM, she/he is hardly likely to go into a STEM field. Person inputs are investigated with respect to motivational, emotional, cognitive, or socio-demographic aspects. In the studies presented in this Research Topic motivational aspects relate to goal orientations (Wolter et al.), emotional aspects to empathy as predictor for math achievement (Ghazy et al.), and cognitive aspects relate to visuospatial skills that are often seen as key aptitude for STEM with a clear gender difference in favor of men. Abad et al. as well as Sanchis-Segura et al. look deeper into the issue of visuospatial skills raising the point that tests for the respective skills are often subject to gender framing, a background context, that affects their results. Such background contexts can be also found in Hsieh et al.'s work that focuses on ethnicity and immigrant status as predictors for motivational beliefs.

(2) **Background Contextual Affordances**
Background contextual affordances relate to culture and cultural norms in which a person is embedded. They provide an indirect impact on all contributions of this Research Topic. The impact of this factor becomes most obvious in the contributions by Sachnis-Segura et al., who discuss how tests for visuospatial performance are constructed (favoring predominantly men), and by Hsieh et al. and Watson et al. who both discuss differences between ethnicities that may include different cultural values.

(3) **Learning Experiences**
Learning experiences play a major role in the model of Lent et al. (1994) by shaping a student's self-efficacy and outcome expectations. They shape a students' feeling of belonging to a learning domain or not (Banchefsky et al.; Deiglmayr et al.; Höhne and Zander). In a long-term process, learning experiences are related to attainments, outcomes, and subsequent feedback and thus mostly need a longitudinal design for their investigation. In the Research Topic, several longitudinal studies are concerned with these variables, especially Dietrich and Lazarides and Hsieh et al. who investigate the development of motivational belief patterns, and Vinni-Laakso et al. as well as Watson et al. who focus on the long-term development of students' self-concept.

(4) **Self-Efficacy and Self-Concept Expectations**
Self-efficacy expectations, for example, are subject-specific academic self-concept or ability beliefs. Self-efficacy expectations are a crucial aspect of career paths into STEM and often vary by gender. Large scale studies such as PISA (OECD, 2015) confirm that—even in case of identical academic outcomes and assessments—the self-concept for STEM is lower for female than for male students. Consequently, several contributions delve deeper into self-concept and ability beliefs. For example, Watson et al. looks closer into the gender-related decline of the self-concept in mathematics. Factors contributing to such processes and to the development of a student's self-concept for STEM in general are investigated by Heyder et al. who explore the impact of teacher expectations as well as by Höhne and Zander who analyze the impact of belongingness. The impact of the self-concept on further developments is investigated by Han who analyzes the relationship between self-concept and achievements, by Luttenberger et al. as well as by Sobieraj and Krämer who focus on the relationship between self-concept and motivation in STEM, and by Saß and Kampa who investigate the impact of self-concept profiles on course selection. Finally, Dietrich and Lazarides as well as Vinni-Laakso et al. analyze to which degree motivational belief patterns are associated with math-related career plans.

(5) **Expectations About Outcomes**
While self-efficacy expectations focus on the estimation of one's own ability, outcome expectations result from an assessment to which degree one's own skills are sufficient to achieve satisfactory outcomes in a field. In this sense, Kessels (2015) discusses women's belief that success in STEM careers is based on innate talent or even brilliance (which women typically believe not to have in STEM) as opposed to hard work and diligence. Consequently, female students often shy away from STEM career choices even if they achieve good grades. Such field-specific ability beliefs, for example that a successful STEM career requires brilliance, have impacts on women's emotions and motivation in STEM fields. Therefore, Deiglmayr et al. as well as Höhne and Zander investigate in a sample of female students and the degree to which such beliefs are associated with uncertainty and feelings of not belonging to the domain of STEM even though these students

major in a STEM field. Luttenberger et al. investigate the degree to which such beliefs predict motivation in STEM and Lazarides and Lauermann explore how beliefs may affect students' career plans. Hsieh et al. as well as Dicke et al. analyze how such beliefs develop over a long range in different STEM subjects and finally Bailey et al. investigate STEM and non-STEM undergraduates as well as academics discussing to which degree undergraduates' beliefs about talent in academia mirror those of academics.

(6) **Contextual Influences Proximal to Choice Behavior**
Thereby, Bailey et al. investigate contextual influences exerted by others and take up the hypothesis, that university graduates transfer their (stereotypical) ability beliefs to undergraduates. Such phenomena are in the scope of several contributions of this Research Topic focusing on the influence of teachers (Heyder et al.), parents (Hoferichter and Raufelder; Luttenberger et al.; Schorr), or the peer-group (Sáinz et al.). Apart from these personal influences, STEM subjects are often generally attributed stereotypically as being male, an aspect that is taken up by Makarova et al. as well as by Watson et al., and consequently female students in STEM often choose contexts that are to a lower degree regarded as being typically male subjects, for example biology instead of physics contexts if they are able to choose (see Wheeler and Blanchard). Such stereotypical perspectives may be reinforced by representations in TV, which is investigated by Wille et al. Generally, such stereotypical as well as traditional gender role beliefs taken up from personal contexts predict lower educational attainment and less inclination for studying STEM subjects—an issue that is investigated by Dicke et al.

(7) **Interests and Task Values**
Interests develop and deepen partly due to an individual's self-efficacy and outcome expectations—however, they are also shaped by contextual influences, for example, when interests are regarded as being inappropriate for a specific gender or when pursuing them seems to require too much effort (see for example Gottfredson, 2005) or task values (see Eccles et al., 1983). In this line, Song et al. investigate the impact of interest and effort on persistence. However, as Schorr discusses, interest is often subject to pre-conditions including personal competency and outcome expectations. Similarly, Sobieraj and Krämer analyze to which degree self-perceptions are conjoined with interest-related characteristics such as intrinsic motivation. In this sense, Ertl and Hartmann as well as Watt et al. bridge the gap between interest and motivational profiles and respective choice goals and actions. Lazarides and Lauermann investigated this relation with respect to task values and career aspirations.

(8) **Choice Goals**
Choice goals can be defined as students' career aspirations that either can go along with a student's interests or reveal deviations. Here, Ertl and Hartmann analyze to which degree students' interests fit to their career aspirations and they find a worse fit between interest and aspirations for STEM than for other subjects. Watt et al. identify different motivational profiles and discuss that especially disengaged students show lower STEM aspirations. Motivation and motivational belief-patterns and their impact on career plans are also discussed by Dietrich and Lazarides as well as by Lazarides and Lauermann, while Vinni-Laakso et al. analyze the impact of self-concept profiles on science course selection. Makarova et al., finally, expand the view on choice-goal section by discussing the impact of gender-science stereotypes on students' choice goals.

(9) **Choice Actions**
Specific choice actions are less predominant in the Research Topic, possibly because the transformation from a choice goal to a choice action is difficult to observe and to operationalize. Despite of such difficulties, Saß and Kampa aim at explaining science course selection by the impact of self-concept profiles. Sobieraj and Krämer apply a retrospective approach for explaining differences between STEM and non-STEM master students with respect to competence, motivational, and volitional variables.

(10) **Performance Domains and Attainment**
Ideally, those who have embarked on a STEM career and show persistence should experience satisfying outcomes such as high attainments (grades, professional success), feelings of belonging, joy, or life satisfaction. The two contributions which look closer into these concepts show that students' persistence in STEM is related to a feeling of belonging (Banchefsky et al.) and to interest (Song et al.). Regarding outcomes, Ghazy et al. analyze the role of empathy for math achievement and math scores and find different effects for male and for female students. Han also focuses on math performance scores and find a stereotype effect impeding female pupils. Similarly, Wille et al. investigate STEM stereotypes in a learning context and their differential effects on scores, stereotype endorsement, and belongingness. Sanchis-Segura et al. find similar effects for visuospatial tasks. The effect of framing tasks differently is investigated by Wheeler and Blanchard focusing on how biological contexts may facilitate female students' familiarizing with the context of force rather than traditional physics contexts. Hoferichter and Raufelder investigate the impact of parents' support and pressure on STEM performance and disclose differential effects for female and male students.

DEVELOPMENT OF STEM PATHWAYS

A specific aim of this Research Topic is to shed light on critical incidents or milestones on a STEM pathway over the life-span. Therefore, the topic covers evidence from kindergarten (Abad et al.) to adult STEM professionals (Dicke et al.). In between, all stages of formal education are well-covered including primary school (Han; Heyder at al.; Vinni-Laakso et al.; Watson et al.), lower secondary school (Hoferichter and Raufelder; Saß and Kampa; Song et al.; Wille et al.), and upper secondary school (Dietrich and Lazarides; Hsieh et al.; Lazarides and Lauermann; Makarova et al.; Schorr; Watt et al.; Wheeler and Blanchard) with a clear focus on university education (Bailey et al.; Banchefsky et al.; Deiglmayr et al.; Ertl and Hartmann; Ghazy et al.; Höhne and Zander; Luttenberger et al.; Sáinz et al.; Sanchis-Segura et al.; Sobieraj and Krämer; Wolter et al.). Some of these contributions focus on longitudinal developments, for example Abad et al. on the development of visuospatial skills, Dietrich and Lazarides as well as Hsieh et al. on the development of motivational belief patterns, Hoferichter and Raufelder on grades and parental influences, Vinni-Laakso et al. and Watson et al. on students' self-concept, and Dicke et al. on the development of STEM professionals.

HETEROGENEITY OF STEM SUBJECTS

Although the term STEM raises the impression of being a homogeneous academic domain, there are different definitions which vary in their broadness and some of them even include life sciences and social sciences into STEM (for a discussion, see Ertl et al., 2017). This Research Topic focuses on the core of STEM that covers natural sciences, technology, engineering, and mathematics. However, also within this narrow definition of STEM, authors point at differences between the subjects. They can be distinguished with respect to the proportion of women in a field (Ertl and Hartmann; Luttenberger et al.), with respect to specific subjects as for example in comparisons of mathematics and biology (Hoferichter and Raufelder); research can also refer to science in general (Watt et al.) or to a range of different subjects in the field of STEM (Deiglmayr et al., Hsieh et al.; Makarova et al.). Contributions that focus on one subject mostly investigate mathematics as key subject in STEM (Dietrich and Lazarides; Ghazy et al.; Han; Heyder et al.; Song et al.; Watson et al.; Wille et al.), followed by computer science (Höhne and Zander; Schorr) and physics (Wheeler and Blanchard). To uncover the special characteristics of STEM, some authors compare STEM subjects with NON-STEM (Bailey et al.; Dicke et al.; Ertl and Hartmann; Lazarides and Lauermann; Sanchis-Segura et al.; Sobieraj and Krämer; Wolter et al.). The remaining contributions focus on rather general aspects regarding STEM (Abad et al.; Banchefsky et al.; Luttenberger, Steinlechner et al.; Sáinz et al.; Saß and Kampa, as well as Vinni-Laakso et al.).

FACILITATION GENDERED PATHWAYS INTO STEM

Luttenberger, Steinlechner et al. finally comment on the development on an individual's STEM pathway from interests to a career goal and choice actions and its respective facilitation by shedding light on the importance of early career-related learning experiences as well as on removing external barriers on the path into STEM.

AUTHOR CONTRIBUTIONS

BE, SL, RL, MJ, and MP wrote this editorial.

REFERENCES

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). "Expectancies, values, and academic behaviors," in *Achievement and Achievement Motives,* ed. J. T. Spence (San Francisco, CA: Freeman), 75–146.

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Gottfredson, L. S. (2005). "Applying Gottfredson's theory of circumscription and compromise in career guidance and counseling," in *Career Development and Counseling. Putting Theory and Research to Work,* eds S. D. Brown and R. W. Lent (Hoboken, NJ: John Wiley and Sons), 71–100.

Kessels, U. (2015). Bridging the gap by enhancing the fit: how stereotypes about STEM clash with stereotypes about girls. *Int. J. Gend. Sci. Technol.* 7, 280–296.

Lent, R. W., Brown, S. D., and Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Voc. Behav.* 45, 79–122.

Lent, R. W., Sheu, H.-B., Miller, M. J., Cusick, M. E., Penn, L. T., and Truong, N. N. (2018). Predictors of science, technology, engineering, and mathematics choice options: a meta-analytic path analysis of the social-cognitive choice model by gender and race/ethnicity. *J. Couns. Psychol.* 65, 17–35. doi: 10.1037/cou0000243

OECD (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. Paris: OECD Publishing.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 24 July 2018
doi: 10.3389/fpsyg.2018.01261

Check for updates

Do Gender-Related Stereotypes Affect Spatial Performance? Exploring When, How and to Whom Using a Chronometric Two-Choice Mental Rotation Task

Carla Sanchis-Segura, Naiara Aguirre, Álvaro J. Cruz-Gómez, Noemí Solozano and Cristina Forn*

Departamento de Psicologia Básica, Clínica y Psicobiología, Facultad de Ciencias de la Salud, Universitat Jaume I, Castellón, Spain

OPEN ACCESS

Edited by:
Bernhard Ertl,
Universität der Bundeswehr München, Germany

Reviewed by:
Uner Tan,
Çukurova University, Turkey
David Stuart Smith,
Robert Gordon University, United Kingdom
Sigrid Schmitz,
Humboldt-Universität zu Berlin, Germany

***Correspondence:**
Carla Sanchis-Segura
csanchis@uji.es

Specialty section:
This article was submitted to Gender, Sex and Sexuality Studies, a section of the journal Frontiers in Psychology

Received: 17 May 2018
Accepted: 29 June 2018
Published: 24 July 2018

Citation:
Sanchis-Segura C, Aguirre N, Cruz-Gómez ÁJ, Solozano N and Forn C (2018) Do Gender-Related Stereotypes Affect Spatial Performance? Exploring When, How and to Whom Using a Chronometric Two-Choice Mental Rotation Task. Front. Psychol. 9:1261. doi: 10.3389/fpsyg.2018.01261

It is a common belief that males have superior visuospatial abilities and that differences in this and other cognitive domains (e.g., math) contribute to the reduced interest and low representation of girls and women in STEM education and professions. However, previous studies show that gender-related implicit associations and explicit beliefs, as well as situational variables, might affect cognitive performance in those gender-stereotyped domains and produce between-gender spurious differences. Therefore, the present study aimed to provide information on when, how and who might be affected by the situational reactivation of stereotypic gender-science beliefs/associations while performing a 3D mental rotation chronometric task (3DMRT). More specifically, we assessed the explicit beliefs and implicit associations (by the Implicit Association Test) held by female and male students of humanities and STEM majors and compared their performance in a 3DMRT after receiving stereotype- congruent, incongruent and nullifying experimental instructions. Our results show that implicit stereotypic gender-science associations correlate with 3DMRT performance in both females and males, but that inter-gender differences emerge only under stereotype-reactivating conditions. We also found that changes in self-confidence mediate these instructions' effects and that academic specialization moderates them, hence promoting 3DMRT performance differences between male and female humanities, but not STEM, students. Taken together, these observations suggest that the common statement "males have superior mental rotation abilities" simplifies a much more complex reality and might promote stereotypes which, in turn, might induce artefactual performance differences between females and males in such tasks.

Keywords: gender stereotypes, stereotype threat, mental rotation, implicit association test, STEM

INTRODUCTION

Although in elementary, middle, and high school, girls and boys take math and science courses in roughly equal numbers, only around 20 percent of STEM graduates are women, a number that declines even further in the workplace (Hill et al., 2010; European Commission, 2016). Because STEM related careers are expected to grow faster than the average rate for all occupations

(National Science Board, 2010) and are among the best paid jobs (National Association of Colleges and Employers, 2009; European Commission, 2016), the underrepresentation of women in STEM studies severely increases the risk of exclusion and precarization in their future incorporation into the labor market. Yet, this is not only a problem for women. The absence of women from STEM education and careers is a waste of talent for those fields (European Commission, 2014; Norland et al., 2016) and also an economic cost for society as a whole. Indeed, it has been estimated that closing the gender gap in the STEM field would increase the EU GDP per capita by 0.7–0.9% in 2030 and by 2.2–3.0% in 2050 (Maceira, 2017). Accordingly, the gender segregation that characterizes the STEM field at the educational and professional level is seen with increasing social and institutional concern (Hill et al., 2010; European Commission, 2016).

The underrepresentation of women in STEM studies and professions has been traditionally considered a consequence of an innate higher proficiency of males in math and visuospatial abilities (Benbow et al., 2000; Baron-Cohen, 2004). Popularized through expressions such as "math is hard for girls" (Barbie, 1994, see Ben-Zeev et al., 2005) or "women cannot read maps" (Pease and Pease, 2004), the notion that males excel over females in these cognitive domains has become a widely shared social belief. However, scientific evidence does not support these claims and presents a much more complex reality (Ceci et al., 2009; Wang and Degol, 2013).

Thus, although older studies regularly identified a males' advantage in math performance (Glennon and Callahan, 1968; Maccoby and Jacklin, 1974; Benbow and Stanley, 1980, 1983), more recent large-sample studies and metaanalyses have revealed that gender differences in mathematics achievement tend to be inconsistent and small (d = 0.05, Lindberg et al., 2010; d = 0.06; Voyer and Voyer, 2014). Moreover, the size and even the direction of average gender differences in math performance widely varies among countries (ds ranging −0.42 to 0.40, Else-Quest et al., 2010; OECD, 2016) and they are correlated to national gender equity indexes (Reilly, 2012). Similarly, the proportion of females and males scoring at the 95th or 99th percentiles also differ among countries (Guiso et al., 2008; Machin and Pekkarinen, 2008) and they are highly correlated to national gender equality indexes, (Guiso et al., 2008; Hyde and Mertz, 2009). Finally, theoretical models demonstrate that the number of women in STEM studies and professions is substantially lower than that predicted from their math performance (Hyde and Mertz, 2009). Taken together, these and other data (reviewed in Spelke, 2005; Halpern et al., 2007; Ceci et al., 2009; Wang and Degol, 2013) strongly argue against the notion that males have innate or "hard-wired" superior math abilities that could account for the underrepresentation of women in STEM studies and occupations.

On the other hand, spatial ability is the cognitive domain in which differences between males and females are most commonly replicated and reported (Voyer et al., 1995; Hyde, 2014). Among the tasks in which such differences are observed, mental rotation tasks (MRT) produce the largest effects (Linn and Petersen, 1985; Halpern, 2013), which meta-analyses and large-sample studies have estimated as being medium to large (Linn and Petersen, 1985; Voyer et al., 1995; Peters et al., 2007; Hyde, 2014). Conversely to what has been observed for math, gender differences in MRT are observed in all countries (Silverman et al., 2007) and their size do not seem to have declined over time (Masters and Sanders, 1993).

Given their high replicability, males-females' differences in MRT performance have been traditionally regarded as "sex differences" in visuospatial competence that arise from brain specializations imposed by the organizing actions of testosterone during prenatal development (Grimshaw et al., 1995; Baron-Cohen, 2004; Kempel et al., 2005; Peters et al., 2007; Vuoksimaa et al., 2010) and/or from the sexual division of labor in human early evolutionary history (Silverman and Eals, 1992; Silverman et al., 2007). However, despite its popularity both inside and outside the scientific realm, the empirical evidence that supports these views is far from conclusive (Fausto-Sterling, 2003; Jones et al., 2003; Jordan-Young, 2010). Indeed, there is a poor correlation between visuospatial abilities and the indirect indices of prenatal testosterone exposure (Puts et al., 2008) and the "sex differences" regularly observed in this cognitive domain are moderated by subjects' age (Geiser et al., 2008; Titze et al., 2010), experience and training (Uttal et al., 2012) as well as by task-related factors [e.g., time constrains (Voyer, 2011; Maeda and Yoon, 2013; kinds of stimuli (Alexander and Evardone, 2008; Ruthsatz et al., 2017)]. Furthermore, the biological and socio-cultural factors traditionally assigned to sex and gender are irremediably entangled and, in practice, it is not possible to separate their relative contribution to males and females' behaviors as they form a complex set of intertwined influences, referred to as sex/gender (Fausto-Sterling, 2003; Kaiser et al., 2009; Springer et al., 2012). Accordingly, the study of behavioral and cognitive similarities, and the differences between females and males, require more complex and integrative formulations than those provided by traditional categorical divides (e.g., male *vs.* female; biological *vs.* social, etc.), and should incorporate the interactions among predisposing, experiential and situational variables (Jordan-Young and Rumiati, 2012; Springer et al., 2012; Rippon et al., 2014).

In line with this, accumulated evidence indicates that factors traditionally assigned to "gender" might boost the differences in MRT performance ordinarily attributed to "sex." Indeed, it is well known that stereotypic beliefs about cognitive female-male differences can exert long-term effects on the acquisition of both interests and skills (Eccles, 1987; Bussey and Bandura, 1999), but may also have more immediate effects by affecting performance when situationally activated. Thus ever since childhood, self- or others' endorsement of commonly held stereotypic beliefs and implicit associations about genders (e.g., "science-male"; Nosek et al., 2002) reduce female performance in cognitive domains culturally viewed as "masculine" (e.g., math; Ambady et al., 2001; Beilock et al., 2010; Cvencek et al., 2011), and dwindle their interest in pursuing STEM-related studies and professions (Schmader et al., 2004; Kiefer and Sekaquaptewa, 2007a; Watt et al., 2012; Wang and Degol, 2013; Ertl et al., 2017).

Cognitive performance may also be affected by mere awareness of, rather than belief in, stereotypes of the different abilities of targeted groups of persons. Thus when situational variables implicitly or explicitly activate stereotypes, they might induce a so-called 'stereotype threat' in the negatively stereotyped group members, and promote a reduction in their confidence and cognitive performance in those tasks perceived as being relevant to the activated stereotype (Steele and Aronson, 1995; Maass and Cadinu, 2003; Pennington et al., 2016). Accordingly, several studies have shown that the situational cues (e.g., received task instructions) that explicitly state or implicitly activate gender-related stereotypes reduce females' performance in experimental tasks and tests measuring visuospatial abilities (McGlone and Aronson, 2006; Moè and Pazzaglia, 2006; Campbell and Collaer, 2009; Hausmann et al., 2009; Heil et al., 2012; Neuburger et al., 2015). However by encouraging downward social comparisons with a denigrated outgroup, the same situational conditions to promote stereotype reactivation might boost self-confidence and performance in non-negatively stereotyped groups (Blanton et al., 1999; Walton and Cohen, 2003). Accordingly, the explicit or implicit activation of stereotypes on the allegedly different visuospatial abilities of males and females also results in increased male performance in MRT (Moè and Pazzaglia, 2006; Campbell and Collaer, 2009; Hausmann et al., 2009), and in other cognitive domains ordinarily perceived as "masculine" (e.g., math; Kiefer and Sekaquaptewa, 2007b).

Although these and other studies clearly establish that endorsement, implicit interiorization or situational activation of gender-related stereotypes might promote opposite effects in males and females' performance in math and visuospatial tasks, less is known about the individual variables that can moderate these effects (Maass and Cadinu, 2003). This is partly due to the generalized experimental treatment of females and males as being two distinct, but internally, homogenous groups and is also owing to focalization on average-based comparisons. Therefore, in the present study, we decided to compare subgroups of females and males with presumably different degrees of visuospatial abilities (STEM-Males ≥ STEM-Females > HUM-Males ≥ HUM-Females) and stereotypic gender-science beliefs/associations (STEM-Males = HUM-Females > HUM-Males > STEM-Females; see Nosek and Smyth, 2011) in a mixed design that allowed us to establish statistical relationships within, between and across groups.

More specifically, we assessed the relationship between the implicit and explicit gender-science biases held by a single cohort of female and male students of STEM and humanities' majors and their MRT performance after receiving stereotype-congruent ("males will do better"), stereotype-incongruent ("females will do better") or stereotype-nullifying ("no gender differences are expected") experimental instructions. After taking into account the results of previous studies, we hypothesized that 3DMRT performance should relate to the interactive effects between the academic trajectory (STEM *vs.* humanities) and situational variables (received instructions) rather than their raw categorization as females or males. In this way, by reactivating preexisting gender-related explicit beliefs/implicit associations, the received instructions should differentially modify 3DMRT performance in each group and promote specific constellations of between-group differences in each experimental condition. These differences were expected to be larger after receiving stereotype-congruent instructions, when task difficulty increased and among participants endorsing stereotypic views of females and males (a more specific hypotheses' formulation is provided in the different subheadings of the Results section). Moreover, correlational and linear-regression analyses were used to specifically explore whether the influence of gender-science biases on the participants' 3DMRT performance was: (1) similar in females and males; (2) similar in STEM and humanities students; (3) similar across the different experimental conditions. Finally, mediation analyses were used to test the *a priori* hypothesis that these gender-related biases influence 3DMRT performance by decreasing/ increasing the participants' self-confidence.

MATERIALS AND METHODS

This study was carried out in accordance with the recommendations of the ethical standards of the American Psychological Association. The protocol was approved by the Ethics Standards Committees of the Universitat Jaume I. In accordance with the Declaration of Helsinki all subjects gave written informed consent prior to participating.

Participants

Participants were university students at the Universitat Jaume I (Spain) who self-volunteered in response to an invitational email. To be included in the study, participants had to meet the following inclusion criteria: (1) to be in their first university year; (2) to maintain a consistent academic specialization in STEM or humanities since the last two high school years. The initial sample comprised 110 subjects, but five subjects were excluded from the statistical analysis due to incomplete reports of relevant demographic data or to violations of the inclusion criteria. Thus, 105 participants were included in this study (see **Table 1** for the sample details), which were subdivided into four groups according to their self-reported gender and college major. Two of these groups, STEM males (STEM-M; $N = 30$) and Humanities females (HUM-F; $N = 25$), had stereotypic gender-major combinations and the other two, STEM females (STEM-F; $N = 28$) and Humanities males (HUM-M; $N = 22$), had non-stereotypic gender-major combinations. All the participants signed informed consent and their collaboration was awarded with €20.

Measures

All the experimental tasks were programmed and presented in individual personal computers using the Millisecond Inquisit software package 4.0 (Millisecond©). The experimental tasks completed by all the participants included in presentation order: a demographic data form (on which participants reported their gender, age and university major), a mental rotation task, the Gender-Science implicit association test (IAT) and a single-item question to assess explicit beliefs on the suitability of females and males for scientific studies/professions.

TABLE 1 | The sample's demographic and academic characteristics.

	Males	Females
Computer sciences	10	8
Engineering	20	20
Total STEM	30	28
Journalism	8	8
Education	5	15
Other humanities studies	9	2
Total HUMANITIES	22	25
Total participants	52	53
Age	19.10 ± 1.20	18.96 ± 1.30

This table presents the number of participants of each gender and major. Male and female participants' ages (mean ± SD) are in the bottom row.

3D Mental Rotation Task (3DMRT)

To construct our 3DMRT task, we used the stimuli set developed and validated by Ganis and Kievit (2015). As in the classical paper-and-pencil mental rotation task designed by Shepard and Metzler (1971), each stimulus displays two abstract figures (a baseline object and a target object) composed of 7–11 cubes, arranged on four arms and connected end-to-end in a sequence. Ganis and Kievit (2015) provided eight different stimuli variations, grouped into two main categories: four "same" stimuli (those at which the baseline and target objects can be made to coincide with each other through a 0°, 50°, 100°, or 150° rotation on the vertical axis) and four "different" stimuli [whenever this is not possible, one figure arm (or more) is flipped]. Thus, by using the different rotation angles of a single figure, this set of stimuli allows the parametric manipulation of task difficulty. Furthermore, since the number of cubes and other characteristics of figures are identical in "different stimuli" and "same stimuli," the task cannot be carried out merely by taking into account the number of cubes in the objects or any other spurious cue.

Our 3DMRT comprised three phases, which correspond to three experimental conditions, each preceded by a different set of instructions (see Procedure). In all these experimental phases, we used six versions (2 categories × 3 rotation angles, 50°, 100°, and 150°) of eight different stimuli across 48 time-restricted trials (duration: 7.5 s; ITI: 0.5 s). These time parameters were the same as those used by Ganis and Kievit (2015) when validating the current stimuli set. Their inclusion was a necessary control to ensure a similar task performance pace for all the participants, which allows administering the necessary instructions before each experimental phase. In each trial, the computer screens displayed a baseline (left) and a target figure (right). The target figure could be a "same" or a "different" rotated (50, 100, or 150°) version of the baseline figure, but both figures had the same number of cubes and arms arrangement in all cases. The participants were asked to respond by pressing the "b" key (masked with a green tag) on their computer keyboard if they decided that the objects in a pair were the same, or by pressing the "n" key (masked with a red tag) if they decided that the two objects differed. Accuracy (number of correct responses) and latency to respond were automatically measured and, at the end of each phase, subjects were asked to provide (by means of a sliding bar of 10 discrete steps) an estimation of the percentage of correct responses achieved. This additional requirement provided an overall measure confidence in task execution, similar to that used by Estes and Felker (2012).

Implicit Association Test

The Implicit Association Test (Greenwald et al., 1998) is commonly used to assess implicit stereotypic associations, such as those which differentially link males to sciences and females to humanities (Nosek et al., 2002; Smyth and Nosek, 2015). For this study, the Gender-Science IAT script provided at http://www.millisecond.com/download/library/ (the Milisecond Test Library) was adapted to and translated into Spanish for this study (see Supplementary Table 1). This provided script implements the standard IAT procedure, which consists of 7 phases.

Phase 1 (Target category sorting training; 20 trials)

Participants are asked to discriminate and classify the target stimuli (male/female names) that appear at the center on the screen into one of the two categories (female/male) displayed in top corners by pressing the left ("E") or the right ("I") key on the computer's keyboard.

Phase 2 (Attribute sorting training; 20 trials)

Participants are asked to similarly classify attribute stimuli (majors) into one of the two categories (humanities/ STEM) displayed in the top corners of the computer's screen using the same keys than in the previous phase.

Phase 3 (Test block. Stereotype consistent target-attribute pairing; 20 trials)

Participants are asked to perform a combined categorization task by responding with the "E" key to both target and attribute stimuli belonging to the categories (female/humanities) placed on the left top corner and with the "I" key to both target and attribute stimuli belonging to the categories (male/ STEM) displayed on the right top corner of the computer screen.

Phase 4 (Test block; Stereotype consistent target-attribute pairing)

This phase is identical to the previous one but consists of 40 trials.

Phase 5 (Target category sorting training; 20 trials)

This phase is identical to phase 1 but the target sides are switched, so participants must classify male names by pressing the "I" key and the female names by pressing the "E" key. Twenty trials.

Phase 6 (Test block. Stereotype inconsistent target-attribute pairing; 20 trials)

This phase is identical to phase 3, but the category-attribute pairs are reversed. Thus, female names and STEM majors share the same response key ("E") whereas the male names and humanities majors are classified by pressing the "I" key.

Phase 7 (Test block. Stereotype inconsistent target-attribute pairing)

This phase is identical to the previous one but consists of 40 trials.

The provided script automatically counterbalances the order presentation of phases 3–4 and 5–6, so half of the participants perform first the test blocks containing stereotype consistent trials and the other half the stereotype inconsistent test blocks. This script also automatically calculates the so-called *d*-scores (Greenwald et al., 2003). *d*-scores are standardized deviation scores that range between +2 and −2, whose interpretation is similar to that of Cohen's *d* statistics. Following the general convention, the IAT protocol used herein were arranged to provide positive *d* values for stereotype-consistent associations (e.g., "science = male/humanities = female") and negative d values for stereotype-inconsistent associations.

Explicit Beliefs

Participants were asked to explicitly declare and quantify their beliefs as to whether males and females differ in their suitability for "scientific tasks." We literally posed this question as "Who is better suited for science?" and the participants provided answers by a sliding bar of 10 discrete steps. Thus, setting the bar at 1 and 10 indicated that males/females were maximally suited for science, respectively (while setting it at 5 indicated no differences in this respect). Individual scores were computed as 5, minus the provided answer. In this way, and similarly to the IAT *d*-scores, positive (1–4) values quantified the presumed differences to favor males and negative (−1 to −4) values quantified the presumed differences to favor females.

Procedure

The experiment was carried out during six different experimental sessions, and each session involved 15–20 participants. As group composition might create a threatening environment for negatively stereotyped groups (Inzlicht and Ben-Zeev, 2000), we matched the participants in each session for gender and academic specialization into four similarly sized groups (see Supplementary Table 2). At the beginning of each session, three female experimenters greeted the participants in the laboratory, and they randomly assigned them to an individual desk equipped with a personal computer. After giving their informed consent, the experimenters asked the participants to fill in the *Demographic data* form and to wait for further instructions.

After all the participants had completed this first step, a senior researcher introduced the 3DMRT task, and informed them that it comprised three successive phases that should be initiated after her explicit instructions. Before starting each phase, and with the help of a video projection system, the researcher explained the task generalities (goal, response keys, etc.) and provided the specific instructions for each experimental condition. Phases were labeled and presented to the participants as "optimized for women," "optimized for men" and "neutral." The experimenter also emphasized that the selection of the stimuli of each phase was in accordance with previous studies in which they proved to be differently processed and resulted in enhanced performance for females or males, or had led to similar results between genders, respectively. These explanations came along with faked figures of brain scans and bar graphs, which displayed such differential results, which also appeared in the written instructions that the participants had to individually read on their computer screens before starting each phase. In order to increase distinguishability between conditions, the stimuli of the "optimized for women," "optimized for men" and "neutral" conditions appeared on a pink, a blue and a white background, respectively (see Supplementary Figure 1). The order of these three experimental conditions was randomized across the six experimental sessions as a strategy to prevent any practice/learning effect (see Supplementary Table 3).

After finishing the 3DMRT task, the same leading researcher introduced the IAT as a word-sorting task by carefully avoiding any reference to gender or gender-related differences and provided the pertinent instructions for its completion. This cautious introduction to the IAT intended to minimize the chance of any carry-over effects from previous experimental phases. The provided instructions, which emphasized responding quickly, but accurately, also came in writing, shown on the individual screen of each participant's computer before starting the IAT.

Finally, participants were instructed to answer a single explicit question to assess their beliefs as to whether males or females are more capacitated for science (see *Explicit Beliefs* in the Measures section). Once they answered this question, participants were thanked and economically rewarded for their participation.

Data and Statistical Analyses

All the data included in the present study are provided as Supplementary Material (Supplementary Data Sheet S1). Data were analyzed using SPSS 23 (IBM Corp.) and PRISM 7.0 (GraphPad Inc.) for Mac OS X. Figures were constructed using PRISM 7.0 GraphPad Inc.).

One-sample Student's *t*-tests were used to evaluate whether or not the explicit beliefs and implicit associations held by each group of participants were significantly different from zero. Between-group differences in these variables as well as in the observed and expected 3DMRT performance were evaluated by design-appropriate ANOVAs, followed by Tuckey HSD *post hoc* comparisons. The relationship between explicit and implicit gender-related biases and observed/expected 3DMRT performance was initially evaluated by means of Pearson's *r* correlation index. However, in a second step, linear-regression-based procedures were used to explore in further detail the relationship between the IAT scores and observed/ expected 3DMRT performance scores. These more fine-grained analyses included: (1) the evaluation of a possible moderating effect of academic specialization on the influence of implicit gender-science associations over the observed and expected 3DMRT performance scores; (2) The evaluation of a possible mediatory role of confidence on the effects of these implicit associations on the observed 3DMRT performance.

RESULTS

Explicit Beliefs and Implicit Associations

H1: The participants, especially those of groups with gender-major stereotypic combinations, will hold explicit beliefs and implicit associations that preferentially link science to males and humanities to females.

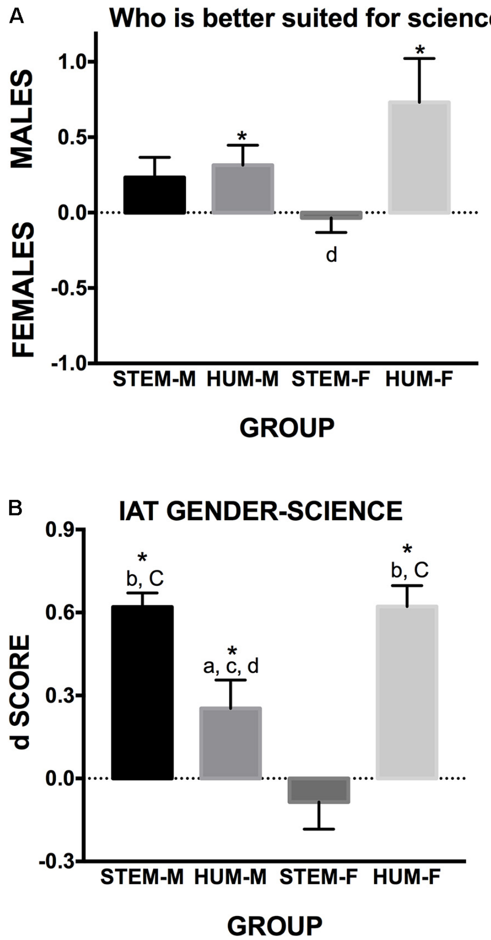


FIGURE 1 | Explicit beliefs and implicit associations. **(A)** Depicts explicit beliefs as to a different suitability of males and females for science studies/professions. **(B)** Illustrates the *d*-scores for IAT Gender-Science. *Significantly different from zero, $p < 0.05$; letters denote statistically significant differences between groups: A different from STEM-Males, B different from HUM-Males, C different from STEM-Females and D different from HUM-Females (capital letters, $p < 0.01$; lowercase letters, $p < 0.05$).

To ascertain whether or not the participants held explicit beliefs as to a differential suitability of females and males for science, one-sample Student's *t*-tests were used. As shown in **Figure 1A**, the size of this belief significantly differed from zero in HUM-Males ($t_{21} = 2.309$, $p < 0.05$) and HUM-Females ($t_{24} = 2.520$, $p < 0.05$), and approached statistical significance in the STEM-Males group ($t_{29} = 1.756$, $p = 0.09$). In a second step, we analyzed the between-group differences by means of one-way ANOVA. The group factor reached statistical significance ($F_{3,101} = 2.86$; $p < 0.05$; $\eta_p^2 = 0.091$) which, as revealed by the Tukey HSD *post hoc* comparisons, was driven solely by a difference between the HUM-Females and the STEM-Females groups ($p < 0.05$, Cohen's $d = 0.739$).

On the other hand, one-sample Student's *t*-test revealed that STEM-Males ($t_{21} = 12.29$, $p < 0.001$), HUM-Males ($t_{21} = 2.46$, $p < 0.05$) and HUM-Females ($t_{24} = 8.24$, $p < 0.001$), but not STEM-Females ($t_{27} = -0.872$, $p = 0.391$), exhibited a significant implicit "male-science/female-humanities" stereotypic association (**Figure 2B**). A one-way ANOVA ($F_{3,101} = 18.12$, $p < 0.001$; $\eta_p^2 = 0.350$) yielded a group effect on the size of this bias. The Tukey HSD *post hoc* comparisons revealed that this bias was larger in groups with gender-major stereotypic combinations than in those with non-stereotypic combinations (STEM-Males > HUM-Males: $p < 0.05$, Cohen's $d = 1.01$; STEM-Males > STEM-Females: $p < 0.01$, Cohen's $d = 1.74$; HUM-Females > HUM-Males: $p < 0.05$; Cohen's $d = 0.89$; HUM-Females > STEM-Females: $p < 0.01$; Cohen's $d = 1.57$). This bias was also larger in HUM-males than in STEM-Females ($p < 0.05$; Cohen's $d = 0.65$).

These results confirmed Hypothesis 1. However, explicit beliefs and implicit associations are two distinct cognitive and poorly correlated ($r = 0.147$, $p = 0.134$) constructs and, only in the second one, the groups with gender-major stereotypic combinations (STEM-Males and HUM-Females) clearly obtained higher bias scores than those with non-stereotypic combinations (STEM-Females and HUM-Males).

3DMRT Observed Performance

H2: The experimental groups will differ in their observed 3DMRT performance (STEM-Males ≥ STEM-Females > HUM-Males ≥ HUM-Females).

H3: The received instructions will differentially modify 3DMRT performance in each group and will hence lead to specific constellations of between-group differences in each experimental phase.

H4: The ability of the experimental instructions to promote gender-related differences in 3DMRT performance will increase with task difficulty.

3DMRT performance was assessed by two main variables: latency to respond and the number of correct responses. As latencies to respond did not differ between groups for any experimental condition (Supplementary Table 4), we do not discuss them further.

Regarding the number of correct responses (**Figure 2A**), a two-way repeated measures ANOVA revealed a significant effect for the group factor ($F_{3,101} = 17.16$, $p < 0.001$; $\eta_p^2 = 0.338$), but not for the condition factor ($F_{2,202} = 3.08$, $p = 0.18$), although the interaction between both factors was significant ($F_{6,202} = 2.98$, $p < 0.001$; $\eta_p^2 = 0.160$). This significant group × condition interaction allowed us to explore how the performance of each group varied across the three experimental conditions (within group comparisons) as well as the between group differences for each one.

Effects of the Received Instructions in Each Experimental Phase (Within Group Comparisons)

The Tukey HSD-based comparisons showed that the performance of the two STEM groups remained largely stable across the three experimental conditions. However, the less conservative Student's *t*-tests for related samples revealed a slight enhancement of STEM-Females' performance for the

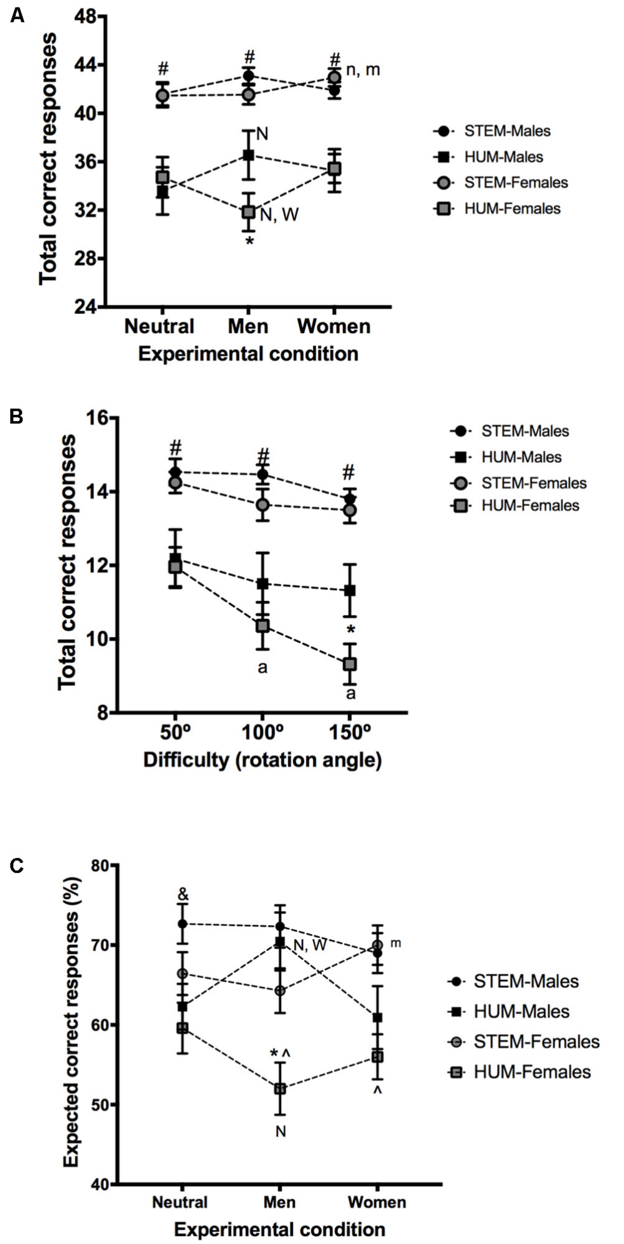


FIGURE 2 | 3DMRT observed and expected performance. **(A)** Depicts the observed performance (mean ± SEM of the number of correct responses) in a mental rotation task run under three experimental conditions ("neutral," "optimized for men" and "optimized for women"). **(B)** Illustrates the relationship between task difficulty (rotation angle) and 3DMRT observed performance (mean ± SEM of correct responses) for each participant's group for the "optimized for men" condition. **(C)** Shows expected performance (mean ± SEM of the participants' expected percentage of correct responses) as a measure of task execution confidence in each experimental phase (see the Measures section for details). Note that in **(A,B)**, the *Y*-axis were adjusted to denote optimal and chance levels performance [#Significantly different from HUM groups; ^Significantly different from STEM groups; *Significantly different from HUM-Males; &Significantly different from HUM-Females; N, n Significantly different from the "neutral" condition ($p < 0.01$ and $p < 0.05$, respectively); W Significantly different from the "optimized for women condition" ($p < 0.01$); m Significantly different from the "optimized for men" condition ($p < 0.05$); a Significantly different from the 50° rotation angle ($p < 0.05$)].

"optimized for women" condition compared to the other two ("neutral": $t_{27} = 2.075$, $p = 0.048$; Cohen's $d = 0.332$; "optimized for males": $t_{27} = 2.655$, $p = 0.013$; Cohen's $d = 0.352$). The same *t*-test based analysis did not reveal any significant variation in the STEM-Males group.

The experimental phase had more pronounced effects on the HUM groups. The intra-group Tukey HSD-based comparisons revealed that HUM-Females' performance dropped for the "optimized for men" condition to become lower than under the "neutral" ($p < 0.05$, Cohen's $d = -0.364$) and the "optimized for women" ($p < 0.05$ Cohen's $d = -0.528$) conditions. Conversely, HUM-Males displayed increased performance under the "optimized for men" condition, which became significantly higher ($p < 0.05$, Cohen's $d = 0.318$) than for the "neutral" condition.

Between-Group Differences in Each Experimental Phase

Under all the experimental conditions STEM-Females and STEM-Males outperformed HUM-Females and HUM-Males (Tukey HSD $p < 0.01$; Cohen's d, ranging from 0.98 to 1.83), but no differences between the two STEM groups were observed. HUM-Males outperformed HUM-Females for the "optimized for men" condition ($p < 0.05$, Cohen's $d = 0.557$), but not for any other experimental condition.

Taken together, these results confirmed Hypotheses 2 and 3 by showing that the 3DMRT performance of STEM-Males, STEM-Females, HUM-Males and HUM-Females differed, and that some of their differences (remarkably those between genders) only arose when receiving gender-loaded task instructions.

Task Difficulty and Gender-Related Differences in 3DMRT Observed Performance

Figure 2B depicts the relationship between task difficulty (rotation angle) and the observed 3DMRT performance for each participants group for the "optimized for men" condition (the only one at which we observed gender-related differences). A two-way repeated measures ANOVA yielded significant group ($F_{3,101} = 17.16$, $p < 0.001$; $\eta_p^2 = 0.338$), rotation angle ($F_{2,202} = 14.90$, $p < 0.001$; $\eta_p^2 = 0.128$) and interaction ($F_{6,202} = 2.29$, $p < 0.05$; $\eta_p^2 = 0.064$) effects.

All the groups showed rotation-related decreases in performance but, as revealed by the Tukey HSD *post hoc* comparisons, this effect was statistically significant only in HUM-Females (50° *vs.* 100° $p < 0.01$, Cohen's $d = 0.557$; 50° vs. 150° $p < 0.001$, Cohen's $d = 0.977$). The between-group comparisons revealed that the two STEM groups outperformed both HUMs groups, regardless of the rotation angle ($p < 0.01$ in all cases). Moreover, when difficulty was maximal (150°) HUM-Females gave fewer correct responses than HUM-Males ($p < 0.05$, Cohen's $d = 0.656$), which hence confirms Hypothesis 4.

Participants' Expected 3DMRT Performance

H5: The received instructions will differentially modify the self-reported expected performance (confidence) in each group, which will then result in specific patterns of between-group differences in each experimental phase.

Figure 2C depicts the participants' expected percentage of correct responses for each experimental condition. A two-way repeated measures ANOVA yielded significant effects for the group factor ($F_{3,101}$ = 6.94, $p < 0.001$; $\eta_p^2 = 0.171$) and for the group × experimental condition interaction ($F_{6,202}$ = 5.36, $p < 0.01$; $\eta_p^2 = 0.137$). This significant group x condition interaction allowed us to explore how this self-reported index of the participants confidence varied within each group across the three experimental conditions as well as the between-group differences in this variable under each experimental condition.

Effects of the Received Instructions in Participants' Expected Performance in Each Experimental Phase (Within-Group Comparisons)

STEM males showed stable levels of expected performance across all the experimental phases. Conversely, all the other groups exhibited significant variations of expected performance depending on the received instructions. Thus HUM-Females' expected performance dropped under the "optimized for men condition" and hence became significantly lower than for the "neutral" condition; Tukey HSD $p < 0.05$, Cohen's d = 0.471). The opposite effect appeared for HUM-Males, with enhanced expected performance under the "optimized for men" condition (Tukey HSD $p < 0.01$, Cohen's d = 0.566 and Tukey HSD $p < 0.001$, Cohen's d = 0.657 compared to the "neutral" and the "optimized for women" conditions, respectively). Finally, the Student's t-tests for related samples, but not the Tukey HSD-based comparisons, revealed a selective increase in STEM-Females' expected performance under the "optimized for women" condition (t_{27} = 2.741, p = 0.01; Cohen's d = 0.261 and t_{27} = 1.780, p = 0.08 compared to the "optimized for men" and "neutral" condition, respectively).

Between-Group Differences in Each Experimental Phase

The Tukey HSD *post hoc* comparisons revealed that HUM-Females had the lowest expected performance (confidence) scores in all the experimental phases. Thus, for the "neutral" condition, only the HUM-Females and the STEM-Males groups significantly differed ($p < 0.01$, Cohen's d = 0.881). For the "optimized for women condition," HUM-Females reported lower expected performance scores than STEM-Males ($p < 0.01$, Cohen's d = 0.932), and also than STEM-Females ($p < 0.01$, Cohen's d = 1.028). Finally, under the "optimized for men" condition, the HUM-Females group differed from all the other groups: STEM-Males ($p < 0.001$, Cohen's d = 1.314), STEM-Females ($p < 0.02$, Cohen's d = 0.789) and HUM-Males ($p < 0.001$, Cohen's d = 1.138).

Taken together, the results of Sections "Effects of the Received Instructions in Participants' Expected Performance in Each Experimental Phase (Within-Group Comparisons)" and "Between-Group Differences in Each Experimental Phase" confirmed hypothesis 5.

Relationships Between Variables

H6: Observed and expected 3DMRT performance will be directly related between them, and will also show gender-dependent correlations with explicit beliefs and implicit associations preferentially linking males and science.

Observed and expected 3DMRT performance directly correlated with one another: ("neutral" condition r = 0.536, $p < 0.000$; "optimized for the men" condition r = 0.596, $p < 0.000$; "optimized for the women" condition r = 0.468, $p < 0.000$). Moreover, these performance-related variables correlated in a gender-dependent manner with the explicit and, more notably, the implicit "gender-science" biases (**Table 2**).

The **Table 2** results confirm Hypothesis 6 and also show that the implicit "male-science/female humanities" associations are more closely related to 3DMRT performance than explicit beliefs. These results also indicate that the same "male-science/female-humanities" association might have opposite functional consequences on female and male 3DMRT performance.

In order to confirm this last observation, we calculated an IAT-derived "influence" index. More specifically, females IAT scores were multiplied by −1, and those of males by 1. This transformation does not change the strength of the implicit Gender-Science associations revealed by the IAT, but slightly modifies the interpretation of the performed correlations, which now provide an index of the expectable "influence" of these implicit gender-related associations on 3DMRT performance rather than a plain measure of their co-variation. As expected, this IAT-derived "influence" index correlated directly with the observed 3DMRT performance ("neutral" condition: r = 0.246, $p < 0.02$; "optimized for males" condition: r = 0.425, $p < 0.000$; "optimized for females" condition: r = 0.319, $p < 0.001$). Similar correlations were found for expected 3DMRT performance ("neutral" condition: r = 0.204, $p < 0.04$; "optimized for males" condition: r = 0.400, $p < 0.000$; "optimized for females" condition: r = 0.277, $p < 0.005$).

By means of this IAT-derived "influence" index, we sought to investigate three additional research questions:

(Q1) Does the implicit "male-science/female humanities" association equally affect 3DMRT observed and expected performance in STEM and humanities students?

To answer this question, we used the regression-based moderation testing procedure proposed by Baron and Kenny (1986). Because the highest correlations between IAT-derived "influence" scores and performance measures were observed for the "optimized for men," we focused on this condition. Regarding observed performance (**Figure 3A**), the slope of the regression line for the HUM group significantly differed from zero ($F_{1,45}$ = 4.47, p = 0.04), unlike that calculated for the STEM group ($F_{1,56}$ = 1.01, p = 0.31). These slopes showed a clear trend toward being significantly different between them (Z = 1.48, p = 0.06). Similarly, as shown in **Figure 3B**, the slope of the regression line for the expected performance of the HUM ($F_{1,45}$ = 16.25, $p < 0.001$), but not that of the STEM groups ($F_{1,56}$ = 1.24, p = 0.26), significantly differed from zero, and yielded a significant inter-groups difference in this case (Z = 2.19,

TABLE 2 | Correlations by gender.

	Explicit		Implicit	
	Females	**Males**	**Females**	**Males**
Correct responses "neutral" condition	***r* = −0.277**	*r* = 0.053	***r* = −0.299**	***r* = 0.330**
	***p* = 0.044**	*p* = 0.710	***p* = 0.030**	***p* = 0.017**
Correct responses "optimized for men" condition	***r* = −0.366**	*r* = −0.030	***r* = −0.433**	***r* = 0.304**
	***p* = 0.007**	*p* = 0.834	***p* = 0.001**	***p* = 0.029**
50°	***r* = −0.311**	*r* = −0.042	***r* = −0.352**	*r* = 0.260
	***p* = 0.023**	*p* = 0.765	***p* = 0.010**	*p* = 0.060
100°	***r* = −0.280**	*r* = −0.146	***r* = −0.356**	*r* = 0.124
	***p* = 0.042**	*p* = 0.302	***p* = 0.009**	*p* = 0.383
150°	***r* = −0.327**	*r* = −0.215	**r = −0.452**	*r* = 0.239
	***p* = 0.017**	*p* = 0.125	***p* = 0.001**	*p* = 0.087
Correct responses "optimized for women" condition	***r* = −0.276**	*r* = −0.153	***r* = −0.470**	***r* = 0.345**
	***p* = 0.045**	*p* = 0.280	***p* < 0.000**	***p* = 0.012**
Expected correct responses "neutral" condition	*r* = −0.139	*r* = 0.005	*r* = −0.018	***r* = 0.303**
	p = 0.322	*p* = 0.996	*p* = 0.898	***p* = 0.029**
Expected correct responses "optimized for men" condition	*r* = −0.253	*r* = −0.022	*r* = −0.260	*r* = 0.199
	p = 0.067	*p* = 0.878	*p* = 0.060	*p* = 0.157
Expected correct responses "optimized for women" condition	***r* = −0.318**	*r* = −0.014	***r* = −0.310**	***r* = 0.337**
	***p* = 0.020**	*p* = 0.920	***p* = 0.024**	***p* = 0.014**

Pearson's r index was used to quantify the correlation between Gender-Science explicit beliefs and implicit associations (IAT d scores) and the different indexes of 3DMRT performance. Statistically significant correlations (p < 0.05) are highlighted in bold.

p = 0.01). These results confirmed that academic specialization is a significant moderator of the IAT Gender-Science "influence" on the observed and expected 3DMRT performance, and revealed that this bias exclusively affected HUM students.

Confirming these results, we found significant correlations between the IAT "influence" scores and the observed and expected performance in HUM, but not in STEM, students (**Table 3**). The same correlational analysis revealed that the "influence" of the implicit "male-science/female-humanities" association on observed and expected 3DMRT performance varied for the different experimental conditions (see below).

(Q2) When do implicit biases affect expected and observed 3DMRT performance?

Several results of the present study were suggestive of a specific effect of the implicit "male-science/female humanities" association on the observed and expected 3DMRT performance of HUM, but not of STEM, students for the "optimized for men" condition. In order to confirm these effects and to explore their specificity, we ran a series of regression analyses.

As shown in **Table 4A**, the IAT "influence" scores (but not gender, age, university major, or explicit beliefs) achieved statistical significance as predictors of the observed 3DMRT performance of HUM students for the "optimized for men" condition. Similarly, the IAT "influence" was the only significant predictor of the expected performance of HUM students under this experimental condition (**Table 4B**). Conversely, neither the IAT "influence," nor gender, age, university major or explicit beliefs achieved statistical significance as predictors of 3DMRT observed or expected performance of HUM students for the "neutral" or the "optimized for women" conditions, nor as predictors of STEM students' performance. Therefore, a specific effect of the implicit "male-science/female humanities" association on HUM students' 3DMRT performance was confirmed.

(Q3) Does expected performance (confidence) mediate the effects of implicit "male-science/female-humanities" association on 3DMRT observed performance?

Previous studies (Steele, 1997; Walton and Cohen, 2003; Estes and Felker, 2012) have suggested that, by reducing confidence, gender stereotypes promote decrease female performance in "male cognitive domains," such as math or mental rotation. Therefore, we sought to explore whether our measure of confidence (expected performance) would mediate the "influence" of implicit gender-science associations on the 3DMRT observed performance. Taking into account all the previous results of our own study, we should solely observe this effect in HUM students (the only ones who displayed gender-related differences) and under the "optimized for men" condition (the only one at which we observed these differences). We tested this *a priori* hypothesis following the regression method for simple mediation described by Baron and Kenny (1986).

As shown in **Figure 4**, when the IAT "influence" and expected performance scores were simultaneously included in a single regression equation, only the second remained a strong predictor ($\beta = 0.674$, $p < 0.000$) of observed performance, while the predictive value of IAT "influence" scores' came very close to zero ($\beta = -0.047$, $p = 0.727$. That is, when the effect of confidence was taken into account, the influence of the

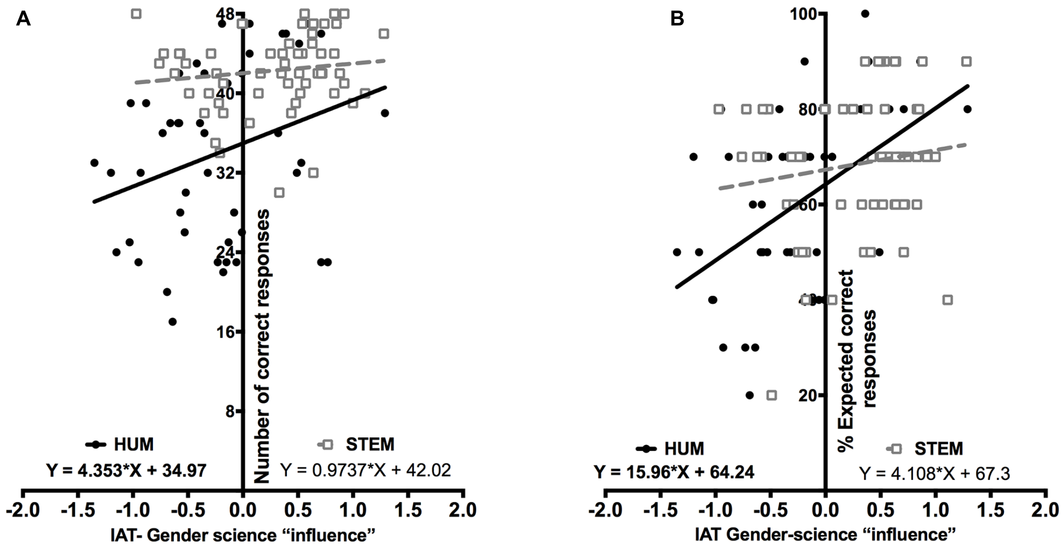


FIGURE 3 | Academic specialization moderates the effects of implicit biases. The data depict individuals' position in a bivariate space defined by the IAT Gender-Science "influence" (*X*-axis) and observed/expected 3DMRT performance (*Y*-axis). To assess the statistical moderation of academic specialization on the effects of implicit associations on MRT performance, separate linear regressions were calculated for HUM (black circles) and STEM (white squares) students. In these analyses, the IAT "influence" scores were used as regressors of the 3DMRT observed **(B)** and expected **(A)** performance scores. Only the slope of the regression equations of the HUM groups significantly differed from zero (in bold, $p < 0.05$ and $p < 0.01$ for observed and expected performance, respectively).

implicit "male-science/female-humanities" association in HUM students' 3DMRT performance was entirely eliminated. The specificity of this mediatory effect was ratified by testing several alternative models with the same regression-based procedure. These additional tests included assessing: (1) the same model in STEM students under the "optimized for men" condition; (2) the same model in HUM students under the "neutral" and "optimized for women" conditions; (3) the reverse model (observed performance mediates expected performance of HUM students under the "optimized for men" condition). As expected, the results of all these tests were negative (for details, see the figures and text included in the Supplementary Materials Image 1 file).

DISCUSSION

Our main results can be summarized as follows: (1) university students hold explicit beliefs and implicit associations that preferentially link science to males and humanities to females; (2) participants' science-related beliefs and associations vary according to an academic specialization (STEM *vs.* humanities) *per* gender interaction; (3) under experimental conditions specifically aimed to nullify or counteract these participants' stereotypic beliefs and associations, academic specialization was the only relevant predictor of 3DMRT performance; (4) when the received experimental instructions reactivated participants' stereotypes on gender-visuospatial abilities, explicit beliefs and, more significantly, gender-science implicit associations, were able to affect 3DMRT performance; (5) changes in confidence mediated these effects and academic specialization moderated them.

Explicit and Implicit Gender-Science Biases

The stereotypical notion of males being more suited for science was explicitly endorsed by the HUM-Males and, to a larger extent, by the HUM-Females groups (**Figure 1A**). As expected from previous studies (Greenwald and Banaji, 1995; Nosek et al., 2002), this explicit belief did not significantly correlate with the implicit Gender-Science associations revealed by the IAT and correlated solely with 3DMRT performance in females, but not in males (**Table 2**). This observation, together with the results of our linear regression-based analyses (see **Table 4**, but also Supplementary Tables 5–7) and those of some previous studies (Hyde et al., 1990; Schmader et al., 2004; Nosek et al., 2009), suggest that explicit gender-science beliefs are less accurate predictors and/or less powerful influencers of cognitive performance than implicit attitudes.

The participants also exhibited an implicit "science-male/humanities-female" association that correlated significantly with the 3DMRT performance in both females and males (**Table 2**). This bias was larger among the gender-major stereotypic combination groups (STEM-Males = HUM-Females; **Figure 1B**) than in those with non-stereotypic combinations (HUM-Males > STEM-Females). This observation is in agreement with cognitive-consistency principles (Nosek et al., 2002), with the results of a massive online survey conducted with college-educated people (Smyth and Nosek, 2015), and also with studies which show that STEM-majoring females hold weaker implicit gender-math stereotypes than both males from the same field and female and male humanities students (Nosek and Smyth, 2011; Smeding, 2012). Taken together, these studies suggest that the implicit "science-

TABLE 3 | Correlations between the IAT "influence" scores and 3DMRT performance in HUM and STEM students.

	HUM	STEM
Correct responses "neutral" condition	$r = -0.044$ $p = 0.771$	$r = 0.171$ $p = 0.200$
Correct responses "optimized for men" condition	**$r = 0.300$** **$p = 0.04$**	$r = 0.121$ $p = 0.367$
50°	$r = 0.157$ $p = 0.802$	$r = 0.123$ $p = 0.357$
100°	$r = 0.107$ $p = 0.473$	$r = 0.123$ $p = 0.341$
150°	**$r = 0.287$** **$p = 0.05$**	$r = 0.127$ $p = 0.341$
Correct responses "optimized for women" condition	$r = 0.072$ $p = 0.630$	$r = 0.123$ $p = 0.357$
Expected correct responses "neutral" condition	$r = 0.038$ $p = 0.802$	$r = 0.005$ $p = 0.971$
Expected correct responses "optimized for men" condition	**$r = 0.515$** **$p < 0.000$**	$r = 0.092$ $p = 0.494$
Expected correct responses "optimized for women" condition	$r = 0.233$ $p = 0.114$	$r = 0.056$ $p = 0.676$

Pearson's r correlation indices and associated p-values are provided. Statistically significant correlations ($p < 0.05$) are in bold.

male/humanities-female" association is highly related to academic/ professional career orientation. Moreover, since our study was conducted in freshman students, our results show that this implicit association is acquired before starting university and suggest that it might influence the students' choice of college major, hence contributing to the asymmetrical representation of girls and boys in STEM and humanities studies.

Interaction Between Implicit Associations and "Neutralizing," "Stereotypic" and "Counter-Stereotypic" Instructions and Its Effects on 3DMRT Performance

When interacting with situational cues (received instructions), the implicit "male-science/female-humanities" association was able to influence 3DMRT performance. As expected, the effects of this implicit bias were substantially smaller when arranging situational cues to nullify latent stereotypes ("neutral" condition) than under the experimental conditions which aimed to activate them (see the correlation values in **Tables 2**, **3**). Indeed, under this "stereotypes' neutralizing condition," STEM-students outperformed HUM-students, and no gender-related differences between these high and low performance groups were found (**Figure 2A**). Accordingly, regression analyses revealed that neither gender nor gender-related explicit beliefs or implicit associations were relevant predictors of 3DMRT performance under this experimental condition, which was significantly related only to academic specialization (see Supplementary Table 5). Thus our results confirm those of previous studies (Quinn and Spencer, 2001; Campbell and Collaer, 2009; Marchand and Taasoobshirazi, 2013), which also observed that stereotype nullification by experimenter-controlled cues suppressed gender-related differences in visuospatial abilities and other cognitive domains for which males' superiority has been traditionally reported. As discussed below, these observations have important theoretical implications in the study and interpretation of "sex-differences" but also practical implications when trying to design educational interventions aimed to increase the representation of girls and women in STEM majors and professions.

The introduction of counter-stereotypic gender-related instructions ("optimized for women condition") did not substantially change the groups' 3DMRT performance. In this atypical situation, STEM-students displayed higher task accuracy than HUM-students but, once again, no gender-related differences were found (**Figure 2A**). Accordingly, linear regression-based analyses revealed that academic specialization, but not participants' gender, gender-related beliefs or implicit associations, became a significant predictor of 3DMRT performance under this experimental condition (see Supplementary Table 7). However, the "optimized for women" and "neutral" conditions were not identical as only the former promoted a slight enhancement of observed and expected performance in STEM-, but not HUM-, females (**Figures 2A,C**). The different reaction of STEM- and HUM-Females to counter-stereotypic instructions could lie in their distinct *a priori* beliefs and implicit associations (**Figure 1**). Thus, lacking any explicit or implicit Gender-Science bias, STEM-Females benefited from females' encouraging instructions, whereas the high and self-demoting biases held by HUM-Females made it impossible for them to benefit from the same positive endorsement. These observations replicate those made in previous studies (Moè and Pazzaglia, 2006; Wraga et al., 2006; Moè, 2009; Heil et al., 2012), which also found that instructions which stressed females'

TABLE 4 | Step-forward linear regression of the **(A)** observed and **(B)** expected performance of HUM students for the "optimized for men" condition.

		Beta	*t*	*p*-value
(A) Observed performance				
Included in the model	Constant	–	26.45	<0.000
	IAT "influence"	0.300	2.11	0.04
Excluded from the regression model	Age	0.194	1.33	0.18
	Gender	−0.107	−0.51	0.60
	University major	−0.182	−1.24	0.21
	Gender-science explicit belief	−0.232	1.65	0.10
	Model summary	*R*	Adjusted R^2	*p*-value
		0.300	0.07	0.04
(B) Expected performance				
Included in the model	Constant	–	25.26	<0.000
	IAT "influence"	0.515	4.02	<0.000
Excluded from the regression model	Age	0.197	1.51	0.13
	Gender	−0.27	−1.50	0.14
	University major	−0.124	−1.24	0.35
	Gender-science explicit belief	−0.178	−1.39	0.16
	Model summary	*R*	Adjusted R^2	*p*-value
		0.515	0.249	<0.000

Separate stepwise forward linear regression analyses were conducted in the HUM and STEM groups to compare the predictive power of the IAT "influence" scores and of other possible predictors on the 3DMRT observed and expected performance at each experimental condition. For these analyses, nominal variables were coded as follows: gender (males = 1, females = 2) and university major (computer sciences = 1, engineering = 2, journalism = 3, education = 4, other humanities' studies = 5). Similar regression analyses were conducted for STEM students, but no variable entered in the model).

superiority in mental rotation tasks increased their performance, and that this increase was more marked for those females who did not sustain *a priori* beliefs about males' visuospatial superiority (Moè and Pazzaglia, 2006). However, in line with some (Moè, 2009; Heil et al., 2012), but not with other (Moè and Pazzaglia, 2006; Wraga et al., 2006) preceding studies, counter-stereotypic instructions did not bring about any change in STEM- or HUM-Males task performance. The reasons why these studies found distinct results remain unclear, but they might be indicative of a relatively weaker capacity of counter-stereotypic instructions to induce 3DMRT performance changes, especially if they result in a threat, and/or if subjects subscribe to the stereotypes contradicted by received instructions.

In contrast, stereotype-congruent instructions resulted in significant gender-related changes in 3DMRT performance. More specifically under the "optimized for men" condition, the 3DMRT accuracy of HUM-Females markedly diminished, but substantially increased in HUM-Males, and hence became significantly different between them and from their own performance under the other two experimental conditions (**Figure 2A**). Thus our results agree with those of previous studies, which have shown that experimental instructions which explicitly state females' inferiority in visuospatial abilities reduce females' performance in mental rotation tasks (Martens et al., 2006; Moè and Pazzaglia, 2006; Wraga et al., 2006; Campbell and Collaer, 2009; Hausmann et al., 2009; Moè, 2009; Heil et al., 2012), but increases males' performance (Moè and Pazzaglia, 2006; Campbell and Collaer, 2009; Hausmann et al., 2009).

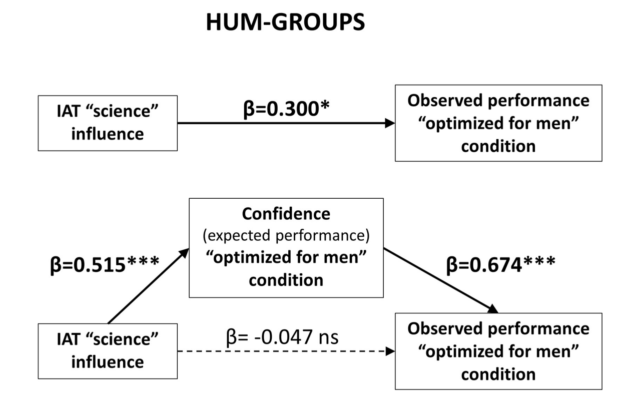


FIGURE 4 | Expected performance (confidence) mediates the effects of implicit gender-science associations. Mediation analysis was performed according to the 3-steps regression method described by Baron and Kenny (1986). First, we confirmed that the IAT "influence" scores predicted 3DMRT observed performance ($\beta = 0.300$, $p < 0.05$). Then we confirmed that these scores also predicted confidence (expected performance; $\beta = 0.515$, $p < 0.000$) and that confidence predicted observed performance ($\beta = 0.650$, $p < 0.000$). Finally, we simultaneously included the IAT "influence" and confidence scores as predictor variables of 3DMRT observed performance in a single regression equation. In this crucial step, only the confidence (expected performance) scores remained as strong predictors ($\beta = 0.674$, $p < 0.000$) of observed performance, while the predictive value of IAT "influence" scores became almost zero ($\beta = -0.047$, $p = 0.727$), hence revealing a near complete mediatory effect of confidence (*$p < 0.05$, ***$p < 0.000$; ns, non-significant).

In line with this, it has been proposed that confidence might underlie between gender differences in 3DMRT performance (Estes and Felker, 2012) as well as instructions-driven performance changes in gender-stereotyped cognitive domains (Steele, 1997; Walton and Cohen, 2003). More specifically, it has been suggested that stereotype reactivation might induce a self-confidence threat that disrupts task performance in the negatively stereotyped group (Schmader et al., 2008), but may induce a self-confidence boost that increases performance in the non-negatively stereotyped group (Blanton et al., 1999; Walton and Cohen, 2003). In agreement with this proposal, we observed that (probably by re-activating previously held stereotypic associations; **Figure 1B** and **Table 3**) the stereotype-congruent instructions of the "optimized for men" condition promoted disparate changes not only in the 3DMRT performance of the HUM-Females and HUM-Males groups (**Figure 2A**), but also in their confidence (**Figure 2C**), and that confidence mediates the influence of implicit associations on 3DMRT observed performance (**Figure 4**).

However, stereotype-congruent instructions do not uniformly affect females or males' performance as academic specialization moderates their effects (**Figure 3**). Accordingly, gender as a binary category did not come over as a significant predictor of 3DMRT performance for the "optimized for men" condition, which was instead mainly predicted from participants' academic specialization (Supplementary Table 6). Moreover, although the IAT "influence" scores were also significant predictors of 3DMRT performance under this experimental condition (Supplementary Table 6), their effects were restricted to HUM students (**Table 4**). Thus, despite having very different implicit Gender-Science associations (**Figure 1B**), the 3DMRT performance of STEM-Females and STEM-Males under the "optimized for men" condition was high, similarly to that observed for the "neutral" and "optimized for women" conditions and was indistinguishable between them (**Figure 2A**). These results, together with those of **Table 3** and Supplementary Figure 2, suggest that academic training or related academic experiences that result in a high level of task performance and/or confidence are able to suppress the influence of the gender-related implicit associations triggered by stereotypic experimental instructions. Our results and conclusions agree with those of a previous study (Hausmann, 2014), which showed that female arts, but not female STEM or male, students, reduced their 3DMRT performance after the reactivation of gender stereotypes. Similarly, gender stereotypes reactivation promotes a reduction of math performance of female psychology, but not of female engineering, students (Crisp et al., 2009).

Limitations and Implications

Under the different experimental conditions of the present study, academic specialization, but not the participants' gender, was the most relevant variable to predict 3DMRT performance. Our results also reveal that the within-gender differences that derived from academic specialization (STEM vs. HUM) are larger than those observed between genders. Indeed, we only observed between-gender differences in 3DMRT performance in HUM, but not STEM, students, and these differences solely emerged in response to stereotype-reactivating experimental instructions. These findings contrast with the common belief that males have better spatial abilities than females (Devlin, 2001; Blanton et al., 2002) and with the ordinarily reported higher performance of males in mental rotation tasks in studies that specifically aim to identify "sex differences" (Linn and Petersen, 1985; Silverman and Eals, 1992; Grimshaw et al., 1995; Kempel et al., 2005; Peters et al., 2007; Silverman et al., 2007; Vuoksimaa et al., 2010; Halpern, 2013; Hyde, 2014; National Science Foundation, 2015).

At this respect, it should be noted that while we used a chronometric two-choice task, most research into sex differences in mental rotation use the pen-and-paper Mental Rotations Test (MRT) developed by Vandenberg and Kuse (1978). The MRT tends to produce larger sex differences (average d = 1) than chronometric tasks (average d = 0.3) and many studies using this second kind of procedures did not observe between genders differences (Voyer, 2011). Therefore, it might be argued that we did not observe the regularly reported gender differences because we did not use the "right" task for this. However, mental rotation chronometric tasks are as valid as psychometric tests (Voyer et al., 2006) and the MRT should not be considered as a benchmark when assessing and comparing the mental rotation abilities of males and females. In fact, the MRT does not seem to provide a pure measure of mental rotation abilities, and its singular ability to detect between gender differences might be related to the specific aspects of this test rather than to the responders' visuospatial abilities (Kerkman et al., 2000; Voyer and Hou, 2006; Hooven et al., 2008; Bors and Vigneau, 2011). Thus, while the results obtained with either chronometric or psychometric MRTs may differ and have a limited generalizability between each other, the use of a chronometric task does not limit the validity of the results observed in the present study.

Yet, it might be argued that, because gender differences observed in mental rotation chronometric tests are small (average d = 0.3), our study may lack the necessary statistical power to detect them. Therefore, the results of the present study should be interpreted with caution and replicated in a larger sample of participants. However, it should be noted that, although some small effects might have failed to reach statistical significance, these power limitations did not preclude by identifying the effects of academic specialization and stereotype-reactivating experimental instructions. This hence reveals that 3DMRT performance (at least as measured in our chronometric task) is much more dependent on these factors than on the participants' gender. Moreover, it should be also noted that the present study was not primarily intended to assess overall gender differences in visuospatial abilities but to identify a possible relationship between gender-science stereotypes and the participants performance in a specific 3DMRT task and that our study has power enough to detect even small to moderate correlations ($\approx\rho = 0.26$ if involving all participants and $\approx\rho = 0.32$ for any two subgroups of participants).

In this regard, it should also be emphasized that our study did not fail to identify between-gender differences in MRT performance but showed that these differences seem to emerge under particular testing conditions and involve some, but not all, male and female participants. Yet, precisely because gender differences in 3DMRT performance depend on task and respondents' characteristics (Sharps et al., 1994; Levine et al., 2005; Jansen-Osmann and Heil, 2007; Alexander and Evardone, 2008; Lippa et al., 2010), it might be concluded that the "sex differences" in mental rotation abilities do not arise from "sex" *per se,* but from its interaction with biographical (e.g., academic specialization) and situational variables (e.g., received instructions). In this way, our results also argue against the attempt to explain the scarce representation of women in STEM studies and professions as a result of "hardly-wired" sex differences in visuospatial and math abilities. On the contrary, our results suggest that gender socialization and stereotypes might have a larger impact in situational performance in these cognitive domains and, thereby, in shaping the perceived competence and motivation to pursuit STEM careers. These conclusions fall in line with those of other studies that have indicated an important role of females and males' differential preferences, experiences and activities in the development of their visuospatial abilities (Flaherty, 2005; Feng et al., 2007; Sander et al., 2010; Nazareth et al., 2013; Moè, 2016). Moreover, the results and conclusions of our study also align with recent proposals which have suggested that in brain and behavior-related studies, sex and gender or, more properly, their composite resultant (sex/gender), should be considered a source of differential interactive effects with other variables rather than a binary-independent factor (Springer et al., 2012; Rippon et al., 2014; Joel and Fausto-sterling, 2016).

CONCLUSION

We observed that experimental instructions might reactivate implicit biases and promote increased/decreased 3DMRT performance, but training and/or other experiences related to academic specialization moderate these effects. In this way, the present study provides evidence about when (after receiving stereotype-congruent, but not stereotype-incongruent or stereotype-nullifying instructions), how (by increasing or reducing confidence) and who (HUM, but not STEM students) might be influenced by implicit gender-science associations while performing a chronometric mental rotation task. Our results also highlight that within-gender differences might be as large as, or even bigger than, those observed between genders and, therefore, that males and females are not two uniform populations (neither in their mental rotation abilities, nor in their reaction to gender-stereotypes reactivation). Therefore, stating that "males have higher visuospatial abilities than females" is a misleading simplification that might contribute to perpetuate stereotypes. Those stereotypes and their detrimental impact on individual performance might progressively undermine the confidence and self-perceived competence of girls in cognitive domains ordinarily labeled as "masculine," hence reducing their interest in pursuing STEM-related academic and professional careers. However, as also suggested by some results of the present study (STEM/HUM females comparison), training and positive academic experiences in those cognitive domains promote resilience against pervasive gender-science stereotypes and provide a promising avenue when trying to enhance the number of women enrolled in STEM majors.

AUTHOR CONTRIBUTIONS

All authors contributed to the data collection. NA programmed the experimental tasks. ÁC-G and NS contributed to manuscript editing. CF and CS-S took part in each and every step of the experiment design, implementation, data analysis/interpretation, and manuscript writing. All authors have read and approved the final manuscript.

FUNDING

This research was supported by grant P1-1B2014-15 provided by Universitat Jaume I and PSI2015-67285-R provided by Ministerio de Educación, Cultura y Deporte to CF and by grant UJI B2017-05, awarded to CS-S.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01261/full#supplementary-material

REFERENCES

Alexander, G. M., and Evardone, M. (2008). Blocks and bodies: sex differences in a novel version of the Mental Rotations Test. *Horm. Behav.* 53, 177–184. doi: 10.1016/j.yhbeh.2007.09.014

Ambady, N., Shih, M., Kim, A., and Pittinsky, T. L. (2001). Stereotype susceptibility in children: effects of identity activation on quantitative performance. *Psychol. Sci.* 12, 385–390. doi: 10.1111/1467-9280.00371

Barbie. (1994). "Responses of threat vs. challenge mediated arousal to stereotypes alleging intellectual inferiority," in *Gender Differences in Mathematics*, eds A. Gallagher and J. C. Kaufman (Cambridge, MA: Cambridge University Press).

Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social the moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers.* 51, 1173–1182. doi: 10.1037/0022-3514.51.6.1173

Baron-Cohen, S. (2004). *The Essential Difference*. London: Penguin Books.

Beilock, S. L., Gunderson, E. A., Ramirez, G., and Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proc. Natl. Acad. Sci. U.S. A.* 107, 1860–1863. doi: 10.1073/pnas.0910967107

Benbow, C. P., Lubinski, D., Shea, D. L., and Eftekhari-Sanjani, H. (2000). Sex differences in mathematical reasoning ability at age 13: their status 20 years later. *Psychol. Sci.* 11, 474–480. doi: 10.1111/1467-9280.00291

Benbow, C. P., and Stanley, J. C. (1980). Sex differences in mathematical ability: fact or artifact? *Science* 210, 1262–1264. doi: 10.1126/science.7434028

Benbow, C. P., and Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: more facts. *Science* 222, 1029–1031. doi: 10.1126/science.6648516

Ben-Zeev, T., Carrasquillo, C. M., Ching, A. M. L., Kliengklom, T. J., McDonald, K. L., Newhall, D. C., et al. (2005). "Math is hard!" (Barbie, 1994). Responses of threat vs. challenge-mediated arousal to stereotypes alleging intellectual inferiority"," in *Gender Differences in Mathematics*, eds A. M. Gallagher and J. C. Kaufman (New York, NY: Cambridge University Press), 48–72.

Blanton, H., Buunk, B. P., Gibbons, F. X., and Kuyper, H. (1999). When better-than-others compare upward: choice of comparison and comparative evaluation as independent predictors of academic performance. *J. Pers. Soc. Psychol.* 76, 420–430. doi: 10.1037/0022-3514.76.3.420

Blanton, H., Christie, C., and Dye, M. (2002). Social identity versus reference frame comparisons: the moderating role of stereotype endorsement 1. *J. Exp. Soc. Psychol.* 38, 253–267. doi: 10.1006/jesp.2001.1510

Bors, D. A., and Vigneau, F. (2011). Sex differences on the mental rotation test: an analysis of item types. *Learn. Individ. Dif.* 21, 129–132. doi: 10.1016/j.lindif.2010.09.014

Bussey, K., and Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychol. Rev.* 106, 676–713. doi: 10.1037/0033-295X.106.4.676

Campbell, S. M., and Collaer, M. L. (2009). Stereotype threat and gender differences in performance on a novel visuospatial task. *Psychol. Women Q.* 33, 437–444. doi: 10.1111/j.1471-6402.2009.01521.x

Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412

Crisp, R. J., Bache, L. M., and Maitner, A. T. (2009). Dynamics of social comparison in counter-stereotypic domains: stereotype boost, not stereotype threat, for women engineering majors. *Soc. Influ.* 4, 171–184. doi: 10.1080/15534510802607953

Cvencek, D., Meltzoff, A. N., and Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x

Devlin, A. S. (2001). *Mind and Maze. Spatial Cognition and Environmental Behavior*. Westport, CT: Greenwood Publishing Group, 328. doi: 10.1002/acp.938

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychol. Women Q.* 11, 135–172. doi: 10.1111/j.1471-6402.1987.tb00781.x

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bulletin* 136, 103–127. doi: 10.1037/a0018053

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Estes, Z., and Felker, S. (2012). Confidence mediates the sex difference in mental rotation performance. *Arch. Sex. Behav.* 41, 557–570. doi: 10.1007/s10508-011-9875-5

European Commission (2014). *Cape Town: Carlos Moedas. Reaching Gender Equality in Science, Technology, Engineering and Mathematics*. Available at: https://ec.europa.eu/commission/2014-2019/moedas/announcements/reaching-gender-equality-science-technology-engineering-and-mathematics_en

European Comission (2016). *SHE Figures 2015*. Available at: https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-final.pdf

Fausto-Sterling, A. (2003). "The problem with sex/gender and nature/nurture," in *Debating Biology*, eds S. J. Williams, L. Birke, and G. Bendelow (London: Routledge), 123–132. doi: 10.4324/9780203987681

Feng, J., Spence, I., and Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychol. Sci.* 18, 850–855. doi: 10.1111/j.1467-9280.2007.01990.x

Flaherty, M. (2005). Gender differences in mental rotation ability in three cultures: Ireland, Ecuador and Japan. *Psychologia* 48, 31–38. doi: 10.2117/psysoc.2005.31

Ganis, G., and Kievit, R. (2015). A new set of three-dimensional shapes for investigating mental rotation processes: validation data and stimulus set. *J. Open Psychol. Data* 3:e3. doi: 10.5334/jopd.ai

Geiser, C., Lehmann, W., and Eid, M. (2008). A note on sex differences in mental rotation in different age groups. *Intelligence* 36, 556–563. doi: 10.1016/j.intell.2007.12.003

Glennon, V. J., and Callahan, L. G. (1968). *Elementary School Mathematics: A Guide to Current Research*. Washington, DC: Association for supervision and Curriculum Development.

Greenwald, A., and Banaji, M. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychol. Rev.* 102, 4–27. doi: 10.1037/0033-295X.102.1.4

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the implicit association test: an improved scoring algorithm. *J. Pers. Soc. Psychol.* 85, 197–216. doi: 10.1037/0022-3514.85.2.197

Grimshaw, G. M., Sitarenios, G., and Finegan, J. A. (1995). Mental rotation at 7 years: relations with prenatal testosterone levels and spatial play experiences. *Brain Cogn.* 29, 85–100. doi: 10.1006/brcg.1995.1269

Guiso, L., Monte, F., Sapienza, P., and Zingales, L. (2008). Culture, gender, and math. *Science* 320, 1164–1165. doi: 10.1126/science.1154094

Halpern, D. F. (2013). *Sex Differences in Cognitive Abilities*, 4th Edn. Hove: Psychology Press. doi: 10.4324/9781410605290

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., and Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychol. Sci. Public Interest* 8, 1–51. doi: 10.1111/j.1529-1006.2007.00032.x

Hausmann, M. (2014). Arts versus science - Academic background implicitly activates gender stereotypes on cognitive abilities with threat raising men's (but lowering women's) performance. *Intelligence* 46, 235–245. doi: 10.1016/j.intell.2014.07.004

Hausmann, M., Schoofs, D., Rosenthal, H. E. S., and Jordan, K. (2009). Interactive effects of sex hormones and gender stereotypes on cognitive sex differences-A psychobiosocial approach. *Psychoneuroendocrinology* 34, 389–401. doi: 10.1016/j.psyneuen.2008.09.019

Heil, M., Jansen, P., Quaiser-Pohl, C., and Neuburger, S. (2012). Gender-specific effects of artificially induced gender beliefs in mental rotation. *Learn. Individ. Dif.* 22, 350–353. doi: 10.1016/j.lindif.2012.01.004

Hill, C., Corbett, C., and St Rose, A. (2010). *Why so Few? Women in Science, Technology, Engineering, and Mathematics title*. Washington, DC: American Association of University Women.

Hooven, C. K., Chabris, C. F., Ellison, P. T., Kievit, R. A., and Kosslyn, S. M. (2008). *The Sex Difference on Mental Rotation Tests Is Not Necessarily a Difference in Mental Rotation Ability*. New York, NY: Christopher F. Chabris.

Hyde, J. S. (2014). Gender similarities and differences. *Annu. Rev. Psychol.* 65, 373–398. doi: 10.1146/annurev-psych-010213-115057

Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., and Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: a Meta-Analysis. *Psychol. Women Q.* 14, 299–324. doi: 10.1111/j.1471-6402.1990.tb00022.x

Hyde, J. S., and Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8801–8807. doi: 10.1073/pnas.0901265106

Inzlicht, M., and Ben-Zeev, T. (2000). A threatening intellectual environment: why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychol. Sci.* 11, 365–371. doi: 10.1111/1467-9280.00272

Jansen-Osmann, P., and Heil, M. (2007). Suitable stimuli to obtain (no) gender differences in the speed of cognitive processes involved in mental rotation. *Brain Cogn.* 64, 217–227. doi: 10.1016/j.bandc.2007.03.002

Joel, D., and Fausto-sterling, A. (2016). Beyond sex differences: new approaches for thinking about variation in brain structure and function. *Philos. Trans. R. Soc. B* 371:20150451. doi: 10.1098/rstb.2015.0451

Jones, C. M., Braithwaite, V. A., and Healy, S. D. (2003). The evolution of sex differences in spatial ability. *Behav. Neurosci.* 117, 403–411. doi: 10.1037/0735-7044.117.3.403

Jordan-Young, R., and Rumiati, R. I. (2012). Hardwired for sexism? Approaches to sex/gender in neuroscience. *Neuroethics* 5, 305–315. doi: 10.1007/s12152-011-9134-4

Jordan-Young, R. M. (2010). *Brain Storm The Flaws in the Science of Sex Differences*. London: Harvard University Press.

Kaiser, A., Haller, S., Schmitz, S., and Nitsch, C. (2009). On sex/gender related similarities and differences in fMRI language research. *Brain Res. Rev.* 61, 49–59. doi: 10.1016/j.brainresrev.2009.03.005

Kempel, P., Gohlke, B., Klempau, J., Zinsberger, P., Reuter, M., and Hennig, J. (2005). Second-to-fourth digit length, testosterone and spatial ability. *Intelligence* 33, 215–230. doi: 10.1016/j.intell.2004.11.004

Kerkman, D. D., Wise, J. C., and Harwood, E. A. (2000). Impossible "mental rotation" problems. *Learn. Individ. Dif.* 12, 253–269. doi: 10.1016/S1041-6080(01)00039-5

Kiefer, A. K., and Sekaquaptewa, D. (2007a). Implicit stereotypes and women's math performance: how implicit gender-math stereotypes influence women's susceptibility to stereotype threat. *J. Exp. Soc. Psychol.* 43, 825–832. doi: 10.1016/j.jesp.2006.08.004

Kiefer, A. K., and Sekaquaptewa, D. (2007b). Implicit stereotypes, gender identification, and math-related outcomes: a prospective study of female college students: research report. *Psychol. Sci.* 18, 13–18. doi: 10.1111/j.1467-9280.2007.01841.x

Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., and Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychol. Sci.* 16, 841–845. doi: 10.1111/j.1467-9280.2005.01623.x

Lindberg, S. M., Hyde, J. S., Petersen, J. L., and Linn, M. C. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psychol. Bull.* 136, 1123–1135. doi: 10.1037/a0021276

Linn, M. C., and Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Dev.* 56, 1479–1498. doi: 10.2307/1130467

Lippa, R. A., Collaer, M. L., and Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Arch. Sex. Behav.* 39, 990–997. doi: 10.1007/s10508-008-9460-8

Maass, A., and Cadinu, M. (2003). Stereotype threat: when minority members underperform. *Eur. Rev. Soc. Psychol.* 14, 243–275. doi: 10.1080/10463280340000072

Maccoby, E. E., and Jacklin, C. N. (1974). *The Psychology of Sex Differences*. Stanford, CA: Stanford University Press.

Maceira, H. M. (2017). Economic benefits of gender equality in the EU. *Intereconomics* 52, 178–183. doi: 10.1007/s10272-017-0669-4

Machin, S., and Pekkarinen, T. (2008). Global sex differences in test score variability. *Science* 322, 1331–1332. doi: 10.1126/science.1162573

Maeda, Y., and Yoon, S. Y. (2013). A meta-analysis on gender differences in mental rotation ability measured by the Purdue spatial visualization tests: visualization of rotations (PSVT:R). *Educ. Psychol. Rev.* 25, 69–94. doi: 10.1007/s10648-012-9215-x

Marchand, G. C., and Taasoobshirazi, G. (2013). Stereotype threat and women's performance in physics. *Int. J. Sci. Educ.* 35, 3050–3061. doi: 10.1080/09500693.2012.683461

Martens, A., Johns, M., Greenberg, J., and Schimel, J. (2006). Combating stereotype threat: the effect of self-affirmation on women's intellectual performance. *J. Exp. Soc. Psychol.* 42, 236–243. doi: 10.1016/j.jesp.2005.04.010

Masters, M. S., and Sanders, B. (1993). Is the gender difference in mental rotation disappearing? *Behav. Genet.* 23, 337–341. doi: 10.1007/BF01067434

McGlone, M. S., and Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. *J. Appl. Dev. Psychol.* 27, 486–493. doi: 10.1016/j.appdev.2006.06.003

Moè, A. (2009). Are males always better than females in mental rotation? Exploring a gender belief explanation. *Learn. Individ. Dif.* 19, 21–27. doi: 10.1016/j.lindif.2008.02.002

Moè, A. (2016). Does experience with spatial school subjects favour girls' mental rotation performance? *Learn. Individ. Dif.* 47, 11–16. doi: 10.1016/j.lindif.2015.12.007

Moè, A., and Pazzaglia, F. (2006). Following the instructions!. Effects of gender beliefs in mental rotation. *Learn. Individ. Dif.* 16, 369–377. doi: 10.1016/j.lindif.2007.01.002

National Association of Colleges and Employers (2009). *Salary Survey*. Houston, TX: NACE.

National Science Board (2010). *Science and Engineering Indicators 2010 (NSB 10-01)*. Arlington, VA: National Science Foundation.

National Science Foundation (2015). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2015*. Arlington, VA: Special Report NSF, 17–310.

Nazareth, A., Herrera, A., and Pruden, S. M. (2013). Explaining sex differences in mental rotation: role of spatial activity experience. *Cogn. Process.* 14, 201–204. doi: 10.1007/s10339-013-0542-8

Neuburger, S., Ruthsatz, V., Jansen, P., and Quaiser-Pohl, C. (2015). Can girls think spatially? Influence of implicit gender stereotype activation and rotational axis on fourth graders' mental-rotation performance. *Learn. Individ. Dif.* 37, 169–175. doi: 10.1016/j.lindif.2014.09.003

Norland, M., Mora, T., and Kostschwar, B. (2016). *Is Gender Diversity Profitable? Evidence from a Global Survey*. Available at: https://ssrn.com/abstract=2729348

Nosek, B. A., Banaji, M., and Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dyn.* 6, 101–115. doi: 10.1037/1089-2699.6.1.101

Nosek, B. A., and Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *Am. Educ. Res. J.* 48, 1125–1156. doi: 10.3102/0002831211410683

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10593–10597. doi: 10.1073/pnas.0809921106

OECD (2016). *PISA 2015 Results: Excellence and Equity in Education, PISA*, Vol. I. Paris: OECD Publishing. Available at: https://www.mecd.gob.es/inee/dam/jcr:54fd088e-f421-49c7-8ee2-852aff57682f/pisa2015-results-eng-vol1.pdf

Pease, A., and Pease, B. (2004). *Why Men Don't Listen and Women Can't Read Maps: How We're Different and What to Do About It*. Emmaus, PA: Rodale, INC.

Pennington, C. R., Heim, D., Levy, A. R., and Larkin, D. T. (2016). Twenty years of stereotype threat research: a review of psychological mediators. *PLoS One* 11:e0146487. doi: 10.1371/journal.pone.0146487

Peters, M., Manning, J. T., and Reimers, S. (2007). The effects of sex, sexual orientation, and digit ratio (2D:4D) on mental rotation performance. *Arch. Sex. Behav.* 36, 251–260. doi: 10.1007/s10508-006-9166-8

Puts, D. A., McDaniel, M. A., Jordan, C. L., and Breedlove, S. M. (2008). Spatial ability and prenatal androgens: meta-analyses of congenital adrenal hyperplasia and digit ratio (2D:4D) studies. *Arch. Sex. Behav.* 37, 100–111. doi: 10.1007/s10508-007-9271-3

Quinn, D. M., and Spencer, S. J. (2001). The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *J. Soc. Issues* 57, 55–71. doi: 10.1111/0022-4537.00201

Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. *PLoS One* 7:e39904. doi: 10.1371/journal.pone.0039904

Rippon, G., Jordan-Young, R., Kaiser, A., and Fine, C. (2014). Recommendations for sex/gender neuroimaging research: key principles and implications for research design, analysis, and interpretation. *Front. Hum. Neurosci.* 8:650. doi: 10.3389/fnhum.2014.00650

Ruthsatz, V., Neuburger, S., Rahe, M., Jansen, P., and Quaiser-Pohl, C. (2017). The gender effect in 3D-Mental-rotation performance with familiar and gender-stereotyped objects – a study with elementary school children. *J. Cogn. Psychol.* 29, 717–730. doi: 10.1080/20445911.2017.1312689

Sander, E., Quaiser-Pohl, C., and Stigler, C. (2010). Factors influencing the development of mental-rotation ability the role of socio-cultural background. *Int. J. Dev. Sci.* 4, 18–30. doi: 10.3233/DEV-2010-4102

Schmader, T., Johns, M., and Barquissau, M. (2004). The costs of accepting gender differences: the role of stereotype endorsement in women's experience in the math domain. *Sex Roles* 50, 835–850. doi: 10.1023/B:SERS.0000029101.74557.a0

Schmader, T., Johns, M., and Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychol. Rev.* 115, 336–356. doi: 10.1037/0033-295X.115.2.336.An

Sharps, M. J., Price, J. L., and Williams, J. K. (1994). Spatial cognition and gender: instructional and stimulus influences on mental image rotation performance. *Psychol. Women Q.* 18, 413–425. doi: 10.1111/j.1471-6402.1994.tb00464.x

Shepard, R. N., and Metzler, B. (1971). Mental rotation of three-dimensional objects. *Science* 171, 702–703. doi: 10.1126/science.171.3972.701

Silverman, I., Choi, J., and Peters, M. (2007). The Hunter-gatherer theory of sex differences in spatial abilities: data from 40 countries. *Arch. Sex. Behav.* 36, 261–268. doi: 10.1007/s10508-006-9168-6

Silverman, I., and Eals, M. (1992). "Sex differences in spatial abilities: evolutionary theory and data," in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, eds J. Barkow and L. Cosmides (New York, NY: Oxford University Press), 533–549.

Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): an investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles* 67, 617–629. doi: 10.1007/s11199-012-0209-4

Smyth, F. L., and Nosek, B. A. (2015). On the gender-science stereotypes held by Scientists: explicit accord with gender-ratios, implicit accord with scientific identity. *Front. Psychol.* 6:415. doi: 10.3389/fpsyg.2015.00415

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *Am. Psychol.* 60, 950–958. doi: 10.1037/0003-066X.60.9.950

Springer, K. W., Mager Stellman, J., and Jordan-Young, R. M. (2012). Beyond a catalogue of differences: a theoretical frame and good practice guidelines for researching sex/gender in human health. *Soc. Sci. Med.* 74, 1817–1824. doi: 10.1016/j.socscimed.2011.05.033

Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *Am. Psychol.* 52, 613–629. doi: 10.1037/0003-066X.52.6.613

Steele, C. M., and Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* 69, 797–811. doi: 10.1037/0022-3514.69.5.797

Titze, C., Jansen, P., and Heil, M. (2010). Mental rotation performance and the effect of gender in fourth graders and adults. *Eur. J. Dev. Psychol.* 7, 432–444. doi: 10.1080/17405620802548214

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., et al. (2012). The malleability of spatial skills: a meta-analysis of training studies. *Psychol. Bull.* 139, 352–402. doi: 10.1037/a0028446

Vandenberg, S. G., and Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Percept. Mot. Skills* 47, 599–604. doi: 10.2466/pms.1978.47.2.599

Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: a meta-analysis. *Psychon. Bull. Rev.* 18, 267–277. doi: 10.3758/s13423-010-0042-0

Voyer, D., Butler, T., Cordero, J., Brake, B., Silbersweig, D., Stern, E., et al. (2006). The relation between computerized and paper-and-pencil mental rotation tasks: a validation study. *J. Clin. Exp. Neuropsychol.* 28, 928–939. doi: 10.1080/13803390591004310

Voyer, D., and Hou, J. (2006). Type of items and the magnitude of gender differences on the Mental Rotations Test. *Can. J. Exp. Psychol.* 60, 91–100. doi: 10.1037/cjep2006010

Voyer, D., and Voyer, S. D. (2014). Gender differences in scholastic achievement: a meta-analysis. *Psychol. Bull.* 140, 1174–1204. doi: 10.1037/a0036620

Voyer, D., Voyer, S. S., and Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychol. Bull.* 117, 250–270. doi: 10.1037/0033-2909.117.2.250

Vuoksimaa, E., Kaprio, J., Kremen, W. S., Hokkanen, L., Viken, R. J., Tuulio-Henriksson, A., et al. (2010). Having a male co-twin masculinizes mental rotation performance in females. *Psychol. Sci.* 21, 1069–1071. doi: 10.1177/0956797610376075

Walton, G. M., and Cohen, G. L. (2003). Stereotype lift. *J. Exp. Soc. Psychol.* 39, 456–467. doi: 10.1016/S0022-1031(03)00019-2

Wang, M. T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001

Watt, H. M., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., and Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: a comparison of samples from Australia, Canada, and the United States. *Dev. Psychol.* 48, 1594–1611. doi: 10.1037/a0027838

Wraga, M., Duncan, L., Jacobs, E. C., Helt, M., and Church, J. (2006). Stereotype susceptibility narrows the gender gap in imagined self-rotation performance. *Psychon. Bull. Rev.* 13, 813–819. doi: 10.3758/BF03194048

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 04 December 2018
doi: 10.3389/fpsyg.2018.02435

Gender Stereotypes in a Children's Television Program: Effects on Girls' and Boys' Stereotype Endorsement, Math Performance, Motivational Dispositions, and Attitudes

Eike Wille[1]*, Hanna Gaspard[1], Ulrich Trautwein[1], Kerstin Oschatz[1], Katharina Scheiter[2,3] and Benjamin Nagengast[1]

[1] Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany, [2] Leibniz-Institut für Wissensmedien, Tübingen, Germany, [3] University of Tübingen, Tübingen, Germany

Television programs are a central part of children's everyday lives. These programs often transmit stereotypes about gender roles such as "math is for boys and not for girls." So far, however, it is unclear whether stereotypes that are embedded in television programs affect girls' and boys' performance, motivational dispositions, or attitudes. On the basis of research on expectancy-value theory and stereotype threat, we conducted a randomized study with a total of 335 fifth-grade students to address this question. As the experimental material, we used a television program that had originally been produced for a national TV channel. The program was designed to show children that math could be interesting and fun. In the experimental condition, the program included a gender stereotyped segment in which two girls who were frustrated with math copied their math homework from a male classmate. In the control condition, participants watched an equally long, neutral summary of the first part of the video. We investigated effects on boys' and girls' stereotype endorsement, math performance, and different motivational constructs to gain insights into differential effects. On the basis of prior research, we expected negative effects of watching the stereotypes on girls' performance, motivational dispositions, and attitudes. Effects on the same outcomes for boys as well as children's stereotype endorsement were explored as open questions. We pre-registered our research predictions and analyses before conducting the experiment. Our results provide partial support for short-term effects of gender stereotypes embedded in television programs: Watching the stereotypes embedded in the video increased boys' and girls' stereotype endorsement. Boys reported a higher sense of belonging but lower utility value after watching the video with the stereotypes. Boys' other outcome variables were not affected, and there were also no effects on girl's performance, motivational dispositions, or attitudes. Results offer initial insights into how even short segments involving gender stereotypes in television shows can influence girls' and boys' stereotype endorsement and how such stereotypes may constitute one factor that contributes to gender differences in the STEM fields.

Keywords: stereotypes, gender differences, television, math motivation, math performance

Edited by:
*Bernhard Ertl,
Universität der Bundeswehr München,
Germany*

Reviewed by:
*Tomas Jungert,
Lund University, Sweden
Sog Yee Mok,
University of Zurich, Switzerland*

****Correspondence:***
*Eike Wille
eike.wille@uni-tuebingen.de*

Specialty section:
*This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology*

Received: *06 July 2018*
Accepted: *19 November 2018*
Published: *04 December 2018*

Citation:
Wille E, Gaspard H, Trautwein U, Oschatz K, Scheiter K and Nagengast B (2018) Gender Stereotypes in a Children's Television Program: Effects on Girls' and Boys' Stereotype Endorsement, Math Performance, Motivational Dispositions, and Attitudes. Front. Psychol. 9:2435. doi: 10.3389/fpsyg.2018.02435

INTRODUCTION

Women are underrepresented in domains that require intensive mathematical skills (National Science Foundation, 2015; National Science Board, 2016). This bias is crucial to the larger economy and contributes to gender inequity in income: More women in science, technology, engineering, and mathematics (STEM) would diversify the workforce, and mathematically intensive STEM fields usually provide high-status career options (National Science Foundation, 2015). Drawing on expectancy-value theory (Eccles et al., 1983), gender differences in STEM careers can be linked to early emerging gender differences in math motivational dispositions. These are rooted in different socialization processes for girls and boys such as the gender stereotypes children encounter in their environments (see Wigfield et al., 2015). Research on stereotype threat has provided insights into the potential mechanisms behind how gender stereotypes might affect girls and boys, indicating that girls can show lower math performance and motivation in the short-term if they are reminded of the stereotype that females perform worse than males in math, whereas boys' performance can benefit from such stereotypes (for a review, see Spencer et al., 2016).

Television programs are one potential source of gender stereotypes for children. Despite the wide diversity of media available nowadays, television continues to be one of the most popular and widely used media among children (Rideout, 2015; Feierabend et al., 2017). Television shows and programs with STEM content have increased in availability (National Reserach Council., 2009) and popularity (Patten, 2013) within the last decade. They transmit certain beliefs and stereotypes about gender roles in the STEM field, such as showing females as underperforming in math and science (Collins, 2011). It is not yet clear, though, whether stereotypes in television programs affect girls' and boys' performance and motivational dispositions in math. So far, research on expectancy-value theory has focused primarily on the role of stereotypes that are implicitly conveyed by parents, teachers, or peers (see Wigfield et al., 2015), whereas research on stereotype threat has traditionally investigated effects of stereotypes presented as isolated stimuli in laboratory settings with a primary focus on adult samples (see Spencer et al., 2016).

In the present study, we aimed to contribute to closing this gap in the literature by examining effects of traditional gender stereotypes in a math television program for children. To increase the ecological validity of the study, we used a television program that was broadcast on a German national TV channel. Specifically, the end of this program showed two girls who were not doing well in math and copied their homework from a male classmate. To examine the effects of these stereotypes, we conducted a randomized study with a pretest–posttest design in which fifth graders watched this television program about math either with or without the segment in which these gender stereotypes were portrayed. In order to comprehensively investigate possible effects, we studied effects on both girls' and boys' stereotype endorsement as well as their performance, motivational dispositions (i.e., expectancy and value beliefs), and attitudes toward math (i.e., sense of belonging, feelings about the domain).

Gender Differences in Motivational Dispositions and Achievement in Math From an Expectancy-Value Theory Perspective

Expectancy-Value Theory

Eccles et al. (1983) expectancy-value theory is one of the most widely used frameworks for investigating gender differences in motivational dispositions in math and has been highly effective in explaining women's underrepresentation in the STEM fields (Watt and Eccles, 2008; Schoon and Eccles, 2014).

In general, motivation can be defined as "the process whereby goal-directed activity is instigated and sustained" (Schunk et al., 2008, p.4). However, current work on motivation from the perspective of expectancy-value theory focusses mainly on expectancy and value beliefs as motivational dispositions (Eccles et al., 1983; Eccles, 2005). Specifically, Eccles et al. (1983) suggested that the expectation of success in a specific domain as well as several aspects of subjective task values would predict academic decision making and thereby also specific educational outcomes, such as later achievement or educational choices. Young people should thus choose math-intensive STEM careers if they expect to be good at math and science activities and have high values in these domains.

Eccles and Wigfield (2002) defined expectancies for success as a person's beliefs about his or her success in a task in the immediate or long-term future. Expectancy beliefs are therefore closely related to other competence beliefs, such as academic self-concept, which has often been used to measure expectancies for success (see Marsh, 2007; Nagengast et al., 2011). Eccles et al. (1983) differentiated four different components of subjective task values: intrinsic value, attainment value, utility value, and cost. Intrinsic value is defined as enjoyment while performing a task (Eccles, 2005). It is thus similar to other motivational constructs such as intrinsic motivation as defined by Deci and Ryan (1985)—which refers to reasons for engaging in a task, such as inherent satisfaction—or interest as defined by Renninger and Hidi (2011). Attainment value refers to the personal importance of doing well on a task or in a domain (Eccles, 2005). Utility value captures more extrinsic reasons for engaging in a task, namely the perceived usefulness of a task or domain (Eccles, 2005). Finally, cost captures negative aspects of engaging in a task or domain, such as required effort or time (Eccles, 2005).

Gender Differences in Motivational Dispositions and Achievement in Math

Ample research drawing upon expectancy-value theory has consistently indicated that girls exhibit lower expectancy and value beliefs (and higher cost) for math than boys from an early age on (for reviews, see Wang and Degol, 2013; Wigfield et al., 2015). By contrast, meta-analyses investigating gender differences in math achievement have shown rather small advantages for boys compared with girls (e.g., Else-Quest et al., 2010; Reilly et al., 2015). Moreover, these analyses have indicated that such gender differences seem to occur only on math achievement tests (Reilly et al., 2015), whereas girls even show an advantage in teacher-assigned school marks (Voyer and Voyer, 2014).

The Role of Stereotypes in the Development of Children's Motivational Dispositions and Achievement

According to expectancy-value theory, socializers' beliefs and behaviors as well as cultural milieu influence individuals' task perceptions and interpretations of previous academic achievement (Eccles et al., 1983). In explaining gender differences in expectancy and value beliefs and achievement, expectancy-value theory thus indicates that girls and boys are socialized through different processes, which are shaped by the surrounding environment and its gender norms and roles, the individuals' beliefs, and the choices females and males make on the basis of their socialization (Eccles, 2009). In particular, gendered socialization refers to specific gender roles or the gender-stereotypical attitudes and expectancies of parents, teachers, and other socializing influences such as the media, all of which transmit gender stereotypes (Wigfield et al., 2015).

Stereotypes can be broadly defined as associations of group members with specific attributes (Greenwald et al., 2002). Regarding gender, there are specific stereotypes about the traits, abilities, and motivation of males and females, specifically in the domain of math (see Leaper, 2015). Math and science are male-typed domains, and gender stereotypes in these domains include assumptions about lower abilities and less talent in math for females compared with males (e.g., Spencer et al., 1999).

According to expectancy-value theory, as a result of the gender stereotypes children face in their socialization, girls disidentify with math and devalue the subject in the long run, whereas boys may particularly identify with and value math (Eccles et al., 1983; Wigfield et al., 2015). Consequently, boys develop higher competence beliefs and values in male-typed domains such as math and math-intensive STEM domains, whereas girls develop higher competence beliefs and values in female-typed domains such as languages and arts (e.g., Wigfield et al., 2015). It is assumed that such gender differences in math competence beliefs and values may lead to gender differences in math achievement in the long run (Wigfield and Eccles, 2000). Previous studies have supported these assumptions by showing that women's gender stereotypes reduced their domain identification (e.g. their positive attitudes and their sense of belonging; Cheryan et al., 2009; see also Thoman et al., 2013 for a review) as well as their future expectancies of success (Smith et al., 2015) and their future task values (Plante et al., 2013; Smith et al., 2015). Expectancy and task values, in turn, have been shown to be important predictors of later achievement (e.g., Marsh et al., 2005; Denissen et al., 2007).

Stereotype Threat as a Potential Mechanism for How Stereotypes Can Influence Children

The repeated experience of stereotypes is one potential mechanism that may explain how stereotypes of others can influence girls' and boys' performance, expectancy and value beliefs, and attitudes toward math. According to expectancy-value theory, such experiences might lead to the internalization of gender-role stereotypes, with the previously described consequences that girls disidentify with and devalue math, and boys particularly identify with and value math in the long run (Eccles et al., 1983; Wigfield et al., 2015).

Research on stereotype threat has provided support for this idea by showing that the activation of traditional gender stereotypes can reduce girls' attitudes and belonging in math as well as their performance and motivational dispositions in the short term (for a review, see Spencer et al., 2016). Steele and Aronson (1995) defined stereotype threat as a situational experience in which group members feel concerned about confirming a negative stereotype that pertained to their own group. They suggested that such concerns might compromise a person's behavior and performance.

Stereotype Threat and Girls' Performance, Motivational Dispositions, and Attitudes

Originally, research on stereotype threat focused on explaining the underperformance of African Americans in performance (Steele and Aronson, 1995), but ample research has also been conducted to examine gender differences in math-intensive domains (e.g., Spencer et al., 1999; Schmader, 2002; Tomasetto et al., 2011). Such research has demonstrated that females show lower math performance if they are reminded of negative stereotypes about women in math, but they perform as well as males if such stereotypes are not made salient before they take a math test (Nguyen and Ryan, 2008; Doyle and Voyer, 2016). Although most of this research has been conducted on college students or older adults, multiple studies have reported similar effects among children or adolescents (e.g., Ambady et al., 2001; Flore and Wicherts, 2015). These studies have demonstrated that children in elementary school are already aware of their own gender and show gender-stereotypical views in the domain of math, as they attribute lower math ability and talent to girls and women than to boys and men (e.g., Signorella et al., 1993; Ambady et al., 2001; Passolunghi et al., 2014). In addition, there is research on the short-term effects of stereotypes on math performance among girls of different ages (Ambady et al., 2001; Muzzatti and Agnoli, 2007; Neuville and Croizet, 2007; Tomasetto et al., 2011; Hermann and Vollmeyer, 2016). A meta-analysis by Flore and Wicherts (2015), for instance, found that girls who are reminded of typical gender stereotypes in math exhibit slightly lower math performance compared to girls who are not reminded of such stereotypes. Such effects have been consistently found for girls younger than 13 years old.

Effects of stereotype threat have also been shown for females' motivational dispositions and attitudes toward a domain, such as their domain identification and their sense of belonging in math and science (e.g., Cheryan et al., 2009; see also Thoman et al., 2013, for a review), their competence beliefs (Cadinu et al., 2003), and their interest (Smith et al., 2007; see also Thoman et al., 2013, for a review). Again, much of this work has been conducted on adult samples. However, there are a few studies reporting similar effects for girls. A study by Muzzatti and Agnoli (2007) indicated stereotype threat effects on 8th grade girls' competence beliefs in math, although no effects were found for 3rd and 5th graders. Furthermore, Master et al. (2015) found stereotype threat effects on 15-years-old female high school students' interest and sense of belonging in STEM courses.

Stereotype Threat and Boys' Performance, Motivational Dispositions, and Attitudes

Effects of stereotypes on boys' performance, motivational dispositions, and attitudes toward a domain are less clear, as there are only a few studies on such effects and contradictory findings have been reported. Muzzatti and Agnoli (2007), for example, found no effects of presenting stereotypes on boys' math performance in Grades 3, 5, and 8 as well as their math competence beliefs in Grades 3 and 5 (see also Hermann and Vollmeyer, 2016 for similar results on boys in elementary school). However, among 8th graders, they found higher competence beliefs among boys who were confronted with the stereotype of males' advantage in math compared to the control group (Muzzatti and Agnoli, 2007). Similarly, Master et al. (2015) found no effects of stereotypes on male adolescents' sense of belonging and interest in enrolling in computer courses.

In addition, there is some work on the effects of stereotypes on males using adult samples that also suggest that males are not much affected by stereotypes (Walton and Cohen, 2003; Cheryan et al., 2009; Fogliati and Bussey, 2013; Doyle and Voyer, 2016). Although a meta-analysis by Walton and Cohen (2003) indicated positive effects of traditional gender stereotypes for men's math performance, a more recent meta-analysis by Doyle and Voyer (2016) found no effects. Furthermore, no effects of traditional gender stereotypes have been reported with respect to men's interest and belonging in computer science (Cheryan et al., 2009) or their motivation to improve in math (Fogliati and Bussey, 2013).

In sum, several studies indicate effects of stereotypes on females' performance, motivational dispositions, and attitudes toward math, whereas most studies have reported no effects for males. Nevertheless, the abovementioned studies on stereotype threat effects should be interpreted with caution because the robustness of such effects has recently been called into question due to indications of publication bias in a meta-analysis of this research (Flore and Wicherts, 2015).

Effects of Stereotypes Presented in the Media

Research on expectancy-value theory has focused primarily on the influence of parents, teachers, or peers on children's endorsement of stereotypes and their expectancy and value beliefs (see Wigfield et al., 2015), but research in the area of media psychology and communication studies has suggested that television programs and movies can contribute to children's gender-role learning in terms of their perceptions of gender-typical occupations (Steinke et al., 2007) or their gender-role values and interpersonal attraction (Aubrey and Harrison, 2004). In addition, research on stereotype threat has indicated a wide range of situations, such as newspaper articles (Cheryan et al., 2013), images in schoolbooks (Good et al., 2010), and photographs (Muzzatti and Agnoli, 2007), in which stereotypes about females' underperformance in math can affect both females and males.

In a recent meta-analysis, Appel and Weber (2017) investigated how stereotypes in mass media (e.g., newspapers, cartoons, advertisements) can affect stereotyped and non-stereotyped groups. In this analysis, negative effects of $d = -0.38$ for members of the stereotyped group and positive effects of $d = 0.17$ for members of the non-targeted group were reported.

Additionally, there are a few studies specifically investigating effects of stereotypes in videos and television advertising (Davies et al., 2002; Murphy et al., 2007; Bond, 2016). Bond (2016) presented short clips of different television shows (about 2 min long) to elementary school girls in a gender stereotype condition, a counter-stereotype condition, and a neutral control condition. No effects of the stereotypes were found on math and science competence beliefs or interest in STEM-related careers. However, girls in the stereotype condition reported more interest in stereotypical careers than those in the other two conditions.

In an adult sample, Murphy et al. (2007) found negative effects of reminding women of their underrepresentation in math-intensive STEM fields via video on their sense of belonging as well as intention to participate in a STEM-related conference. In this study, women in the stereotyped condition watched a video in which the male-female ratio reflected the proportion of women in these fields, whereas women in the control condition watched a video with a gender-balanced proportion.

Davies et al. (2002) showed that women experience stereotype threat when they are reminded of existing stereotypes about women in television advertising. In this study, participants watched commercials in which women were very excited about buying cosmetic products or trying a new baking recipe. After watching these commercials, women performed worse on a math test compared with men who watched the same commercials and compared with women who watched gender-neutral commercials. The results furthermore showed that women preferred verbal tasks and avoided math-related tasks after watching such commercials compared with the control group and men in the experimental group. Women also showed less interest in educational and vocational areas that are typically male-stereotyped but higher interest in typically female-stereotyped domains.

The reported studies indicate that stereotypes in videos can have negative effects on females. However, these findings provide only initial insights into the effects of television. Furthermore, these studies investigated stereotypes that were presented in isolated situations. Thus, they were not able to provide insights into how stereotypes might affect children when experienced in their daily lives in more complex situations, for instance, as one part of a whole television program.

The Present Study

In the present study, we investigated effects of gender stereotypes in a STEM television program on girls' and boys' stereotype endorsement, their math performance, their motivational dispositions (i.e., expectancy and value beliefs), and their attitudes (i.e., sense of belonging and feeling) toward math. Despite the importance of television programs in children's everyday lives and the relevance of such programs for children's informal science learning, there is a lack of research on how girls' and boys' reception of STEM television programs might be affected in different ways by presentations of traditional gender stereotypes in such programs. Research on expectancy-value theory and stereotype threat has provided initial insights into how stereotypes might affect children. However, research on expectancy-value theory has mainly focused on the role of stereotypes that are conveyed by parents, teachers, or peers (see Wigfield et al., 2015), and research on stereotype threat has traditionally investigated effects of stereotypes presented as isolated stimuli in laboratory settings on adults (see Spencer et al., 2016). Furthermore, there are indications of publication bias in the stereotype threat literature (Flore and Wicherts, 2015). Accordingly, it is unclear whether and how stereotypes embedded in children's daily activities such as in a television program might affect girls and boys.

Therefore, we conducted a randomized study in which fifth-grade students watched a children's television program about math that either contained or did not contain a clip in which traditional gender stereotypes were made salient. We chose this age group because of specific developmental processes in children's expectancy and value beliefs during that age. During their elementary school years, children become increasingly better at understanding, interpreting, and integrating the feedback of others (for a review, see Wigfield et al., 2015). Therefore, they become more realistic in evaluating their own strengths and weaknesses during that period and link their expectancy and value beliefs more closely to environmental experiences than younger elementary school children (for a review, see Wigfield et al., 2015). Additionally, children become increasingly aware of social gender roles and how behavior might reflect such roles (for a review, see Leaper, 2015). In order to link the study as closely as possible to what children are likely to watch in their everyday lives, we used a television program that was broadcast on a national TV channel in Germany as the experimental material. The chosen program was designed to show children that math could be interesting and fun and included a section with stereotypes in which two girls were frustrated that they had to do math and then decided to copy their homework from a male classmate.

According to expectancy-value theory, experiencing gender stereotypes leads girls to disidentify with math and devalue the subject, whereas boys may particularly identify with and value math. As a result of such processes, boys develop higher competence beliefs and values in male-typed domains such as math and math-intensive STEM domains than girls (e.g., Wigfield et al., 2015). In order to obtain a comprehensive picture of how stereotypes can affect such socialization processes, we examined effects of the experimental manipulation on different outcomes. First, we explored how the stereotypes affect children's stereotype endorsement. Second, we examined effects on sense of belonging in math and feeling toward the domain as indicators of children's identification with the subject. Third, we investigated effects on self-concept (as an indicator of expectancy beliefs), the four task values as well as performance in math. We pre-registered our predictions on the effects for these outcomes before conducting the experiment in order to increase research transparency (https://osf.io/8f7y6/?view_only=d85b73e70f5040b5a54fcf03091811f1). As such, we followed the recommendations of Wagenmakers et al. (2012) and van't Veer and Giner-Sorolla (2016) by pre-registering hypotheses and exploratory research questions as well as information on the experimental design, the sample, the variables, and the analysis strategy.

On the basis of existing literature on effects of stereotypes on math performance (Flore and Wicherts, 2015), self-concept (Cadinu et al., 2003; Muzzatti and Agnoli, 2007), and sense of belonging (Master et al., 2015), we expected that girls who watched the gender-stereotyped television program would show lower math performance, lower math self-concept, and a lower sense of belonging in math compared with girls in the control condition.

We explored effects on girls' task values in math and their feelings about math as open-ended research questions. There is only sparse evidence on how task values might be influenced by gender stereotypes (Plante et al., 2013; Smith et al., 2015), and previous work has not differentiated between the four components (intrinsic value, attainment value, utility value, and cost). Furthermore, to the best of our knowledge, there is no work that has investigated effects of stereotypes on children's feelings about a domain. We therefore did not hypothesize specific effects on task values and feelings about math.

In order to gain insights into possible differential effects of such stereotypes on girls and boys, we explored effects on boys' performance, expectancy and value beliefs, sense of belonging and feeling toward the domain in math-related constructs as well, using the same outcomes measures. Due to the mixed findings from previous research on the effects of stereotypes on such constructs for males, we did not hypothesize specific effects for boys but rather investigated possible effects on these outcomes for boys as exploratory research questions.

We did not formulate any specific hypotheses with respect to the endorsement of gender stereotypes among both girls and boys, because previous research has provided mixed results on the effects of gender stereotypes on children's endorsement of gender stereotypes (Ambady et al., 2001; Schmader et al., 2004; Steffens et al., 2010).

METHODS

Participants

Participants were 335 fifth-grade students. Children were recruited from 18 classes of four academic track schools (Gymnasium) in Baden-Württemberg, Germany. The sample size was based on a power analysis for a randomized block trial with the treatment implemented at the student level using Optimal Design (Raudenbush et al., 2011). We calculated the required number of classrooms by aiming to achieve an acceptable level of power ($\beta = 0.80$) to detect medium-sized intervention effects ($\delta = 0.40$) when comparing the experimental with the control condition. We assumed that 10 girls and 10 boys would participate in each class, and they would be randomly assigned to the control and experimental conditions.

We furthermore assumed an effect size variability of 0.10 (for more details, see the preregistration protocol).

Children participated in the study on a voluntary basis, and for every participant, we obtained written consent from a parent. The mean age of the sample was 10.08 years ($SD = 0.38$), and the number of girls and boys who participated in the study was almost equal (48.7% girls).

Design and Procedure

As preregistered, we collected the data using a pretest–posttest design, and we applied a randomized block design to examine effects of gender stereotypes in a television program. Girls and boys were randomly assigned to the experimental and control conditions within each class (experimental condition: $N = 87$ girls and $N = 85$ boys; control condition: $N = 76$ girls and $N = 87$ boys). Participants were tested in one classroom simultaneously, but every student watched the video separately on an iPad with headphones. We collected the pretest data 1 week before the experimental manipulation and the posttest data directly after the experimental manipulation. The presentation order of the achievement test and the questionnaire was balanced on the class level in both phases of data collection because research on stereotype threat has shown that even small and short manipulations can influence students' performance, motivational dispositions, and attitudes (e.g., Master et al., 2015; i.e., the achievement test might affect students' motivational dispositions and attitudes if assessed first, or the questionnaire might wash out any effects on performance). We randomly assigned the classes to these two conditions ($N = 9$ classes in each condition). Data were collected in June and July 2016 by trained research assistants during school hours (a maximum of one lesson for the pretest, a maximum of two lessons for the experiment and the posttest).

Experimental Manipulation

As experimental material, we used one episode from a German children's television program, which was broadcast on a German national television channel in June 2015. The episode focused on math and was designed to show children that math could be interesting and fun even though it might be experienced as boring in school (KiKa.de, 2015). The episode had a total duration of 23 min. As preregistered, only 15 min of the episode were used in the present study due to time constraints. This included an introduction by a male television presenter (about 1 min) and two different math tasks solved by fifth-grade children (about 13 min). In addition, the video included a clip that implied traditional gender stereotypes in math (about 1 min). This part showed two girls who were very frustrated that they had to do math homework. Instead of doing their homework, one girl copied it from a male classmate, and in exchange, she promised him that her friend would accompany him to the movies. Her friend was horrified about going out with this boy because he seemed rather geeky. He was wearing very large glasses, a shirt that was completely buttoned up, suit trousers, and suspenders. Such stereotypes of the geeky math boy are often presented in movies or television programs (see e.g., Heyman, 2008; Collins, 2011).

The introduction and the math tasks solved by the children were used in both conditions. The experimental manipulation depended on only the last minute of the video. In the experimental condition, participants watched the gender-stereotyped clip. In the control condition, participants watched a neutral summary of the first 14 min of the video. The summary was comparable in length so that the total length of the video would be held constant between the conditions. Consequently, participants experienced the stereotype as a short section within the whole television program so that the ecological validity of the experiment would be high.

Because the television program was broadcast on a national TV channel in Germany, we assessed whether participants had already seen the video beforehand, which was the case for 41 students. As a robustness check, we computed all analyses without these students, but the results did not differ meaningfully (see the **Supplemental Material**).

Instruments

We used an achievement test and a questionnaire to assess effects of the experimental manipulation. The instruments were identical at pre- and posttest, with the exception of questions about the video, which were only assessed at posttest.

Math Performance

We assessed students' math performance with a speed test that consisted of three sections containing basic tasks involving addition, subtraction, and multiplication (basic competence test; Lambert et al., in preparation). Each part consisted of 36 tasks, and for each individual part, we asked the students to solve as many tasks as possible within 2 min. The sum score of all three parts, generated by computing the sum of correctly solved items, was used in the analyses. The test showed high internal consistency (Kuder-Richardson 20 = 0.93/0.94 for the pretest/posttest).

Questionnaire

We assessed children's stereotype endorsement, their motivational dispositions (i.e., self-concept and value beliefs) as well as their attitudes toward math (i.e., sense of belonging and feelings) with a questionnaire to capture whether children (dis)identify with and (de)value this domain after watching the video including the stereotypes. Unless otherwise noted, all items on the questionnaire were measured with a 4-point Likert scale ranging from 1 (*completely disagree*) to 4 (*completely agree*). The 4-point Likert scale was used to avoid confounding response factors in scales containing a middle category (Kaplan, 1972; Dubois and Burns, 1975). Additionally, four response options seems to be optimal for children, as they are not able to differentiate between more categories (Borgers et al., 2004). Due to the small number of response options, we carefully checked the degree of non-normality in our data. Although there was some variation across scales, the skewness and kurtosis values all fell within an acceptable range (average skewness was −0.36, with no scale having a skewness >1.4, and the average kurtosis was 0.59, with only 2 scales having a kurtosis >1). The questionnaire is available at https://osf.io/8f7y6/?view_only=d85b73e70f5040b5a54fcf03091811f1.

Stereotype endorsement

We assessed stereotype endorsement with three items based on items from Schmader et al. (2004). We adapted the items for children by using "boys" and "girls" in the wording instead of "men" and "women" (e.g., "Boys have higher math abilities than girls"; $\alpha = 0.76/0.76$ for the pretest/posttest).

We extended the scale by including two items in which the words "boys" and "girls" were interchanged (e.g., "Girls have better math abilities than boys") and preregistered this extension. We recoded these items before computing the scale score. Because the reliability of the extended scale was rather low ($\alpha = 0.52/0.55$ for the pretest/posttest), we used only the original scale in our analyses.

Task values

We assessed students' value beliefs in math with scales from Gaspard et al. (2015). The items covered all four conceptual dimensions of task values as specified in the expectancy-value model (Wigfield and Eccles, 2000). Intrinsic value (e.g., "I like doing math"; $\alpha = 0.92/0.94$ for the pretest/posttest), attainment value (e.g., "It is important to me to be good at math"; four items; $\alpha = 0.87/0.93$ for the pretest/posttest), and cost (emotional costs, e.g., "Studying math makes me quite nervous"; $\alpha = 0.78/.86$ for the pretest/posttest) were assessed with four items each. For utility value, we differentiated between two facets: utility for daily life (e.g., "Knowing about the subject of math brings me many advantages in my daily life"; $\alpha = 0.82/0.84$ for the pretest/posttest) and social utility (e.g., "Sound knowledge in math counts for something with my classmates"; $\alpha = .68/.80$ for the pretest/posttest), which were both assessed with three items.

Self-concept

We assessed self-concept with a math self-concept scale comprised of four items (e.g., "I am good at math"; $\alpha = .86/.86$ for the pretest/posttest), which has been well-validated in previous studies (see Gaspard et al., 2016).

Sense of belonging

We assessed students' sense of belonging in math with 10 items (e.g., "I feel like a real part of my class in math"), based on the Psychological Sense of School Membership (PSSM; Goodenow, 1993). The items were translated into German and adapted to math class instead of school membership. Due to low item-scale correlations ($r_{it} = 0.03/0.16$ for the pretest/posttest), we excluded 1 item when we computed the scale. The final scale therefore consisted of 9 items and showed an acceptable internal consistency ($\alpha = 0.76/0.84$ for the pretest/posttest). Because we did not preregister the exclusion of the item, we conducted the analysis for this outcome also using the original scale, which included all 10 items. The internal consistency for this scale was acceptable ($\alpha = 0.73/0.83$ for the pretest/posttest), and the results did not differ meaningfully from those computed with the reduced scale (see the **Supplemental Material** for this as well as for model fit indices from confirmatory factor analyses of the scales).

Explicit attitudes toward math

We assessed explicit attitudes toward math with a feeling thermometer as used by Kessels et al. (2006). Students were asked to rate their preferences using scales ranging from 0 (*cold/unfavorable*) to 100 (*warm/favorable*) for math and German. As done by Kessels et al. (2006), we calculated the difference between the two scores as an indicator of students' attitudes toward the domains. Therefore, the final score consisted of possible values ranging from −100 to +100, whereby positive values indicated positive attitudes toward math relative to German, and negative values indicated negative attitudes toward math relative to German.

Additional scales

As preregistered, we additionally assessed stereotype endorsement with measures based on studies by Ambady et al. (2001) and Steffens et al. (2010) in which the participants were asked how much they would like to engage in activities related to math and German. Due to high rates of missing data and the low reliability of these scales, we refrained from conducting additional analyses on these instruments.

We furthermore preregistered analyses with respect to the same set of constructs (i.e., task values, self-concept, sense of belonging) in the domain of German. Dimensional comparisons of complementary domains are important in the development of students' motivational dispositions (Möller and Marsh, 2013), and there are initial findings on how motivational dispositions in a verbal domain might be affected by traditional gender stereotypes in commercials (Davies et al., 2002). Due to space limitations, the results on girls' and boys' motivational dispositions and attitudes in German are reported in the **Supplemental Material**. In summary, we found no effects of the experimental condition on girls' and boys' motivational dispositions and attitudes in German except that girls in the experimental condition reported lower cost in German than those in the control condition.

Statistical Analyses

In order to estimate effects of the gender stereotypes in the television program, we computed multiple regression analyses for the different outcomes in Mplus 7.31 (Muthén and Muthén, 2012) as preregistered. All models included student gender (pacifier coded, boy = 1), the experimental condition (a pacifier-coded variable based on students' original assignment, experimental condition = 1), and the Gender × Condition interaction as predictor variables. In addition, we included the respective pretest measures as covariates to estimate the effect of the experimental manipulation more precisely (Raudenbush, 1997). In order to make it easier to interpret the results, we standardized all continuous predictors (i.e., the pretest scores) and the respective dependent variable.

In our analyses, we conducted an intention-to-treat analysis by taking only the original assignment into account in order to keep the randomization to the experimental and control conditions intact (Shadish et al., 2002). As a robustness check, we ran all analyses without the students who did not correctly answer a question about what they had seen in the last minute of the

TABLE 1 | Descriptive statistics for all study variables on the pretest separated by gender.

Variable	Girls		Boys		d^a	*d* 95% CI	
	M	*SD*	*M*	*SD*			
Stereotype endorsement T1	2.55	0.52	2.73	0.45	0.35	0.20	0.50
Performance T1	51.88	8.14	56.09	8.99	0.48	0.33	0.62
Self-concept T1	3.15	0.73	3.40	0.61	0.37	0.20	0.55
Sense of belonging T1	3.16	0.48	3.19	0.45	0.05	−0.16	0.26
Feeling thermometer T1	1.62	33.48	15.80	32.07	0.42	0.25	0.60
Intrinsic value T1	3.12	0.76	3.27	0.74	0.20	0.03	0.37
Attainment value T1	3.50	0.57	3.46	0.61	−0.06	−0.31	0.20
Utility value—daily life T1	3.24	0.66	3.26	0.68	0.02	−0.15	0.20
Utility value—social T1	2.22	0.68	2.41	0.64	0.29	0.11	0.47
Cost T1	1.60	0.60	1.53	0.52	−0.13	−0.33	0.08

C = confidence intervall.
[a]The dependent variable is standardized.

video, that is, two girls who copied the homework of a classmate in the experimental condition or a summary of the video in the control condition ($n = 13$). This question was assessed at the end of the posttest questionnaire. The results did not differ meaningfully and are presented in the **Supplemental Material**.

To test whether there were any order effects of the instruments, we computed multiple-group regression analyses with the order of the instruments as the grouping variable. We tested the difference between the models for each group with Wald χ^2 tests. If there were no significant differences between the coefficients in the models, we calculated multiple regressions for the whole sample.

Missing data ranged from 2.1% to 9.9% for the different scales because some students were absent when the pre- or post-test was given, and some students did not respond to individual scales. To deal with missing data, we used the full information maximum likelihood approach as implemented in Mplus 7.31 (Muthén and Muthén, 2012).

We considered the clustered structure of the data (students nested in classes) by using the design-based correction of standard errors implemented in Mplus 7.31 (Muthén and Muthén, 2012).

RESULTS

Descriptive Statistics and Randomization Check

The means and standard deviations for all scales are shown by gender and condition in **Tables 1–3**. Compared with boys, girls showed significantly lower math performance and reported lower levels of the feeling thermometer, self-concept, intrinsic value, and social utility value on the pretest. The correlations for the outcome variables indicate that the mean levels were relatively stable across the two measurement points for all outcomes ($0.60 < r < 0.87$; see **Table 4**).

To test whether the randomization in the two conditions had been successful in the baseline measures, we computed multiple regression models as preregistered (pretest values regressed on the experimental condition, gender, and the Gender × Condition interaction). There were no significant differences between the conditions for girls and boys on the pretest values for all variables (all ps > 0.137) except for the boys with respect to sense of belonging. Here, boys in the experimental condition showed lower baseline scores than those in the control condition [$d = 0.36$, 95% CI [0.07, 0.65]]. As preregistered, we controlled for the pretest scores in all analyses to estimate the effect of the experimental manipulation more precisely because of the explanatory power of this covariate.

Effects of the Experimental Manipulation

First, we tested if there were any order effects of the instruments by computing multiple-group regression analyses using the order of the instruments as the grouping variable. Wald χ^2 tests indicated no differences in these models with respect to any of the studied outcomes (all ps > 0.154) except for social utility value, where the coefficients for the Gender × Condition interaction differed significantly, $\chi^2_{(1)} = 11.76$, $p = 0.001$. Consequently, we computed multiple regression analyses using the total sample for all outcomes (i.e., averaged across instrument order) except for social utility value (see **Tables 5–7**).

We specified multiple regressions to test effects of the experimental manipulation (see **Tables 5**, **6**). As girls were coded 0, the main effect of the experimental condition was equal to the simple slope for girls, whereas the Gender × Condition interaction term indicated whether the effects differed between boys and girls. Because we were more interested in investigating effects of the experimental manipulation on girls' and boys' performance, motivational dispositions, and attitudes rather than on gender differences in these outcomes, we additionally estimated the simple slopes for boys for all outcomes using the model constraint in Mplus.

With respect to stereotype endorsement, we did not hypothesize specific effects due to mixed previous results for effects of stereotype threat on this outcome. The results revealed a

TABLE 2 | Descriptive statistics for all outcome variables at T1 separated by gender and condition.

	Girls								Boys							
	Experimental condition				Control condition				Experimental condition				Control condition			
Variable	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max
Stereotype endorsement T1	2.52	0.47	1.00	3.33	2.59	0.58	1.00	4.00	2.70	0.44	1.67	3.67	2.76	0.47	2.00	4.00
Performance T1	51.33	8.28	33.00	71.00	52.51	7.98	31.00	73.00	55.79	8.65	34.00	74.00	56.40	9.38	37.00	86.00
Self-concept T1	3.20	0.64	1.25	4.00	3.09	0.75	1.50	4.00	3.39	0.61	1.50	4.00	3.41	0.61	1.75	4.00
Sense of belonging T1	3.16	0.51	1.63	4.00	3.16	0.46	1.57	4.00	3.10	0.47	1.56	3.89	3.27	0.43	2.11	4.00
Feeling thermometer T1	3.86	34.10	−80.00	100.00	−0.88	32.83	−100.00	90.00	16.40	35.84	−90.00	100.00	15.19	27.89	−70.00	70.00
Intrinsic value T1	3.16	0.74	1.00	4.00	3.07	0.77	1.00	4.00	3.28	0.74	1.00	4.00	3.26	0.75	1.00	4.00
Attainment value T1	3.54	0.49	2.00	4.00	3.45	0.65	1.00	4.00	3.46	0.62	1.50	4.00	3.46	0.60	1.50	4.00
Utility value—daily life T1	3.20	0.64	1.33	4.00	3.29	0.67	1.33	4.00	3.20	0.67	1.00	4.00	3.32	0.68	1.00	4.00
Utility value—social T1	2.22	0.65	1.00	4.00	2.22	0.71	1.00	4.00	2.41	0.61	1.00	3.67	2.42	0.67	1.00	4.00
Cost T1	1.58	0.55	1.00	3.50	1.63	0.66	1.00	4.00	1.53	0.51	1.00	3.00	1.53	0.53	1.00	3.50

TABLE 3 | Descriptive statistics for all outcome variables at T2 separated by gender and condition.

	Girls								Boys							
	Experimental condition				Control condition				Experimental condition				Control condition			
Variable	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max	*M*	*SD*	Min	Max
Stereotype endorsement T2	2.68	0.51	1.00	4.00	2.45	0.56	1.00	3.67	2.75	0.50	1.00	4.00	2.67	0.48	1.67	4.00
Performance T2	54.52	8.42	31.00	73.00	55.28	8.40	36.00	73.00	58.04	8.04	39.00	73.00	58.74	8.19	37.00	73.00
Self-concept T2	3.21	0.66	1.00	4.00	3.08	0.76	1.00	4.00	3.43	0.62	1.75	4.00	3.33	0.61	1.75	4.00
Sense of belonging T2	3.12	0.54	1.56	4.00	3.18	0.53	1.78	4.00	3.12	0.50	1.67	4.00	3.15	0.55	1.67	4.00
Feeling thermometer T2	1.46	35.04	−100.00	100.00	1.00	36.00	−100.00	100.00	18.00	33.22	−80.00	90.00	17.60	32.50	−70.00	100.00
Intrinsic value T2	3.16	0.79	1.00	4.00	3.05	0.86	1.00	4.00	3.26	0.73	1.00	4.00	3.17	0.75	1.50	4.00
Attainment value T2	3.57	0.56	2.00	4.00	3.53	0.62	1.25	4.00	3.49	0.60	1.25	4.00	3.52	0.62	2.00	4.00
Utility value—daily life T2	3.36	0.59	1.33	4.00	3.40	0.63	1.00	4.00	3.29	0.68	1.00	4.00	3.37	0.68	1.33	4.00
Utility value—social T2	2.22	0.71	1.00	4.00	2.06	0.70	1.00	4.00	2.20	0.72	1.00	3.67	2.32	0.78	1.00	4.00
Cost T2	1.54	0.65	1.00	4.00	1.63	0.70	1.00	4.00	1.47	0.53	1.00	3.00	1.49	0.57	1.00	3.50

significant positive effect of the experimental condition for girls. The same result held for boys because the Gender × Condition interaction was not statistically significant (see **Table 5**).

Regarding math performance, math self-concept, and sense of belonging, we hypothesized that girls in the experimental condition would score lower on these outcomes than girls in the control condition. For boys, we did not hypothesize specific effects. For these outcomes, the results revealed no significant effect of the experimental condition for girls. For math performance and math self-concept, there were also no effects of the condition for boys. With respect to sense of belonging, the Gender × Condition interaction was statistically significant, and there was a positive effect of the condition for boys, indicating that in contrast to girls, boys in the experimental condition showed higher values of sense of belonging than boys in the control condition (see **Table 5** and **Figure 1**).

Regarding task values and attitudes toward math assessed with the feeling thermometer, we did not hypothesize specific effects of the experimental condition for girls and boys. With respect to the feeling thermometer, intrinsic value, attainment value, utility value for daily life, and cost, we found no significant effects of the experimental condition for either girls or boys (see **Table 6**).

For social utility, we computed multiple-group regression analyses using the order of the instruments as a grouping variable because a Wald χ^2 test indicated effects of the order of the instruments in the assessment as described above. Because we were interested in the effects of the experimental manipulation on social utility assessed with the questionnaire, the results for the students who were given the questionnaire first in the assessment were of major interest. For the students who were given the questionnaire first, there was no significant effect of the condition for girls, but the Gender × Condition interaction was statistically significant, indicating that boys in the experimental condition reported a significantly lower social utility score than those in the control condition (see **Table 7** and **Figure 1**). For the students who were given the achievement test first, there was no significant effect of the condition for girls or for boys (see **Table 7**).

DISCUSSION

In this experimental study, we examined how stereotypes embedded in a children's television program about math influence girls' and boys' stereotype endorsement, math performance, motivational dispositions and attitudes in math. We used a randomized study with a pretest–posttest design and a relatively large sample size, which enabled us to detect medium-sized effects. The material we chose was a television program that had been broadcast on a German national television channel, thus contributing to the high validity of the study. Television programs play a central role in children's everyday lives and are an important part of their informal science learning, but such programs can provide specific gender stereotypes about math (National Reserach Council., 2009; Collins, 2011; Rideout, 2015). Previous research has indicated that the stereotypes children encounter in their environment can impact young girls' and boys' math performance, motivational dispositions, and attitudes. Yet, such research has primarily been conducted in laboratory settings where stereotypes have been presented as isolated stimuli, rather than integrated into other information as would be the case in children's daily lives, for instance, in television programs.

Overall, our results did not indicate that children's performance, motivational dispositions, and attitudes were strongly affected by the stereotypes presented in one part of a television program. However, girls and boys in the experimental condition reported a higher endorsement of stereotypes compared with the respective control condition. Furthermore, boys showed a higher sense of belonging but lower social utility after watching the video that included the stereotypes compared with boys in the control condition. We did not find any effects on either the other motivational dispositions, attitudes or math performance for boys. We also did not find any effects on math performance, motivational dispositions, and attitudes for girls.

Discussion of the Findings

First of all, the small number of significant effects found in this study support previous research indicating that the short-term effects of stereotypes on performance, motivational dispositions, and attitudes are not as robust as sometimes claimed. For example, Stoet and Geary (2012) reviewed replication attempts of the stereotype threat effect on women's math performance that was found in Spencer et al. (1999) original study. According to this review, only 30% of replication studies confirmed the original finding. In addition, Flore and Wicherts (2015) found indications of publication bias in their meta-analysis on stereotype threat effects in children. In accordance with these findings, the non-significant effects found in our study indicate that stereotype threat effects might occur only in specific situations or might apply only to some children. Here, the negative effect on boys' social utility might add to this discussion because this effect was found only for students who were given the questionnaire first (in the assessment in which we also assessed social utility). We did not find any effects of condition among boys who were given the questionnaire after the achievement test. Therefore, the stereotypes might have affected boys' social utility in the short term, but were washed out after they completed the achievement test, indicating that even if stereotype threat effects occur, they might be very limited in duration and sensitive to other influences.

Nevertheless, specific characteristics of the present study could have also contributed to the small number of effects found. For example, the duration and frequency of the stereotypes presented in the video provide one possible explanation for the fact that we found hardly any effects on girls' and boys' performance, motivational dispositions, and attitudes even though we found an effect on their stereotype endorsement. According to expectancy-value theory, it is through repeated experience that effects begin to accumulate and can lead to the internalization of gender-role stereotypes and to gender differences in expectancy and value beliefs in math in the end (Wigfield and Eccles, 2000; Eccles, 2009). In our study, we used a television program that was broadcast on national television to ensure that the experimental material was strongly linked to children's daily

TABLE 4 | Correlations between all Study Variables.

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	Stereotype endors. T1	—																			
2	Stereotype endors. T2	0.38	—																		
3	Performance T1	(0.11)	(0.09)	—																	
4	Performance T2	0.13	0.15	0.85	—																
5	Self-concept T1	(0.09)	(0.08)	0.36	0.35	—															
6	Self-concept T2	0.12	0.11	0.30	0.33	0.80	—														
7	Sense of belonging T1	(0.10)	(0.10)	0.14	0.19	0.49	0.49	—													
8	Sense of belonging T2	(0.05)	(0.09)	0.12	0.17	0.39	0.50	0.81	—												
9	Feeling thermo. T1	(0.07)	(0.01)	0.19	0.22	0.48	0.44	0.33	0.29	—											
10	Feeling thermo. T2	0.11	(0.04)	0.20	0.26	0.42	0.45	0.32	0.28	0.87	—										
11	Intrinsic value T1	0.11	(0.09)	0.15	0.18	0.75	0.72	0.53	0.46	0.51	0.51	—									
12	Intrinsic value T2	(0.05)	(0.07)	0.18	0.22	0.68	0.76	0.46	0.47	0.39	0.41	0.87	—								
13	Attainment value T1	(0.10)	(0.07)	(−0.05)	(−0.00)	0.29	0.22	0.30	0.25	0.22	0.17	0.33	0.25	—							
14	Attainment value T2	(0.06)	(0.05)	(−0.01)	(0.03)	0.24	0.27	0.28	0.32	0.18	0.19	0.29	0.32	0.70	—						
15	Utility v: daily life T1	0.15	(0.08)	(−0.06)	(−0.03)	0.25	0.22	0.32	0.26	0.23	0.21	0.38	0.30	0.42	0.33	—					
16	Utility v: daily life T2	(0.06)	(−0.01)	(−0.10)	(−0.07)	0.16	0.22	0.30	0.36	(0.11)	0.15	0.30	0.33	0.30	0.45	0.60	—				
17	Utility v: social T1	0.16	(0.07)	(0.10)	0.16	0.34	0.28	0.36	0.31	0.26	0.26	0.39	0.34	0.43	0.29	0.31	0.23	—			
18	Utility v: social T2	0.13	(0.03)	(0.09)	(0.07)	0.24	0.27	0.25	0.28	0.21	0.24	0.35	0.33	0.34	0.28	0.23	0.24	0.70	—		
19	Cost T1	(−0.01)	(−0.06)	−0.16	−0.19	−0.68	−0.68	−0.53	−0.47	−0.42	−0.36	−0.71	−0.68	−0.16	−0.18	−0.23	−0.21	−0.17	(−0.10)	—	
20	Cost T2	(−0.01)	(−0.05)	−0.11	−0.19	−0.59	−0.73	−0.45	−0.51	−0.31	−0.31	−0.56	−0.68	(−0.09)	−0.23	−0.18	−0.25	−0.15	(−0.10)	0.72	—

Nonsignificant correlations are displayed in parentheses; for all other correlations, $p < 0.05$.
Stereotype endors. = Stereotype endorsement; Feeling thermo. = Feeling thermometer; Utility v = Utility value.

TABLE 5 | Multiple regression models 1: effects on stereotype endorsement, performance, self-concept, sense of belonging, and feeling thermometer.

Predictor	Stereotype endorsement		Performance		Self-concept		Sense of belonging		Feeling thermometer	
	β	β 95% CI	β	β 95% CI	β	β 95% CI	β	β 95% CI	β	β 95% CI
Pretest	0.39***	[0.26, 0.52]	0.86***	[0.81, 0.91]	0.81***	[0.73, 0.89]	0.81***	[0.75, 0.87]	0.86***	[0.80, 0.92]
Gender (boys = 1)	0.29†	[0.03, 0.55]	0.10	[−0.06, 0.26]	−0.01	[−0.14, 0.11]	−0.20*	[−0.36,−0.04]	0.04	[−0.10, 0.18]
Condition (exp. = 1)	0.50***	[0.03, 0.74]	0.04[a]	[−0.10, 0.18]	0.03[a]	[−0.09, 0.15]	−0.10[a]	[−0.23, 0.02]	−0.12†	[−0.23, −0.01]
Gender × Condition	−0.28	[−0.58, 0.02]	−0.07	[−0.24, 0.09]	0.12	[−0.03, 0.26]	0.30**	[0.12, 0.49]	0.10	[−0.07, 0.27]
Effect of condition for boys	0.22*	[0.04, 0.40]	−0.03	[−0.15, 0.09]	0.14†	[0.00, 0.29]	0.20*	[0.04, 0.36]	−0.02	[−0.13, 0.10]

All continuous variables are standardized. CI = confidence interval; exp. = experimental condition.
[a]*We formulated a hypothesis for this effect prior to the analysis.*
†*$p < 0.10$.*
**$p < 0.05$.*
***$p < 0.01$.*
****$p < 0.001$.*

TABLE 6 | Multiple regression models 2: effects on intrinsic value, attainment value, utility value for daily life, and cost.

Predictor	Intrinsic value		Attainment value		Utility value: daily life		Cost	
	β	β 95% CI	β	β 95% CI	β	β 95% CI	β	β 95% CI
Pretest	0.87***	[0.80, 0.93]	0.71***	[0.62, 0.79]	0.62***	[0.52, 0.71]	0.71***	[0.62, 0.80]
Gender (boys = 1)	−0.03	[−0.17, 0.11]	−0.05	[−0.19, 0.09]	−0.09	[−0.32, 0.15]	−0.11	[−0.35, 0.12]
Condition (exp. = 1)	0.05	[−0.10, 0.19]	−0.02	[−0.14, 0.10]	0.03	[−0.17, 0.23]	−0.05	[−0.21, 0.11]
Gender × Condition	0.00	[−0.18, 0.18]	0.00	[−0.21, 0.21]	−0.03	[−0.31, 0.26]	0.05	[−0.19, 0.30]
Effect of condition for boys	0.05	[−0.09, 0.18]	−0.02	[−0.18, 0.15]	0.00	[−0.19, 0.19]	0.00	[−0.23, 0.23]

All continuous variables are standardized. CI = confidence interval; exp. = experimental.
****$p < 0.001$.*

TABLE 7 | Multiple-group multiple regression model: effects on social utility value.

Predictor	Utility value—social			
	Questionnaire first		Achievement test first	
	β	β 95% CI	β	β 95% CI
Pretest	0.65***	[0.60, 0.71]	0.76***	[0.70, 0.82]
Gender (boys = 1)	0.30**	[0.14, 0.47]	0.10	[−0.13, 0.34]
Condition (exp. = 1)	0.24†	[0.00, 0.48]	0.21†	[0.03, 0.40]
Gender × Condition	−0.88***	[−1.12, −0.64]	−0.08	[−0.38, 0.22]
Effect of condition for boys	−0.64**	[−0.98, −0.30]	0.14	[−0.05, 0.32]

All continuous variables are standardized. CI = confidence interval; exp. = experimental.
†*$p < 0.10$.*
***$p < 0.01$.*
****$p < 0.001$.*

life experiences. However, the stereotyped clip in this television program had a duration of only about 1 min, and the children in the experimental condition saw this clip only once. Thus, the duration and frequency of stereotype presentation might need to be increased in future studies to substantially affect girls' and boys' motivational dispositions.

Furthermore, when interpreting the results of the present study for girls and for boys, the specific age group of the participants should be taken into consideration. We investigated how stereotypes in a television program affect 5th graders because important processes in the development of children's expectancy and value beliefs and understanding of gender role behavior take place during that age period. Around the age of 10 years old, children become increasingly aware of how gender-stereotypical behavior might reflect social gender roles (for a review, see Leaper, 2015). In addition, children increasingly understand, interpret and integrate others' feedback and become more realistic in evaluating their strengths and weaknesses during their elementary school years (Wigfield et al., 2015). Such processes are believed to influence the development of children's expectancy and value beliefs (Wigfield et al., 2015).

We investigated how stereotypes experienced in the environment might influence students' motivational dispositions among 5th graders because children at that age should be right at the beginning of these developmental processes. In addition, previous research has indicated that even elementary school children can be affected by gender stereotypes—at least with respect to math performance (Flore and Wicherts, 2015). However, the participants' young age could be a reason why we found (almost) no effects on students' expectancy and value

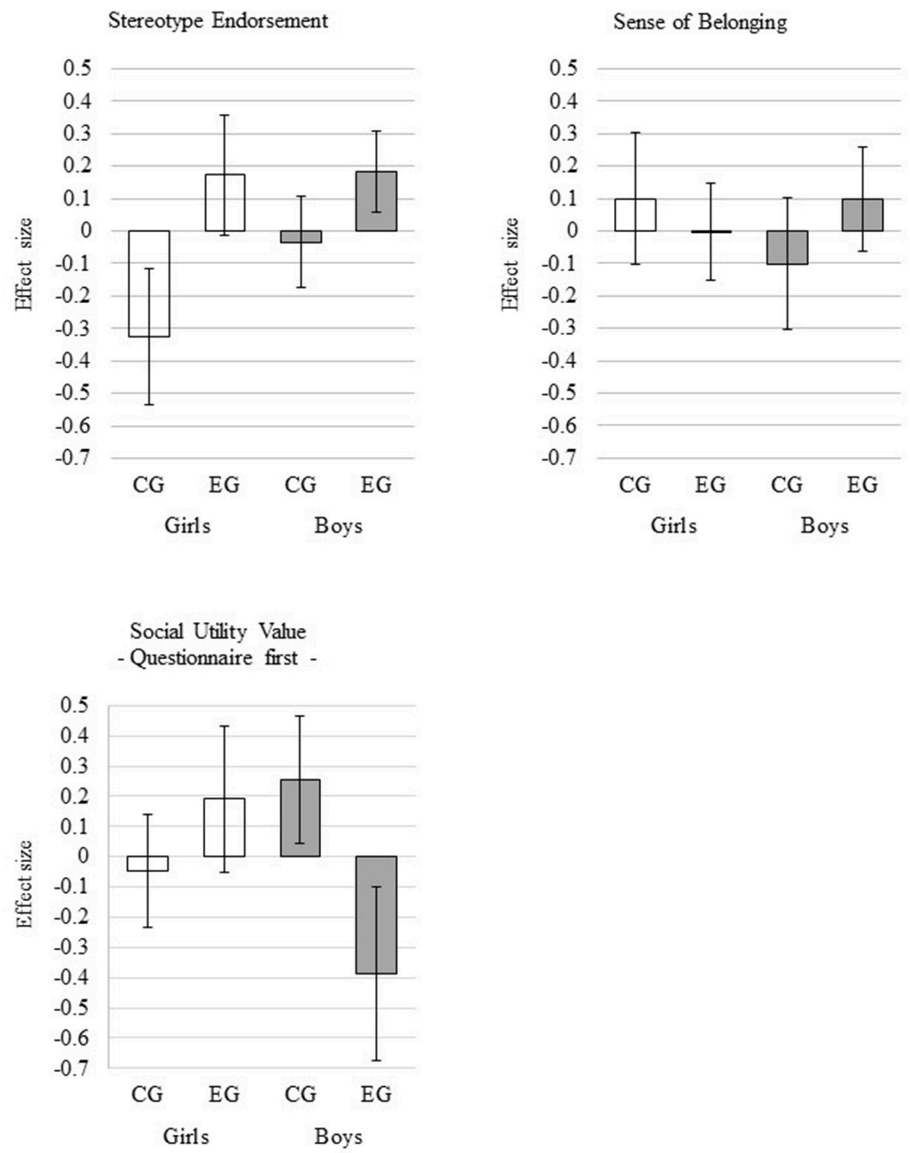


FIGURE 1 | Effects of the experimental manipulation. Error bars represent 95% confidence intervals. CG = control group; EG = experimental group.

beliefs. One reason for this assumption is provided by findings from the stereotype threat literature that have indicated that group and domain identification moderate effects of stereotype threat (e.g, Schmader, 2002; Lewis and Sekaquaptewa, 2016). Given that children increasingly identify with specific school subjects in elementary and middle school but do not differentiate much between the subjects at younger ages (see Wigfield et al., 2015), the participants in our study might have been too young and might not have sufficiently identified with the domain of math.

In addition, the stereotypes that were displayed in the video may provide an explanation for the fact that we did not find any effects on girls' motivational dispositions, attitudes, and performance in math and only a few effects on boys' motivational dispositions and attitudes. With respect to the girls in the video, it was not clear whether the girls in the video thought doing their math homework was boring or whether they were not able to solve the problems; thus, the video might have targeted the low motivation of these girls and not their low performance or talent in math, which has typically been the focus of studies that have investigated the effects of stereotype threat (see e.g., Nguyen and Ryan, 2008).

A video that more directly targets girls' lower performance or talent compared with boys might thus evoke stronger effects on girls' motivational dispositions and attitudes. Such a video might also evoke more positive effects on boys' motivational dispositions and attitudes, effects that would go against previous research that has indicated the experience of *stereotype lift* for male students in situations in which female students' disadvantage in math was made salient. Stereotype lift describes the effect of a boost for the non-targeted group in settings in which stereotypes are activated (e.g., for men after negative stereotypes of women's math performance have been presented; e.g., Walton and Cohen, 2003; Johnson et al., 2012). The positive effect on boys' sense of belonging could be an indication of effects of stereotype lift on this outcome due to the traditional gender stereotypes in the video such as the stereotype that boys are better at math than girls.

However, the negative effect on boys' social utility can hardly be explained by stereotype lift effects. Here, the specific portrait of the boy presented in the stereotyped clip could have played a role. Although the male classmate from whom the girls copied their homework seemed to be mathematically competent, he was also presented as geeky. To the best of our knowledge, effects of this stereotype have not yet been investigated. However, there is research on the stereotypes of math and science. Such research has indicated that favoring these subjects reduces students' perceived social competence and popularity. A study by Hannover and Kessels (2004) showed that students who admitted to liking science were judged as less popular, less attractive, less socially competent, and less integrated than students who claimed they did not like science. As the social utility scale directly referred to social acceptance, the stereotype of the boy as competent but geeky might thus explain the negative effect of the stereotype on boys' social utility.

Strength and Limitations

One major strength of this study is its high ecological validity. In our experiment, we used a television program that was broadcast on national television. Although the experiment took place in the school context, which does not exactly represent the setting in which children watch television programs in their everyday lives, the experimental material perfectly reflected what children encounter in real-world situations. Contrary to previous research on stereotypes, we furthermore investigated effects of stereotypes embedded into a more complex situation, where a lot of other information was presented to the children. Our results therefore provide initial insights into effects of stereotypes embedded in a television program on young girls and boys in a naturalistic setting. Nevertheless, further studies should also investigate such effects in other real-life settings, such as the home, where children might watch television programs together with their families and therefore might discuss the content of these programs.

In conducting the experiment, we applied a strong research design to address our research questions. We used a randomized block design, randomizing male and female students within classes to the different conditions. Thereby, we investigated possible effects on girls' and boys' performance as well as on different motivational dispositions and attitudes with the aim of obtaining a comprehensive picture of possible effects of traditional stereotypes in television programs. The sample size was based on a power analysis, and in order to increase the transparency of our research, we preregistered all of our hypotheses as well as the analyses. By doing so, we attempted to counter any arguments that might suggest that the effects of stereotype threat were built on *p-hacking* (Flore and Wicherts, 2015).

To assess possible effects of the stereotypes embedded in the television program, we included several different outcome measures such as scales for measuring all dimensions of the task values, for instance, or scales for assessing students' sense of belonging. The findings thus provide a comprehensive picture of possible effects on different outcomes, although one should keep in mind that the scale to assess students' sense of belonging was adapted from the original study. However, the measures we used were based on an achievement test and a questionnaire, which consisted of self-report measures. Our results thus provide no insights into how individuals might process the information presented in the video. Other assessment tools such as observational outcome measures (e.g., eye tracking) are necessary for investigating such processes.

The specific stereotypes transmitted in the television program also need to be considered when interpreting the results of our study. Whereas previous studies on stereotype threat mostly investigated stereotypes of girls being less able to do math than boys (see e.g., Nguyen and Ryan, 2008), the girls in the video might have only been too bored to do their math homework and the boy is depicted as being geeky. The effects on stereotype endorsement indicate that the children noticed the stereotype of boys being better in math than girls in the video. Nonetheless, it is still an open question if a video that more explicitly presents girls as being less able to do math than boys and boys not as being geeky would have caused effects on the other outcomes under investigation. For example, there is research indicating that favoring math and sciences reduces students' perceived social competence and popularity (Hannover and Kessels, 2004). Based on such findings, it can be speculated that the negative effect on social utility for boys found in the present study might be due to the presentation of the boy as being geeky in the video because the social utility scale directly referred to social acceptance. Additionally, it might be possible that the stereotype of the geeky math boy prevented girls from being negatively affected by the video because girls might have experienced this presentation as a negative stereotype against boys. However, such assumptions are rather speculative and further research is necessary to investigate whether other presentations of gender stereotypes affect girls and boys differently than those used in the present study.

Another limitation refers to the sample, which consisted of academic track students (students attending Gymnasiums). We used this sample because academic track schools are the most frequented type of school in Germany (more than 40% of students attend this type of school after primary school), and the school-leaving certificate from academic track schools entitles students to attend university (State Statistical Office of Baden-Württemberg., 2016). When investigating the influence of stereotypes on gender differences in important predictors of STEM careers, it is therefore most informative to assess samples of academic track students. Nevertheless, further research is required to investigate how the results can be generalized to students from other types of schools.

CONCLUSION

This study suggests that stereotypes in television can increase children's stereotype endorsement, but hardly affect their motivational dispositions, attitudes, and performance. Consequently, one could argue that traditional gender stereotypes presented in a television programs do not seem to affect young girls in math. This might be positive, particularly in light of the huge amount of time children spend watching television every day (Rideout et al., 2010; Rideout, 2015). However, in our study, we investigated effects of stereotypes in a television program in which only about 1 min of the material had been manipulated, and it might be repeated experience that causes effects to accumulate and sustainably affect boys and girls in the end (Wigfield and Eccles, 2000; Eccles, 2009). Additionally, even such a short clip containing stereotypes presented only once increased children's stereotype endorsement (at least in the short term). The results therefore suggest that television can activate and increase stereotypes about males' advantage in math in children, which might ultimately lead to gender differences in mathematically-intensive STEM fields (Eccles, 2009). Even though we did not find effects on children's motivational dispositions and attitudes, program developers might therefore nonetheless wish to carefully consider including stereotypes in television programs for children.

Our research adds to the discussion of the relevance of stereotype threat effects, particularly with respect to motivational dispositions (see Spencer et al., 2016). Despite effects of the experimental condition on girls' and boys' stereotype endorsement, we found hardly any effects on children's performance, motivational dispositions, and attitudes. Again, it might be repeated experience that renders effects of stereotype threat potentially harmful, and more research is needed to explore the duration of possible effects. Nevertheless, given failed attempts to replicate the original findings on stereotype threat (Stoet and Geary, 2012) and indications of publication bias in the literature on stereotype threat (Flore and Wicherts, 2015), the findings from the present study cast doubt on the robustness of stereotype threat effects. To continue this discussion, it is imperative that non-significant findings are not hidden away in the file drawer.

DATA AVAILABILITY STATEMENT

Datasets are available on request: The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of American Psychological Association with written informed consent from all subjects. All subjects and their parents gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Ethics Committee for Psychological Research of the University of Tübingen.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This research was funded by the LEAD Graduate School & Research Network [GSC1028], a project of the Excellence Initiative of the German federal and state governments. This research project was supported by the Postdoc Academy of the Hector Research Institute of Education Sciences and Psychology, Tübingen, funded by the Baden-Württemberg Ministry of Science, Research and the Arts. We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tübingen.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02435/full#supplementary-material

REFERENCES

Ambady, N., Shih, M., Kim, A., and Pittinsky, T. L. (2001). Stereotype susceptibility in children: effects of identity activation on quantitative performance. *Psychol. Sci.* 12, 385–390. doi: 10.1111/1467-9280.00371

Appel, M., and Weber, S. (2017). Do mass mediated stereotypes harm members of negatively stereotyped groups? a meta-analytical review on media-generated stereotype threat and stereotype lift. *Commun. Res.* doi: 10.1177/0093650217715543. [Epub ahead of print].

Aubrey, J. S., and Harrison, K. (2004). The gender-role content of children's favorite television programs and its links to their gender-related perceptions. *Media Psychol.* 6, 111–146. doi: 10.1207/s1532785xmep0602_1

Bond, B. J. (2016). Fairy godmothers and gt; robots. *Bull. Sci. Technol. Soc.* 36, 91–97. doi: 10.1177/0270467616655951

Borgers, N., Hox, J., and Sikkel, D. (2004). Response effects in surveys on children and adolescents: the effect of number of response options, negative wording, and neutral mid-point. *Qual. Quant.* 38, 17–33. doi: 10.1023/B:QUQU.0000013236.29205.a6

Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., and Latinotti, S. (2003). Stereotype threat: the effect of expectance on performance. *Eur. J. Soc. Psychol.* 33, 267–285. doi: 10.1002/ejsp.145

Cheryan, S., Plaut, V. C., Davies, P. G., and Steele, C. M. (2009). Ambient belonging: how stereotypical cues impact gender participation in computer science. *J. Pers. Soc. Psychol.* 97, 1045–1060. doi: 10.1037/a0016239

Cheryan, S., Plaut, V. C., Handron, C., and Hudson, L. (2013). The stereotypical computer scientist: gendered media representations as a barrier to inclusion for women. *Sex Roles* 69, 58–71. doi: 10.1007/s11199-013-0296-x

Collins, R. L. (2011). Content analysis of gender roles in media: where are we now and where should we go? *Sex Roles* 64, 290–298. doi: 10.1007/s11199-010-9929-5

Davies, P. G., Spencer, S. J., Quinn, D. M., and Gerhardstein, R. (2002). Consuming images: how television commercials that elicit stereotype threat can restrain women academically and professionally. *Pers. Soc. Psychol. Bull.* 28, 1615–1628. doi: 10.1177/014616702237644

Deci, E. L., and Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York, NY; Boston; Dordrecht; London; Moscow: Kluwer Academic/Plenum Publishers.

Denissen, J. J. A., Zarrett, N. R., and Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Dev.* 78, 430–447. doi: 10.1111/j.1467-8624.2007.01007.x

Doyle, R. A., and Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: a meta-analysis. *Learn. Individ. Diff.* 47, 103–116. doi: 10.1016/j.lindif.2015.12.018

Dubois, B., and Burns, J. A. (1975). An analysis of the meaning of the question mark response category in attitude scales. *Educ. Psychol. Meas*. 35, 869–884.

Eccles, J. S. (2005). "Subjective task values and the Eccles et al. model of achievement related choices," in *Handbook of Competence and Motivation*, edS A. J. Elliot and C. S. Dweck (New York, NY: Guilford), 105–121.

Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368

Eccles, J. S., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). "Expectancies, values, and academic behaviors," in *Achievement and Achievement Motivation,* ed J. T. Spence (San Francisco, CA: W. H. Freeman and Co), 75–121.

Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Feierabend, S., Plankenhorn, T., and Rathgeb, T. (2017). *KIM-Studie 2016: Kindheit, Internet, Medien [KIM-Study 2016: Childhood, internet, media]. Stuttgart*. Available online at: https://www.mpfs.de/fileadmin/files/Studien/KIM/2016/KIM_2016_Web-PDF.pdf

Flore, P. C., and Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *J. Sch. Psychol.* 53, 25–44. doi: 10.1016/j.jsp.2014.10.002

Fogliati, V. J., and Bussey, K. (2013). Stereotype threat reduces motivation to improve effects of stereotype threat and feedback on women's intentions to improve mathematical ability. *Psychol. Women Q.* 37, 310–324. doi: 10.1177/0361684313480045

Gaspard, H., Dicke, A.-L., Flunger, B., Häfner, I., Brisson, B., Trautwein, U., et al. (2016). Side effects of motivational interventions? Effects of an intervention in math classrooms on motivation in verbal domains. *AERA Open*2, 1–14. doi: 10.1177/2332858416649168

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., et al. (2015). More value through greater differentiation: gender differences in value beliefs about math. *J. Educ. Psychol.* 107, 663–677. doi: 10.1037/edu0000003

Good, J. J., Woodzicka, J., a, and Wingfield, L. C. (2010). The effects of gender stereotypic and counter-stereotypic textbook images on science performance. *J. Soc. Psychol.* 150, 132–147. doi: 10.1080/00224540903366552

Goodenow, C. (1993). The psychological sense of school membership among adolescents: scale development and educational correlates. *Psychol. Sch.* 30, 79–90.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., and Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychol. Rev.* 109, 3–25. doi: 10.1037/0033-295X.109.1.3

Hannover, B., and Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learn. Instruct.* 14, 51–67. doi: 10.1016/j.learninstruc.2003.10.002

Hermann, J. M., and Vollmeyer, R. (2016). Stereotype threat in der grundschule [stereotype threat in elementary school]. *Zeitschr. Entwicklungspsychol. Pädagogische Psychol.* 48, 42–49. doi: 10.1026/0049-8637/a000143

Heyman, K. (2008). Talk nerdy to me. *Science* 320, 740–741. doi: 10.1126/science.320.5877.740

Johnson, H. J., Barnard-Brak, L., Saxon, T. F., and Johnson, M. K. (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *J. Exp. Educ.* 80, 137–149. doi: 10.1080/00220973.2011.567312

Kaplan, K. J. (1972). On the ambivalence-indifference problem in attitude theory and measurement: a suggested modification of the semantic differential technique. *Psychol. Bull.* 77, 361–372. doi: 10.1037/h0032590

Kessels, U., Rau, M., and Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *Br. J. Educ. Psychol.* 76, 761–780. doi: 10.1348/000709905X59961

KiKa.de (2015). *Pur+. Mathe ist Magie [Pur+. Math is magic]*. Germany. Available online at: http://www.kika.de/pur/sendungen/sendung84728.html

Leaper, C. (2015). "Gender and social-cognitive development," in *Handbook of Child Psychology and Developmental Science*, ed R. M. Lerner (Hoboken, NJ: John Wiley and Sons, Inc), 1–48. doi: 10.1002/9781118963418.childpsy219

Lewis, N. A., and Sekaquaptewa, D. (2016). Beyond test performance: a broader view of stereotype threat. *Curr. Opin. Psychol.* 11, 40–43. doi: 10.1016/j.copsyc.2016.05.002

Marsh, H. W. (2007). *Self-Concept Theory, Measurement and Research Into Practice: The Role of Self-Concept in Educational Psychology*. Leicester: British Psychological Society.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., and Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Dev.* 76, 397–416. doi: 10.1111/j.1467-8624.2005.00853.x

Master, A., Cheryan, S., and Meltzoff, A. N. (2015). Computing whether she belongs: stereotypes undermine girls' interest and sense of belonging in computer science. *J. Educ. Psychol.* 108, 424–437. doi: 10.1037/edu0000061

Möller, J., and Marsh, H. W. (2013). Dimensional comparison theory. *Psychol. Rev.* 120, 544–560. doi: 10.1037/a0032459

Murphy, M. C., Steele, C. M., and Gross, J. J. (2007). Signaling threat. *Psychol. Sci.* 18, 879–885. doi: 10.1111/j.1467-9280.2007.01995.x

Muthén, B. O., and Muthén, L. K. (2012). *Mplus User's Guide, 7th Edn.* Los Angeles, CA: Muthén and Muthén.

Muzzatti, B., and Agnoli, F. (2007). Gender and mathematics: attitudes and stereotype threat susceptibility in Italian children. *Dev. Psychol.* 43, 747–759. doi: 10.1037/0012-1649.43.3.747

Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T., and Trautwein, U. (2011). Who took the "×" out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychol. Sci.* 22, 1058–1066. doi: 10.1177/0956797611415540

National Reserach Council. (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. Washington, DC: The National Academic Press.

National Science Board (2016). *Science and Engineering Indicators 2016*.

National Science Foundation (2015). *Women, Minorities, and Persons With Disabilities in Science and Engineering 2015*. Special Report NSF.

Neuville, E., and Croizet, J.-C. (2007). Can salience of gender identity impair math performance among 7-8 years old girls? The moderating role of task difficulty. *Eur. J. Psychol. Educ.* 22, 307–316. doi: 10.1007/BF03173428

Nguyen, H.-H. D., and Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *J. Appl. Psychol.* 93, 1314–1334. doi: 10.1037/a0012702

Passolunghi, M. C., Rueda Ferreira, T. I., and Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learn. Individ. Differ.* 34, 70–76. doi: 10.1016/j.lindif.2014.05.005

Patten, D. (2013). *Full 2012-2013 TV Season Series Rankings*. Available online at: http://deadline.com/2013/05/tv-season-series-rankings-2013-full-list-506970/ (Accessed February 23, 2017).

Plante, I., De la Sablonnière, R., Aronson, J. M., and Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: the role of competence beliefs and task values. *Contemp. Educ. Psychol.* 38, 225–235. doi: 10.1016/j.cedpsych.2013.03.004

Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychol. Methods* 2, 173–185. doi: 10.1037/1082-989X.2.2.173

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. -,f., Martinez, A., and Bloom, H. S. (2011). *Optimal Design Software for Multi-Level and Longitudinal Research* (Version 3.01). Available online at: www.wtgrantfoundation.org

Reilly, D., Neumann, D. L., and Andrews, G. (2015). Sex differences in mathematics and science achievement: a meta-analysis of national assessment of educational progress assessments. *J. Educ. Psychol.* 107, 645–662. doi: 10.1037/edu0000012

Renninger, K. A., and Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educ. Psychol.* 46, 168–184. doi: 10.1080/00461520.2011.587723

Rideout, V. J. (2015). *The Common Sense Census: Media Use by Tweens and Teens*. San Francisco, CA: Common Sense Media.

Rideout, V. J., Foehr, U. G., and Roberts, D. F. (2010). *Media in the Lives of 8-18 Year Olds*. Menlo Parc, CA: Kaiser Family Foundation Study.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *J. Exp. Soc. Psychol.* 38, 194–201. doi: 10.1006/jesp.2001.1500

Schmader, T., Johns, M., and Barquissau, M. (2004). The costs of accepting gender differences: the role of stereotype endorsement in women's experience in the math domain. *Sex Roles* 50, 835–850. doi: 10.1023/B:SERS.0000029101.74557.a0

Schoon, I., and Eccles, J. S. (2014). *Gender Differences in Aspiration and Attainment: A Life Course Perspective*. Cambridge: Cambridge University Press.

Schunk, D. H., Pintrich, P. R., and Meece, J. L. (2008). *Motivation in Education: Theory, Research and Applications , 3rd Edn*. Upper Saddle River, NJ: Pearson.

Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inferences*. Berkeley, CA: Houghton Mifflin.

Signorella, M. L., Bigler, R. S., and Liben, L. S. (1993). Developmental differences in children′ s gender schemata about others: a meta-analytic review. *Dev. Rev.* 13, 147–183.

Smith, J. L., Brown, E. R., Thoman, D. B., and Deemer, E. D. (2015). Losing its expected communal value: how stereotype threat undermines women's identity as research scientists. *Soc. Psychol. Educ.* 18, 443–466. doi: 10.1007/s11218-015-9296-8

Smith, J. L., Sansone, C., and White, P. H. (2007). The stereotyped task engagement process: the role of interest and achievement motivation. *J. Educ. Psychol.* 99, 99–114. doi: 10.1037/0022-0663.99.1.99

Spencer, S. J., Logel, C., and Davies, P. G. (2016). Stereotype threat. *Annu. Rev. Psychol.* 67, 415–437. doi: 10.1146/annurev-psych-073115-103235

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35, 4–28. doi: 10.1006/jesp.1998.1373

State Statistical Office of Baden-Württemberg. (2016). *Pressemitteilung. Zahl der Übergänge auf Weiterführende Schulen zum Schuljahr 2015/16 bleibt Stabil.* [Press release. Number of transitions to secondary schools remains stable in school year 2015/2016]. Available online at: https://www.statistik-bw.de/Presse/Pressemitteilungen/2016020 (Accessed January 30, 2018).

Steele, C. M., and Aronson, J. M. (1995). Stereotype threat and the intellectual test performance of African Americans. *J. Pers. Soc. Psychol.* 69, 797–811. doi: 10.1037/0022-3514.69.5.797

Steffens, M. C., Jelenec, P., and Noack, P. (2010). On the leaky math pipeline: comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *J Educ. Psychol.* 102, 947–963. doi: 10.1037/a0019920

Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., et al. (2007). Assessing media influences on middle school aged children's perceptions of women in science using the draw-a-scientist test (DAST). *Sci. Commun.* 29, 35–64. doi: 10.1177/1075547007306508

Stoet, G., and Geary, D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement? *Rev. Gen. Psychol.* 16, 93–102. doi: 10.1037/a0026617

Thoman, D. B., Smith, J. L., Brown, E. R., Chase, J., and Lee, J. Y. K. (2013). Beyond performance: a motivational experiences model of stereotype threat. *Educ. Psychol. Rev.* 25, 211–243. doi: 10.1007/s10648-013-9219-1

Tomasetto, C., Alparone, F. R., and Cadinu, M. (2011). Girls' math performance under stereotype threat: the moderating role of mothers' gender stereotypes. *Dev. Psychol.* 47, 943–949. doi: 10.1037/a0024047

van't Veer, A. E., and Giner-Sorolla, R. (2016). Pre-registration in social psychology—a discussion and suggested template. *J. Exp. Soc. Psychol.* 67, 2–12. doi: 10.1016/j.jesp.2016.03.004

Voyer, D., and Voyer, S. D. (2014). Gender differences in scholastic achievement: a meta-analysis. *Psychol. Bull.* 140, 1174–1204. doi: 10.1037/a0036620

Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., and Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspect. Psychol. Sci.* 7, 632–638. doi: 10.1177/1745691612463078

Walton, G. M., and Cohen, G. L. (2003). Stereotype lift. *J. Exp. Soc. Psychol.* 39, 456–467. doi: 10.1016/S0022-1031(03)00019-2

Wang, M.-T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001

Watt, H. M. G., and Eccles, J. S. (2008). *Gender and Occupational Outcomes: Longitudinal Assessments of Individual, Social, and Cultural Influences*. Washington, DC: American Psychological Association.

Wigfield, A., and Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., and Schiefele, U. (2015). "Development of achievement motivation and engagement," in *Handbook of Child Psychology and Developmental Science,* 7th Edm, eds R. M. Lerner, C. Garcia Coll and M. E. Lamb (Hoboken, NJ: John Wiley and Sons, Inc.), 1–44.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 17 December 2018
doi: 10.3389/fpsyg.2018.02563

Sex Differences in Gains Among Hispanic Pre-kindergartners' Mental Rotation Skills

Carla Abad[1], Rosalie Odean[2] and Shannon M. Pruden[3]*

[1] STEM Transformation Institute, Florida International University, Miami, FL, United States, [2] School of Education, University of Delaware, Newark, DE, United States, [3] Department of Psychology, Florida International University, Miami, FL, United States

The current study explores change in mental rotation skills throughout the pre-kindergarten year in a Hispanic population to better understand the development of early sex differences in mental rotation. Ninety-six Hispanic children (M = 4 years 8 months) completed a mental rotation task at the beginning and end of pre-kindergarten. Results suggest Hispanic boys and girls differed in gains on mental rotation ability, with boys improving significantly more than girls during pre-kindergarten on a mental rotation task. This study highlights the significance of studying mental rotation abilities in a Hispanic population of pre-kindergarten aged children and suggests the importance of examining sex differences in mental rotation over time, rather than at one time-point, to better understand when sex differences in spatial skills develop. We discuss various factors that potentially affect the growth of spatial skills including the role of early education, spatial experiences, and spatial language input.

Keywords: spatial thinking, preschool, sex differences, Hispanic children, mental rotation

OPEN ACCESS

Edited by:
Silke Luttenberger, University College of Teacher Education Styria, Austria

Reviewed by:
Juan Cristobal Castro-Alonso, Universidad de Chile, Chile
David Stuart Smith, Robert Gordon University, United Kingdom
Uner Tan, Çukurova University, Turkey

***Correspondence:**
Carla Abad
caabad@fiu.edu

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: *16 September 2018*
Accepted: *29 November 2018*
Published: *17 December 2018*

Citation:
Abad C, Odean R and Pruden SM (2018) Sex Differences in Gains Among Hispanic Pre-kindergartners' Mental Rotation Skills. Front. Psychol. 9:2563. doi: 10.3389/fpsyg.2018.02563

INTRODUCTION

Spatial thinking is the ability to think about the spatial world and encompasses a number of skills including mentally rotating and transforming objects and shapes, recreating patterns, and navigating around one's environment (Sinton et al., 2013). Children and adults depend on spatial thinking for a variety of everyday situations such as remembering the location of a doll in a play room or a car in the parking lot, fitting toys in a box or suitcases in a trunk, and building block towers or Ikea furniture (Abad, 2018). Aside from being necessary for everyday tasks, spatial thinking is linked to early mathematics ability (Cheng and Mix, 2014; Verdine et al., 2017) and predicts future entry in Science, Technology, Engineering, Mathematics (STEM) fields (Humphreys et al., 1993; Shea et al., 2001; Wai et al., 2009).

Several studies have established sex differences in adults' spatial skills, particularly mental rotation ability, with men consistently outperforming women (Maccoby and Jacklin, 1974; Linn and Petersen, 1985; Voyer et al., 1995). However, when these sex differences develop remains uncertain since sex differences in children's spatial skills are inconsistent (Frick et al., 2014). Previous studies typically examine sex differences at one timepoint in mostly middle- to upper-income populations of primarily White individuals leaving open the question of whether we see sex differences in spatial thinking over time and in underrepresented populations. The projections that Hispanics will make up 28 percent of the United States population by the year 2050 (Colby and Ortman, 2015) and the lack of minority representation in STEM fields (National Science Foundation and National Center for Science and Engineering Statistics, 2013) reveal the importance of exploring sex differences in a Hispanic population. In a recent publication Levine et al. (2016) attempted to "advance the conversation" on sex differences in spatial cognition by laying out the need for more research examining change in development. The current study seeks to fill these gaps in the literature and "advance the conversation" by examining whether there are sex differences in spatial skills *over time* in a typically understudied population, Hispanic pre-kindergarten (pre-k) children of diverse socioeconomic status. Specifically, the current study aims to explore whether changes in mental rotation ability made by Hispanic boys throughout pre-kindergarten are different from changes made by Hispanic girls.

Sex Differences in Spatial Thinking

Research over four decades suggests consistent sex differences in spatial thinking, with men reliably outperforming women on some spatial thinking tasks and the largest effects found on tasks requiring mental rotation ability (Maccoby and Jacklin, 1974; Linn and Petersen, 1985; Voyer et al., 1995; Uttal et al., 2013). However, *when* and *how* these sex differences emerge are more contentious subjects (see review by Levine et al., 2016 as well as Frick et al., 2014).

Studies on sex differences in spatial skills across the lifespan have resulted in inconsistent findings. Consistent with the adult literature, some studies have found an early male advantage on spatial skills (e.g., Johnson and Meade, 1987; Levine et al., 1999; Levine et al., 2005; Ehrlich et al., 2006; Casey et al., 2008; Moore and Johnson, 2008; Quinn and Liben, 2008, 2014; Levine et al., 2012; Lauer et al., 2015). For instance, a study by Johnson and Meade (1987) on children between 6 and 18 years showed boys outperform girls on spatial tasks by age 10. Exploring a younger population, Levine et al. (1999) tested children ranging from 4 to almost 7 years of age on a mental rotation task and found sex differences with boys outperforming girls as early as 4.5 years of age. Sex differences on spatial tasks in 3–4.5 year olds have been replicated in other studies (Levine et al., 2012; Joh, 2016; Pruden and Levine, 2017).

However, studies examining children's spatial thinking have found no consistent sex differences, even on tasks requiring mental rotation skills where the strongest sex differences are found in adults (e.g., Platt and Cohen, 1981; Kaplan and Weisberg, 1987; Caldwell and Hall, 1970; Kaess, 1971; Jahoda, 1979; Kosslyn et al., 1990; Estes, 1998; Lachance and Mazzocco, 2006; Frick et al., 2009; Krüger and Krist, 2009; Jansen and Heil, 2010; Frick et al., 2013; Lehmann et al., 2014; Verdine et al., 2017). For instance, a study by Manger and Eikeland (1998) on sixth graders' spatial visualization skills found no significant sex differences. Frick et al. (2013) explored the performance of children between the ages of 3 and 5 on a mental rotation task and found no consistent sex differences. More recently, Verdine et al. (2017) assessed the spatial thinking abilities of 3 to 5-year-old children on a variety of spatial tasks and found no significant sex differences. Furthermore, due to the file-drawer problem, it is possible that numerous other studies showing no significant sex differences in children's spatial thinking remain unpublished (Rosenthal, 1979).

Sex differences in adults' spatial ability are well established, however, *when* and *how* these sex differences emerge remains uncertain. Biological, hormonal, and evolutionary accounts still permeate the debate, however, environmental factors have been shown to influence and potentially mediate sex differences in spatial skills including: (1) boys are engaged in more activities related to spatial and mathematics achievement than girls (Newcombe et al., 1983; Baenninger and Newcombe, 1995; Nazareth et al., 2013); (2) boys and girls are held to different expectations and standards (i.e., gender stereotypes) by their parents and teachers (Parsons et al., 1982; Eccles and Jacobs, 1986); (3) girls have more anxiety regarding their performance on spatial activities (Lawton, 1994; Baenninger and Newcombe, 1995); and (4) boys hear more spatial language than girls from their parents (Pruden and Levine, 2017).

While no single explanation accounts for the sex differences found in spatial thinking and the timing of the emergence of these sex differences is still debated, it is clear that sex differences exist and are influenced by many environmental factors. The bulk of previous research has addressed sex differences in spatial thinking in middle- to upper-income populations containing primarily non-Hispanic White individuals. However, to better understand sex differences in spatial thinking, it is critical to explore whether these differences exist across populations.

Generalizability of Sex Differences in Spatial Thinking

Studies investigating whether sex differences in spatial thinking are generalizable across diverse populations show conflicting results. Several studies in African, Asian, and Western cultures suggest the male advantage exists across cultures in both child and adult samples (e.g., Jahoda, 1980; Mann et al., 1990; Lynn, 1992; Silverman et al., 2007; Casey et al., 2008; Lippa et al., 2010; Liu and Lynn, 2011). However, other studies utilizing cross-cultural and diverse populations suggest sex differences in spatial thinking may not be generalizable across all populations. For instance, Feingold (1994) examined sex differences in studies conducted after 1980 with participants from outside of the United States and found no consistent sex differences in verbal, math, or spatial skills across cultures. Berry (1966) examined the spatial ability of Eastern Canadian Eskimos from the Baffin Islands and found no sex differences on a variety of spatial assessments. More recently, Icelandic high school girls were found to outperform their male peers on highly spatial sections of a mathematics test (Lemke et al., 2004). Additionally, socioeconomic status was found to mediate sex differences in mental rotation (Levine et al., 2005). Levine and colleagues found sex differences in the mental rotation skills of boys and girls from middle and high SES but no sex differences in the mental rotation skills of children from low SES.

These studies suggest sex differences in spatial skills are generalizable across some nations, cultures, ethnicities, and socioeconomic statuses but may not be universal. Little research to date (though see Casey et al., 2008; Nazareth et al., 2013) has looked at whether there are similar sex differences in Hispanic individuals across varying socioeconomic groups – the aim of the present study. Importantly, no studies have investigated sex differences in spatial thinking in an exclusively Hispanic population of children within the United States (US). Given the growing Hispanic population in the United States and the current underrepresentation of Hispanic women in STEM fields, it is critical to examine whether sex differences in spatial thinking are generalizable to this particular population.

Changes in Spatial Skill

Numerous studies have established that spatial thinking is malleable and can be improved through training in both males and females (Ehrlich et al., 2006; Terlecki et al., 2008; Wright et al., 2008). Additionally, studies with multiple timepoints provide a greater understanding of sex differences in spatial thinking by examining whether these differences change over time. However, these studies generally examine change in spatial thinking after a specific intervention, with few studies investigating naturally occurring changes in spatial thinking in males and females throughout development (e.g., Huttenlocher et al., 1998; Levine et al., 2005; Lachance and Mazzocco, 2006; Verdine et al., 2017).

Huttenlocher et al. (1998) conducted a cross-sectional study where kindergartners and first graders were tested during school months and summer break on several cognitive tasks. Emphasizing the impact of early education for the development of spatial skills, children were found to grow significantly more during school months compared to vacation months on cognitive tasks related to language and spatial operations. While this study examines change in spatial thinking throughout early development, its cross sectional design leaves the question of whether boys and girls make similar or different changes throughout the school year unanswered. In a longitudinal study, Levine and colleagues examined the influence of SES and sex on second and third graders' spatial skills. Children made improvements over time on all spatial tasks measured, however, there were no reported differences in spatial ability by sex or SES over time. A different longitudinal study by Verdine et al. (2017) on the spatial abilities of children between the ages of 3 and 5 found no sex differences in preliminary analysis and therefore did not examine sex differences over time. Another longitudinal study (Lachance and Mazzocco, 2006) where over 100 students were followed from kindergarten to third grade to examine sex differences in math and spatial skills found no persistent sex differences during any year of the study, in any area of math or spatial skills, or in growth rates for math or spatial skills. These studies look at change in spatial skills over time in a naturalistic setting, however, none included an analysis of sex differences in change over time in preschool aged children, a time when many children enter formal schooling and sex differences may be emerging (Levine et al., 1999, 2012; Joh, 2016).

Given the lack of consensus on the age at which sex differences in spatial skills emerge and the fact that spatial skills develop over time and sex differences in spatial skills strengthen over time (Voyer et al., 1995), examining naturally occurring change in boys' and girls' spatial skills over time may provide a greater understanding of when these sex differences develop not available through studies with only one timepoint and/or cross-sectional designs. The current study aims to explore sex differences in change on mental rotation throughout the preschool year to better understand the development of early sex differences in spatial skills.

The Current Study

In sum, prior research finds that sex differences in spatial skills exist (e.g., Maccoby and Jacklin, 1974; Linn and Petersen, 1985; Voyer et al., 1995; Uttal et al., 2013), are generalizable across some populations (e.g., Jahoda, 1980; Mann et al., 1990; Lynn, 1992; Silverman et al., 2007; Casey et al., 2008; Liu and Lynn, 2011), are malleable (e.g., Baenninger and Newcombe, 1989; Uttal et al., 2013), and are influenced by environmental factors (e.g., Newcombe et al., 1983; Baenninger and Newcombe, 1995; Levine et al., 2012; Nazareth et al., 2013). However, little is known regarding change in sex differences in mental rotation ability over time and whether sex differences generalize to an all-Hispanic population. The current study seeks to address this gap by following Hispanic children throughout pre-k to assess early sex differences in mental rotation skills. Specifically, the current study has two aims: (1) to examine whether sex differences in mental rotation skills exist in Fall (time 1) and Spring (time 2) semesters of pre-kindergarten and (2) to explore whether *changes* in mental rotation skills of Hispanic pre-k boys are different from *changes* in mental rotation skills of Hispanic pre-k girls.

MATERIALS AND METHODS

Participants

The sample consisted of 96 children (45 boys; mean age at time 1 = 56 months; SD = 3.69 months) from 27 classrooms (20 schools) enrolled in Florida's state funded pre-k program at private schools. Participants were part of a larger study examining the role of educator language on the development of various spatial skills. One child was excluded due to a diagnosed developmental delay. All children were from Hispanic families in South Florida (16 families or 16.7% refused to report ethnicity) and were Spanish/English bilinguals. An English and Spanish language screener was administered as an additional check that they were being raised in bilingual homes (see section "Materials' and Methods" for a description of the language screener). Socioeconomic status (SES), which has been shown to mediate sex differences in spatial skills (Levine et al., 2005) was diverse, with families reporting variability in two indicators of SES, income and education levels. Given the correlation between family gross income and highest degree of education, $r(94) = 0.630$, $p < 0.001$, gross income was used as a proxy for SES. Eight participants reported earning $100,000 or more a year (8.3%), 5 earning between $75,000 to $99,999 (5.2%), 19 earning between $50,000 and $74,999 (19.8%), 13 earning between $35,000 and $49,999 (13.5%), 18 earning between $15,000 and $34,999 (18.8%), and 13 earning less than $15,000 a year (13.5%). Twenty families did not report gross income (20.8%).

Materials and Procedures

Consent and Demographics

Participants were recruited via letters sent from schools to parents. Interested families returned a signed consent form and received a demographics questionnaire regarding the child's race and ethnicity, primary caregiver's highest level of education, and family gross income.

Language Comprehension Screener

Children with parental consent were administered a brief language comprehension task in both English and Spanish during the first school visit. The screener included five questions in English (i.e., *what is your name, point to the cat, how old are you, show me your nose, what is your favorite color*) and Spanish (i.e., *cómo te llamas, señala el gato, cuántos años tienes, enseñame tu nariz, cuál es tu color favorito*) intended to assess basic language comprehension in both languages. Children were tested in a random order and the first two boys and two girls from each classroom to answer a minimum of four out of five items correctly on both language screeners were included in the study. Only four children from each classroom were selected to ensure a balanced number of children from each classroom and to limit classroom disruption. Forty-four children were excluded from the study for failing the screener in either English or Spanish. This confirmed children were able to understand English for our English-administered mental rotation assessment and were bilingual.

Assessments

Participants completed the *Peabody Picture Vocabulary Test*, *Test de Vocabulario en Imágenes Peabody*, and *Children's Mental Transformation Task* at both Fall and Spring semesters. The average time lag between time 1 and time 2 assessments was approximately four and a half months. Children were tested individually at their preschool and were given a sticker at the end of each testing session as a reward.

Receptive vocabulary

Children completed a measure of receptive vocabulary in English (*Peabody Picture Vocabulary Test, 4th ed* [PPVT]; Dunn and Dunn, 1997) and Spanish (*Test de Vocabulario en Imagenes Peabody* [TVIP]; Dunn et al., 1986) twice during the school year. These measures served as a proxy of children's verbal intelligence; standardized scores for both the PPVT and TVIP were not significantly correlated [$r(94) = -0.195$, $p = 0.07$] and were both included as control variables in analyses. Since bilingual children vary in their relative strength in each language spoken, we found it to be important to include both measures to accurately represent children's vocabulary skills. For each test item, the experimenter asked the child to point to a picture from a set of four pictures (e.g., "point to feather"). Each assessment took approximately 10–15 min to administer. Scores on both the PPVT and TVIP were age-based standardized scores with a mean score of 100 and a standard deviation of 15.

Children's mental transformation task

An abbreviated version of the *Children's Mental Transformation Task* (CMTT; Levine et al., 1999) used by Pruden et al. (2011) was administered at each timepoint. This task evaluates children's ability to mentally rotate and translate two shapes to make a whole object. The CMTT is different from classic embedded figures task in that it requires both rotation and transformation of object parts to form a whole rather than simply identifying parts within a whole. On each of 10 items, children were shown two pieces of shapes and four target shapes, and were asked to point to the shape that the two pieces would make if they were put together (**Figure 1**). Every correct response received 1 point with a possible score range of 0 to 10 points. On average, the CMTT took 5 min to complete and children were administered all 10 items.

RESULTS

Normality, Outliers, and Missing Data

Prior to analysis, SES, child's age at the time of each assessment, and scores on the PPVT, TVIP, and CMTT at each timepoint were examined for normality as well as univariate and multivariate outliers. Histograms were examined for univariate outliers and violations of normality (Tabachnick and Fidell, 2013). No univariate outliers were found in any variables tested. However, several variables were found to be skewed; this was addressed by using bootstrapping in subsequent analyses. No multivariate outliers were identified by using Mahalanobis distance with $p < 0.001$ (Yuan and Hayashi, 2010). Little's Missing Completely At Random (MCAR) test was not significant ($X^2 = 767.92$, $df = 816$, $p = 0.884$) suggesting data were missing at random (Tabachnick and Fidell, 2013). Less than 13% of data were missing, missing data were addressed by conducting multiple imputations using five imputations (Schafer and Graham, 2002). Reported results are from analyses conducted utilizing pooled data from the five imputations.

Descriptive Statistics

Descriptive statistics for children's performance on the CMTT and the receptive vocabulary measures (PPVT; TVIP) show considerable variability at each timepoint and in both sexes (**Table 1**). Average assessment scores suggest no floor or ceiling effects for any of the assessments, and ranges suggest variability in children's performance.

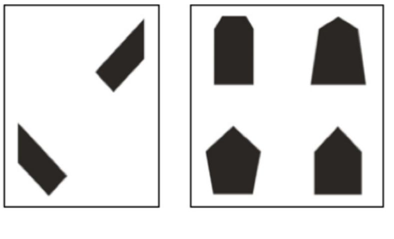


FIGURE 1 | Example from the Children's Mental Transformation Task (CMTT). Children are shown two pieces of shapes and four shapes and are asked to "*Point to the shape the pieces make*."

TABLE 1 | Descriptive statistics for scores at time 1 and time 2.

	Time 1						Time 2					
					Males	Females					Males	Females
	M	*SD*	Min	Max	*M* (SD)	*M* (SD)	*M*	*SD*	Min	Max	*M* (SD)	*M* (SD)
CMTT	4.04	2.54	0.00	9.00	4.45 (2.43)	3.68 (2.61)	5.69	2.07	0.00	10.00	5.58 (2.43)	5.78 (1.72)
PPVT	89.60	14.74	59.00	117.00	87.92 (15.14)	91.23 (14.34)	93.79	15.06	59.00	120.00	94.31 (14.58)	93.36 (15.60)
TVIP	90.23	17.58	55.00	135.00	90.21 (16.41)	90.25 (18.73)	91.13	19.66	55.00	139.00	90.71 (18.79)	91.47 (20.56)

Scores in this table were not imputed.

Main Analyses

Two multiple linear regressions and an ANCOVA style linear regression were conducted with Mplus version 7.31 to assess whether child sex was predictive of performance on the CMTT at time 1 and time 2 (aim 1) and to examine whether child sex was predictive of *changes* in CMTT scores throughout pre-kindergarten (aim 2).

Aim 1. A multiple linear regression was performed at each of the two timepoints to examine whether there were sex differences in mental rotation scores at time 1 (see **Table 2**) and time 2 (see **Table 2**); each regression controlled for SES, age at the time of assessment, and receptive vocabulary (PPVT and TVIP) scores. Results suggest there were no significant sex differences in Hispanic pre-kindergartners' CMTT scores at the time 1 (CMTT: $b = -0.65$, $\beta = -0.13$, $p = 0.18$, $R^2 = 0.201$) or time 2 (CMTT: $b = 0.28$, $\beta = 0.07$, $p = 0.51$, $R^2 = 0.134$). Given the bilingual population of this study, English and Spanish receptive vocabulary scores were both included as controls; however, the results held when controlling for only English and only Spanish scores.

Aim 2. An ANCOVA style linear regression was performed on CMTT scores (see **Table 3**); ANCOVA style linear regressions can be used to determine change with two timepoints by using statistical controls to investigate change over time (Newsom, 2012). An ANCOVA style regression allows for similar conclusions as an ANOVA while permitting the inclusion of additional continuous variables to be controlled (Cohen et al., 2003).

The second aim of the study was to explore whether there are sex differences in the *changes* Hispanic children make in mental rotation skills during prekindergarten. Results show that the interaction between child's sex and CMTT scores at time 1 significantly predicted CMTT scores at time 2 ($b = -0.41$, $\beta = -0.37$, $p = 0.02$, $R^2 = 0.234$), suggesting there were significant sex differences in *gains* made throughout pre-k on this task. Specifically, boys improved 0.41 points more than girls on the CMTT throughout the school year. English and Spanish receptive vocabulary scores were both included as controls; however, results held when controlling for only English or only Spanish scores. These results suggest boys are experiencing significantly greater gains than girls on the CMTT (see **Figure 2**). Furthermore, CMTT scores at time 1 were found to be a significant positive predictor of CMTT scores at time 2 ($b = -0.41$, $\beta = 0.50$, $p = 0.011$, $R^2 = 0.234$), suggesting improvement in CMTT scores throughout pre-k when boys and girls scores are combined.

It is important to note that results from the ANCOVA style linear regressions seem to contradict descriptive statistics previously reported, where the mean score on the CMTT for girls increased more than the mean score on the CMTT for boys across timepoints (**Table 1**); however, simply looking at mean scores does not take into account individual differences in performance. In fact, examining standard deviations for mean scores shows that while there is a similar amount of variability in boys' and girls' scores at time 1, there is much greater variability in boys' scores compared to girls' scores at time 2. Furthermore, simply comparing mean scores does not control for any potential confounds. The literature points to SES, age, and verbal IQ as strong predictors of

TABLE 3 | ANCOVA analyses predicting CMTT at time 2 for imputed pooled scores.

Variable	*B*	SE (*B*)	β	*p*
Constant	−0.488	4.326	−0.237	0.910
Time 1 score	0.410	0.162	0.496	0.011*
Child's sex	0.431	0.433	0.102	0.320
Child's sex * time 1	−0.408	0.177	−0.367	0.021*
Family income	0.191	0.150	0.149	0.203
PPVT	0.023	0.017	0.178	0.179
TVIP	0.010	0.010	0.089	0.348
Age	0.040	0.060	0.068	0.502

*$R^2 = 0.234$; *$p < 0.05$. CI, confidence interval; LL, lower limit; UL, upper limit.*

TABLE 2 | Regression analyses predicting CMTT at time 1 for imputed pooled scores.

	Time 1				Time 2			
Variable	*B*	SE (*B*)	β	*p*	*B*	SE (*B*)	β	*p*
Constant	−17.89	3.49	−7.07	0.00	−8.51	4.29	−4.06	0.05
Child's sex	−0.65	0.49	−0.13	0.18	0.28	0.43	0.07	0.51
Family income	−0.07	0.17	−0.05	0.66	0.23	0.17	0.18	0.18
PPVT	0.04	0.02	0.24	0.05*	0.03	0.02	0.20	0.13
TVIP	0.02	0.01	0.14	0.16	0.02	0.01	0.15	0.12
Age	0.23	0.06	0.32	0.00**	0.07	0.06	0.11	0.26

*Time 1 $R^2 = 0.201$; Time 2 $R^2 = 0.134$; *$p < 0.05$ and **$p < 0.01$.*

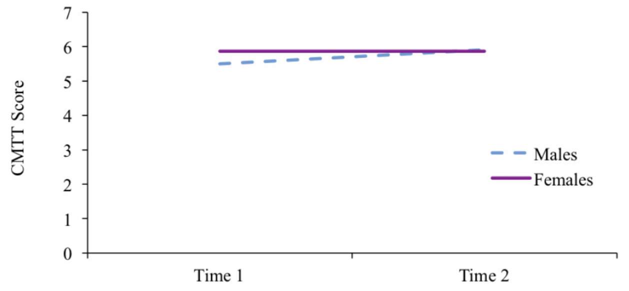


FIGURE 2 | Gains in CMTT scores throughout pre-kindergarten by child's sex based on regressions including age at the time of assessment, family gross income, and English and Spanish receptive vocabulary scores as covariates with mean covariates values.

CMTT scores (e.g., Levine et al., 2005), therefore, in order to accurately assess growth in CMTT it is necessary to control for these variables. Since our control variables are continuous variables, but we are interested in change over time and sex differences, which are categorical variables, ANCOVA, which within the general linear model combines ANOVA and regression, is the most appropriate and least biased approach (see Rutherford, 2011 for a detailed explanation of the uses of ANCOVA). For these reasons, an ANCOVA style regression was run in order to gain a more complete understanding of the development of mental rotation skills throughout pre-k.

DISCUSSION

The current study aimed to examine the development of sex differences in the mental rotation skills of Hispanic children throughout pre-k by exploring (1) whether sex differences exist in Hispanic pre-kindergartners' mental rotation skills at time 1 and time 2; and (2) whether there are sex differences in the *changes* (i.e., *gains*) Hispanic children make on mental rotation skills throughout pre-k. In short, our results suggest that simply looking at sex differences at only one time would suggest there are no sex differences in mental rotation skills at this age; however, by examining sex differences over time with robust analysis utilizing potential covariates, boys were found to make greater gains in mental rotation than girls. The importance of these findings is threefold as they suggest (1) pre-kindergarten is a time of significant change and emergence of sex differences in mental rotation skills; (2) different methodologies such as including multiple timepoints and examining change is critical for understanding when sex differences in mental rotation skills develop; and (3) early sex differences in mental rotation are generalizable to a SES-diverse population of Hispanic children living in the United States.

Sex Differences in Mental Rotation Skills

Our findings that boys make greater gains in mental rotation than girls throughout prekindergarten provide interesting insight into the study of sex differences in childhood. While sex differences in the mental rotation skills of adults are well established, studies examining sex differences in children's spatial skills show inconsistent results (Frick et al., 2014). Our results help explain these discrepancies in the literature, as they show pre-k boys make greater improvements than girls on a task requiring mental rotation skills, but boys and girls do not significantly differ in performance when examined at one single timepoint. Furthermore, while there is a similar amount of variability in boys' and girls' scores at time 1, there is greater variability in boys' scores compared to girls' scores at time 2. These findings suggest that pre-kindergarten is a time when sex differences in mental rotation skills are emerging, though more studies will be needed to determine when exactly these differences emerge.

One possible explanation for boys improving significantly more on the CMTT than girls is that early education may be providing more opportunities for boys, compared to girls, to advance their mental rotation skills. Previous research shows boys are exposed to more activities (e.g., Legos, blocks, construction toys) that promote spatial learning than girls (Newcombe et al., 1983; Baenninger and Newcombe, 1995; Nazareth et al., 2013). While speculative, since we did not gather data on children's exposure to spatial activities in the classroom, it is possible boys may have been exposed to new spatial experiences that improved their mental rotation skills in pre-k, allowing them to make larger gains on the CMTT. Another possible explanation for boys improving more than girls on the CMTT is that boys hear more spatial language than girls from their parents in the home setting (Pruden and Levine, 2017). Similarly, it could be that educators, like parents, use new or a greater quantity and quality of spatial language with boys in the pre-k classroom, contributing to the development of boys' spatial skills. While speculative, these various factors highlight the need for more research on spatial development in the early education setting, the impact of early education on spatial development, and whether boys and girls are exposed to the same kind of spatial experiences (e.g., activities and language) in the early education classroom (Costales et al., 2015). These mechanisms alone, or more likely via complex interactions, may provide a powerful means to promote spatial thinking in both sexes.

Examining Sex Differences at Multiple Timepoints

The current findings suggest the importance of examining change in sex differences in mental rotation through time to better understand these differences and when and how they develop. Had we explored sex differences at only one of the two timepoints, we would have concluded that there were no sex differences in Hispanic pre-kindergartners' mental rotation skills. However, by examining change in boys' and girls' mental rotation skills throughout the school year, we were able to observe that mental rotation ability in pre-k is actually different for boys and girls, with boys making greater improvement on a mental rotation task than girls.

Given our finding that sex differences in mental rotation are growing stronger or emerging at this age, it is possible that studies looking for sex differences at only one timepoint may not always find significant results, even when differences are present. Our results highlight the importance of utilizing different methodologies examining change over time to uncover the earliest indicators of sex differences in spatial thinking and help explain the inconsistencies in the current research on early sex differences in spatial thinking. Given that spatial skills are malleable and can be improved (Uttal et al., 2013), early detection of sex differences in spatial thinking by exploring change is critical and may help us identify when in development spatial experiences and training for both girls and boys should occur.

Generalizing Sex Differences in Mental Rotation to a Hispanic Population

Studies exploring sex differences in diverse populations show inconsistent results (Berry, 1966; Jahoda, 1980; Mann et al., 1990; Lynn, 1992; Feingold, 1994; Lemke et al., 2004; Levine et al., 2005; Silverman et al., 2007; Casey et al., 2008; Liu and Lynn, 2011). To date, few studies have examined sex differences with diverse populations including Hispanic participants (though see Casey et al., 2008; Nazareth et al., 2013). Our findings suggest sex differences in mental rotation skills are generalizable to a Hispanic population of children of varying SES living in the United States. Given the link between spatial thinking and future entry into STEM fields (Humphreys et al., 1993; Shea et al., 2001; Wai et al., 2009) and the exponentially increasing Hispanic population in the United States, our findings may have important implications for future work aimed at understanding the underrepresentation of minorities and women in STEM fields.

Limitations

It is important to note some limitations to this study. First, the current study only measured mental rotation ability at two timepoints, limiting our ability to test non-linear relations and utilize more powerful statistical tools like growth curve modeling. Second, given the small number of children assessed at some of the classrooms and schools, it was not possible to determine whether there were any classroom or school clustering effects. Third, given the observational nature of this study, we are unable to make causal inferences regarding the mechanisms that lead to changes in mental rotation performance and the sex differences seen in changes on the mental rotation task. Fourth, given the nature of this study it was not possible to control for the influence of many factors which have been shown to influence sex differences in mental rotation skills (e.g., engagement with spatial toys and activities, gender stereotypes, spatial anxiety, and spatial language. Finally, our study is the first to examine spatial thinking skills in a bilingual population. While we believe this is an important and understudied area of research, given that our entire sample was English-Spanish bilinguals, it was not possible for us to examine the specific role of bilingualism versus monolingualism on the development of spatial thinking.

CONCLUSION

The current study suggests there are sex differences in the gains made throughout pre-k on mental rotation, with boys making significantly more gains than girls. The current findings point to the need to explore change over time to attain a greater understanding of sex differences in mental rotation ability. Future research should continue to explore the influence of early schooling on the development of spatial skills in diverse populations, with a particular focus on the mechanisms resulting in changes in spatial thinking and the factors leading to sex differences in these changes (e.g., spatial activities and spatial language). Given the link between spatial thinking and future entry into STEM fields (Humphreys et al., 1993; Shea et al., 2001; Wai et al., 2009), a better understanding of the influence early education, among other potential factors, on spatial development in boys and girls from diverse backgrounds is needed. Identifying mechanisms that promote growth in spatial thinking is critical to increasing the number of minorities and women entering STEM fields.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Florida International University's Institutional Review Board. Participants' primary caregivers gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Institutional Review Board at Florida International University.

AUTHOR CONTRIBUTIONS

CA, RO, and SP contributed to conception and design of the study. CA and RO collected the data and performed the statistical analysis. CA wrote the first draft of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

This research was supported by a generous grant from The Ware Foundation to the Project on Language and Spatial Development. This research is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1038321. Any opinion, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

ACKNOWLEDGMENTS

We thank the research assistants who worked on this study and the preschools, educators, and children for their participation.

REFERENCES

Abad, C. (2018). *The Development of Spatial Thinking*. Ph.D. thesis, Florida International University, Miami, FL.

Baenninger, M., and Newcombe, N. (1989). The role of experience in spatial test performance: a meta-analysis. *Sex Roles* 20, 327–344. doi: 10.1007/BF00287729

Baenninger, M., and Newcombe, N. (1995). Environmental input to the development of sex-related differences in spatial and mathematical ability. *Learn. Individ. Differ.* 7, 363–379. doi: 10.1016/1041-6080(95)90007-1

Berry, J. W. (1966). Temne and Eskimo perceptual skills. *Int. J. Psychol.* 1, 207–229. doi: 10.1080/00207596608247156

Caldwell, E. C., and Hall, V. C. (1970). Concept learning in discrimination tasks. *Dev. Psychol.* 2, 41–48. doi: 10.1037/h0028606

Casey, B. M., Andrews, N., Schindler, H., Kersh, J. E., Samper, A., and Copley, J. (2008). The development of spatial skills through interventions involving block building activities. *Cogn. Instr.* 26, 269–309. doi: 10.1080/07370000802177177

Cheng, Y. L., and Mix, K. S. (2014). Spatial training improves children's mathematics ability. *J. Cogn. Dev.* 15, 2–11. doi: 10.1080/15248372.2012.725186

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S. (2003). *Applied Multiple Correlation/Regression Analysis for the Behavioral Sciences*. Milton Park: Taylor & Francis.

Colby, S. L., and Ortman, J. M. (2015). *Projections of the Size and Composition of the US Population: 2014 to 2060*. Available at: https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf

Costales, A., Abad, C., Odean, R., and Pruden, S. M. (2015). "Spatial activities and manipulatives for early childhood classrooms," in *The Sage Encyclopedia of Classroom Management*, ed. G. Scarlett (Thousand Oaks, CA: Sage Publication).

Dunn, L. M., and Dunn, L. M. (1997). *Peabody Picture Vocabulary Test –(PPVT-III)*, 3rd Edn. Circle Pines, MN: American Guidance Service.

Dunn, L. M., Padilla, E., Lugo, D., and Dunn, L. M. (1986). *Test de Vocabulario en Imágenes Peabody—Adaptación Hispanoamericana [Peabody Picture Vocabulary Test—Latin American Adaptation]*. Circle Pines, MN: American Guidance Service.

Eccles, J. S., and Jacobs, J. E. (1986). Social forces shape math attitudes and performance. *Signs* 11, 367–380. doi: 10.1086/494229

Ehrlich, S. B., Levine, S. C., and Goldin-Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Dev. Psychol.* 42, 1259–1268. doi: 10.1037/0012-1649.42.6.1259

Estes, D. (1998). Young children's awareness of their mental activity: the case of mental rotation. *Child Dev.* 69, 1345–1360. doi: 10.1080/15248372.2012.725186

Feingold, A. (1994). Gender differences in variability in intellectual abilities: a cross-cultural perspective. *Sex Roles* 30, 81–92. doi: 10.1007/BF01420741

Frick, A., Daum, M. M., Wilson, M., and Wilkening, F. (2009). Effects of action on children's and adults' mental imagery. *J. Exp. Child Psychol.* 104, 34–51. doi: 10.1016/j.jecp.2009.01.003

Frick, A., Ferrara, K., and Newcombe, N. S. (2013). Using a touch screen paradigm to assess the development of mental rotation between 3½ and 5½ years of age. *Cogn. Process.* 14, 117–127. doi: 10.1007/s10339-012-0534-0

Frick, A., Möhring, W., and Newcombe, N. S. (2014). Development of mental transformation abilities. *Trends Cogn. Sci.* 18, 536–542. doi: 10.1016/j.tics.2014.05.011

Humphreys, L. G., Lubinski, D., and Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *J. Appl. Psychol.* 78, 250–261. doi: 10.1037/0021-9010.78.2.250

Huttenlocher, J., Levine, S., and Vevea, J. (1998). Environmental input and cognitive growth: a study using time-period comparisons. *Child Dev.* 69, 1012–1029. doi: 10.2307/1132360

Jahoda, G. (1979). On the nature of difficulties in spatial-perceptual tasks: ethnic and sex differences. *Br. J. Psychol.* 70, 351–363. doi: 10.1111/j.2044-8295.1979.tb01705.x

Jahoda, G. (1980). Sex and ethnic differences on a spatial–perceptual task: some hypotheses tested. *Br. J. Psychol.* 71, 425–431. doi: 10.1111/j.2044-8295.1980.tb01757.x

Jansen, P., and Heil, M. (2010). The relation between motor development and mental rotation ability in 5- to 6-year-old children. *Eur. J. Dev. Sci.* 4, 67–75. doi: 10.3233/DEV-2010-4105

Joh, A. S. (2016). Training effects and sex difference in preschoolers' spatial reasoning ability. *Dev. Psychobiol.* 58, 896–908. doi: 10.1002/dev.21445

Johnson, E. S., and Meade, A. C. (1987). Developmental patterns of spatial ability: an early sex difference. *Child Dev.* 58, 725–740. doi: 10.2307/1130210

Kaess, D. W. (1971). Measures of form constancy: developmental trends. *Dev. Psychol.* 4:296. doi: 10.1037/h0030443

Kaplan, B. J., and Weisberg, F. B. (1987). Sex differences and practice effects on two visual-spatial tasks. *Percept. Mot. Skills* 64, 139–142. doi: 10.2466/pms.1987.64.1.139

Kosslyn, S. M., Margolis, J. A., Barrett, A. M., Goldknopf, E. J., and Daly, P. F. (1990). Age differences in imagery abilities. *Child Dev.* 61, 995–1010. doi: 10.2307/1130871

Krüger, M., and Krist, H. (2009). Imagery and motor processes—When are they connected? the mental rotation of body parts in development. *J. Cogn. Dev.* 10, 239–261. doi: 10.1080/15248370903389341

Lachance, J. A., and Mazzocco, M. M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learn. Individ. Differ.* 16, 195–216. doi: 10.1016/j.lindif.2005.12.001

Lauer, J. E., Udelson, H. B., Jeon, S. O., and Lourenco, S. F. (2015). An early sex difference in the relation between mental rotation and object preference. *Front. Psychol.* 6:558. doi: 10.3389/fpsyg.2015.00558

Lawton, C. A. (1994). Gender differences in way-finding strategies: relationship to spatial ability and spatial anxiety. *Sex Roles* 30, 765–779. doi: 10.1007/BF01544230

Lehmann, J., Quaiser-Pohl, C., and Jansen, P. (2014). Correlation of motor skill, mental rotation, and working memory in 3- to 6-year-old children. *Eur. J. Dev. Psychol.* 11, 560–573. doi: 10.1080/17405629.2014.888995

Lemke, M., Sen, A., Pahlke, E., Partelow, L., Miller, D., Williams, T., et al. (2004). *International Outcomes of Learning in Mathematics Literacy and Problem Solving: PISA 2003 Results from the U.S. Perspective (NCES Report No.* 2005-003). Washington, DC: U.S. Department of Education.

Levine, S. C., Foley, A., Lourenco, S., Ehrlich, S., and Ratliff, K. (2016). Sex differences in spatial cognition: advancing the conversation. *Wiley Interdiscip. Rev. Cogn. Sci.* 7, 127–155. doi: 10.1002/wcz.1380

Levine, S. C., Huttenlocher, J., Taylor, A., and Langrock, A. (1999). Early sex differences in spatial skill. *Dev. Psychol.* 35, 940–949. doi: 10.1037/0012-1649.35.4.940

Levine, S. C., Ratliff, K. R., Huttenlocher, J., and Cannon, J. (2012). Early puzzle play: a predictor of preschoolers' spatial transformation skill. *Dev. Psychol.* 48, 530–542. doi: 10.1037/a0025913

Levine, S. C., Vasilyeva, M., Lourenco, S. F., Newcombe, N. S., and Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychol. Sci.* 16, 841–845. doi: 10.1111/j.1467-9280.2005.01623.x

Linn, M. C., and Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Dev.* 56, 1479–1498. doi: 10.2307/1130467

Lippa, R. A., Collaer, M. L., and Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Arch. Sex. Behav.* 39, 990–997. doi: 10.1007/s10508-008-9460-8

Liu, J., and Lynn, R. (2011). Factor structure and sex differences on the Wechsler Preschool and Primary Scale of Intelligence in China, Japan and United States. *Pers. Individ. Differ.* 50, 1222–1226. doi: 10.1016/j.paid.2011.02.013

Lynn, R. (1992). Sex differences on the Differential Aptitude Test in British and American adolescents. *Educ. Psychol.* 12, 101–102. doi: 10.1080/0144341920120201

Maccoby, E. E., and Jacklin, C. N. (1974). *The Psychology of Sex Differences*, Vol. 1. Stanford, CA: Stanford University Press.

Manger, T., and Eikeland, O. (1998). The effects of spatial visualization and students' sex on mathematical achievement. *Br. J. Psychol.* 89, 17–25. doi: 10.1111/j.2044-8295.1998.tb02670.x

Mann, V. A., Sasanuma, S., Sakuma, N., and Masaki, S. (1990). Sex differences in cognitive abilities: a cross-cultural perspective. *Neuropsychologia* 28, 1063–1077. doi: 10.1016/0028-3932(90)90141-A

Moore, D. S., and Johnson, S. P. (2008). Mental rotation in human infants: a sex difference. *Psychol. Sci.* 19, 1063–1066. doi: 10.1111/j.1467-9280.2008.02200.x

National Science Foundation and National Center for Science and Engineering Statistics (2013). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2013*. Arlington, VA: National Science Foundation.

Nazareth, A., Herrera, A., and Pruden, S. M. (2013). Explaining sex differences in mental rotation: role of spatial activity experience. *Cogn. Process.* 14, 201–204. doi: 10.1007/s10339-013-0542-8

Newcombe, N., Bandura, M. M., and Taylor, D. G. (1983). Sex differences in spatial ability and spatial activities. *Sex Roles* 9, 377–386. doi: 10.1007/BF00289672

Newsom, J. T. (2012). "Basic longitudinal analysis approaches for continuous and categorical variables," in *Longitudinal data Analysis: A Practical Guide for Researchers in Aging, Health, and Social Sciences*, eds J. T. Newsom, R. N. Jones, and S. M. Hofer (New York, NY: Routledge), 143–179.

Parsons, J. E., Adler, T. F., and Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: parental influences. *Child Dev.* 53, 310–321. doi: 10.2307/1128973

Platt, J. E., and Cohen, S. (1981). Mental rotation task performance as a function of age and training. *J. Psychol.* 108, 173–178. doi: 10.1080/00223980.1981.9915260

Pruden, S. M., and Levine, S. C. (2017). Parents' spatial language mediates a sex difference in preschoolers' spatial-language use. *Psychol. Sci.* 28, 1583–1596. doi: 10.1177/0956797617711968

Pruden, S. M., Levine, S. C., and Huttenlocher, J. (2011). Children's spatial thinking: Does talk about the spatial world matter? *Dev. Sci.* 14, 1417–1430. doi: 10.1111/j.1467-7687.2011.01088.x

Quinn, P. C., and Liben, L. S. (2008). A sex difference in mental rotation in young infants. *Psychol. Sci.* 19, 1067–1070. doi: 10.1111/j.1467-9280.2008.02201.x

Quinn, P. C., and Liben, L. S. (2014). A sex difference in mental rotation in infants: Convergent evidence. *Infancy* 19, 103–116. doi: 10.1111/infa.12033

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychol. Bull.* 86, 638–641. doi: 10.1037/0033-2909.86.3.638

Rutherford, A. (2011). *ANOVA and ANCOVA: a GLM Approach*. Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9781118491683

Schafer, J. L., and Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychol. Methods* 7, 147–177. doi: 10.1037/1082-989X.7.2.147

Shea, D. L., Lubinski, D., and Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: a 20-year longitudinal study. *J. Educ. Psychol.* 93, 604–614. doi: 10.1037/0022-0663.93.3.604

Silverman, I., Choi, J., and Peters, M. (2007). The hunter-gatherer theory of sex differences in spatial abilities: data from 40 countries. *Arch. Sex. Behav.* 36, 261–268. doi: 10.1007/s10508-006-9168-6

Sinton, D., Bendarz, S., Gershmehl, P., Kolvoord, R. A., and Uttal, D. (2013). *The People's Guide to Spatial Thinking*. Washington, DC: National Council for Geographic Education.

Tabachnick, B. G., and Fidell, L. S. (2013). *Using Multivariate Statistics*. Boston, MA: Pearson.

Terlecki, M. S., Newcombe, N. S., and Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: gender differences in growth patterns. *Appl. Cogn. Psychol.* 22, 996–1013. doi: 10.1002/acp.1420

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., et al. (2013). The malleability of spatial skills: a meta-analysis of training studies. *Psychol. Bull.* 139, 352–402. doi: 10.1037/a0028446

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., and Newcombe, N. S. (2017). Links between spatial and mathematical skills across the preschool years: IV. Results—Links between spatial assembly, later spatial skills, and concurrent and later mathematical skills. *Monogr. Soc. Res. Child Dev.* 82, 71–80. doi: 10.1111/mono.12283

Voyer, D., Voyer, S., and Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychol. Bull.* 117, 250–270. doi: 10.1037/0033-2909.117.2.250

Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J. Educ. Psychol.* 101, 817–835. doi: 10.1037/a0016127

Wright, R., Thompson, W. L., Ganis, G., Newcombe, N. S., and Kosslyn, S. M. (2008). Training generalized spatial skills. *Psychon. Bull. Rev.* 15, 763–771. doi: 10.3758/PBR.15.4.763

Yuan, K. H., and Hayashi, K. (2010). Fitting data to model: structural equation modeling diagnosis using two scatter plots. *Psychol. Methods* 15, 335–351. doi: 10.1037/a0020140

Conflict of Interest Statement: The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge, or beliefs) in the subject matter or materials discussed in this manuscript.

frontiers in Psychology

ORIGINAL RESEARCH
published: 22 February 2019
doi: 10.3389/fpsyg.2019.00380

Check for updates

Changes in United States Latino/a High School Students' Science Motivational Beliefs: Within Group Differences Across Science Subjects, Gender, Immigrant Status, and Perceived Support

Ta-yang Hsieh*, Yangyang Liu and Sandra D. Simpkins

School of Education, University of California, Irvine, Irvine, CA, United States

Science motivational beliefs are crucial for STEM (science, technology, engineering, and math) performance and persistence, but these beliefs typically decline during high school. We expanded the literature on adolescents' science motivational beliefs by examining: (1) changes in motivational beliefs in three specific science subjects, (2) how gender, immigrant generation status, and perceived support from key social agents predicted differences in adolescents' science motivational beliefs, and (3) these processes among Latino/as in the United States, whose underrepresentation in STEM is understudied. We used hierarchical linear modeling to estimate the changes in 104 (40% female) Latino/a high school students' physics, chemistry, and biology motivational beliefs from 9th to 11th grade. Subject-specific ability self-concept, interest, and utility were regressed on gender, immigrant generation status, and perceived science support while controlling for family income, parent education, and adolescents' school. Adolescents' utility declined from 9th to 11th grade whereas their interest remained stable for all three science subjects. Adolescents' ability self-concept increased for biology, decreased for physics, but remained stable for chemistry. Gender differences in adolescents' motivational beliefs at 9th grade only emerged for physics utility as well as physics and chemistry interest; yet, there were no gender differences in how adolescents' science motivational beliefs changed over time. Contrary to expectations, immigrant generation status was not significantly associated with adolescents' science motivational beliefs at 9th grade or in terms of how they changed over time. Adolescents who perceived higher science support generally had higher motivational beliefs in 9th grade, but did not differ on their rate of change. Our findings highlight the need to examine specific science subjects, and that typical gender differences in adolescents' motivational beliefs discussed in the literature may not generalize to all racial and ethnic groups.

Keywords: science, motivational beliefs, Latino, gender, immigrant generation status, science support, interest, ability self-concept

OPEN ACCESS

Edited by:
Rebecca Lazarides, Universität Potsdam, Germany

Reviewed by:
Ursula Moffitt, Universität Potsdam, Germany
Sauro Civitillo, University of Duisburg-Essen, Germany

*Correspondence:
Ta-yang Hsieh
tayangh@uci.edu

Specialty section:
This article was submitted to Developmental Psychology, a section of the journal Frontiers in Psychology

Received: _14 December 2018_
Accepted: _07 February 2019_
Published: _22 February 2019_

Citation:
Hsieh T-y, Liu Y and Simpkins SD (2019) Changes in United States Latino/a High School Students' Science Motivational Beliefs: Within Group Differences Across Science Subjects, Gender, Immigrant Status, and Perceived Support. Front. Psychol. 10:380. doi: 10.3389/fpsyg.2019.00380

INTRODUCTION

Gender disparities in science have gained both research and policy interest in the United States. The most recent statistics show that although females have gained fair representation in the biological sciences, they are still significantly underrepresented in the physical sciences (National Science Foundation [NSF], 2017). The female underrepresentation in the physical sciences is even more pronounced for certain ethnic minorities such as Latino/as, who are the fastest growing underrepresented minority population in the United States. Latino/as account for 18% of the United States population and are projected to account for more than half of the school-aged Americans by 2050. Latino/as, however, only account for 5% of all computer/mathematical scientists and physical scientists in the United States (Fry and Gonzales, 2008; National Science Foundation [NSF], 2017). Most research to date focuses on ethnic group comparisons, which highlight the underrepresentation of Latino/as. It is also important to understand the differences *among* Latino/as as some Latino/as are succeeding and pursuing STEM (science, technology, engineering, and math) pathways. Studies focusing on a specific ethnic group can help determine the various developmental pathways for that group, interindividual differences in those processes, and what supports positive development. In other words, studies examining the heterogeneity *within* the Latino/a population that identify factors predicting favorable outcomes are needed. The aim of the current study, hence, was not to generalize findings to the entire population, but instead to provide an in-depth analysis on a specific ethnic group that is often underrepresented in the literature and in STEM fields.

The focal outcomes of the current study were Latino/a adolescents' science motivational beliefs from 9th to 11th grade, which according to the expectancy-value theory are critical predictors of science achievement-related outcomes. Considering that high school is a time when United States students face life-altering decisions, such as attending college and selecting a major, motivational beliefs could significantly influence their pursuit into science majors or careers during this turning point (Sadler et al., 2012). Additionally, we examined both demographic (i.e., gender and immigrant generational status) and contextual factors (i.e., perceived science support from significant social agents) as predictors of those science motivational beliefs.

Expectancy-Value Theory

According to the expectancy-value theory (Wigfield and Eccles, 2000), individuals will be more likely to choose, persist, and excel in a task or domain if they hold high motivational beliefs about that domain. With regard to science, people who have strong science motivational beliefs are more likely to pick science classes and college majors (Buday et al., 2012). Earning high grades and demonstrating achievement or mastery in science is not enough to support individuals' continued pursuit of science. Individuals also need to value and believe they are good at science to select into and persist in science-related fields (Jacobs et al., 2005).

Three of the key motivational beliefs as described by the expectancy-value theory are interest, utility, and ability self-concept. Interest, in this case, is how enjoyable individuals find science to be. Utility pertains to how useful individuals think science is. Ability self-concept is how good individuals think they are in science. It is theorized that the more value people see in science and the more people expect themselves to excel, the more likely they would choose to engage and persist in science-related fields. Prior studies have indeed supported the positive associations between science motivational beliefs and favorable outcomes, such as higher achievement, engagement, and aspirations in science (Lau and Roeser, 2002; Singh et al., 2002; Osborne et al., 2003; Maltese and Tai, 2011; Nagengast and Marsh, 2012).

Individuals' motivational beliefs are expected to change over time as they develop and gain more experience (Wigfield et al., 2015). Several studies consistently suggest that students' academic intrinsic motivation, ability self-concept, and perceived value generally decline with age (e.g., Lepper and Henderlong, 2000; Gottfried et al., 2001; Fredricks and Eccles, 2002; Jacobs et al., 2002; Lepper et al., 2005; Wigfield and Wagner, 2005; Lazarides and Raufelder, 2017; Lazarides and Watt, 2017). However, changes in youth's science motivational beliefs have received less attention and the existing findings are inconsistent. When science motivation was examined as science overall, white American adolescents' science motivational beliefs declined over time (Gottfried et al., 2009). But, when science motivational beliefs were examined specifically within the physical sciences, most adolescents reported either stability or decrease in their ability self-concept and value (Wang et al., 2017). Although the Wang and colleague's study (2017) is highly relevant to the current study, their participants were predominately white students. The science gaps and predictors vary across racial/ethnic groups (National Science Foundation [NSF], 2017) and the patterns found for white adolescents may not generalize to underrepresented groups (Hsieh and Simpkins, 2018). Thus, the first aim of the current study is to describe how Latino/a adolescents' science motivational beliefs changed in high school.

Another limitation is that researchers have typically examined youth's motivational beliefs in science overall without differentiating between the specific science subjects. Not only does the expectancy-value theory argue for the importance of domain-specificity, middle and high school students' motivational beliefs differ based on the specific science subject. For example, middle school students, on average, were more interested in and placed higher value on biology compared with chemistry and physics (Bennett and Hogarth, 2009). Relatedly, high school girls reported stronger ability self-concepts in earth science than boys, whereas no gender differences were found in the physical and life sciences (Britner, 2008). Finally, both white and Latino/a 9th grade students held different ability self-concepts and values across biology, chemistry, and physics (Simpkins et al., 2015b). Indicated by these findings, collapsing biology, chemistry, and physics into 'science' could mask the differences youth see in these subjects and could instead provide an average that is not representative of their true beliefs about any particular science subject. Moreover, females are overrepresented in the biological sciences and underrepresented in the physical sciences, making collapsing multiple science subjects particularly problematic. To address these gaps in the literature, the current study examined the three major science subjects in United States high school curricula, namely biology, chemistry, and physics. In addition, we also took the average of the three subjects for the purpose of showing how overall 'science' might fail to represent the nuances that each subject offers.

The expectancy-value theory also posits that motivational beliefs are shaped by individual characteristics (e.g., gender), the cultural milieu and family demographics (e.g., immigrant generational status), and socialization by others (e.g., perceived support). In the following paragraphs, we reviewed the theory and literature on gender, immigrant generational status, and perceived support, respectively, as potential determinants of motivational beliefs.

Gender Differences in the Trajectories of Students' Motivational Beliefs

One of the most salient demographic factors that might influence students' motivational beliefs in science is gender. According to the expectancy-value theory, gender stereotypes and expectations shape socializers' behaviors and individuals' personal and social identities, which influence their expectancies and the values they associate with specific domains. Indeed, female students tend to rate themselves lower on math and science motivational beliefs than male students (Stake and Nickens, 2005; Simpkins et al., 2006, 2015a,b; Kurtz-Costes et al., 2008; Sax et al., 2015; Lazarides and Watt, 2017). However, there is evidence that gender differences in students' motivational beliefs vary across ethnic groups (Else-Quest et al., 2013; Hsieh and Simpkins, 2018). It is unclear if the typical gender differences found among white majority students will emerge between Latinas and Latinos. On one hand, traditional gender roles are a core Latino/a cultural value (Knight et al., 2010) and may amplify the stereotype favoring males in science and may more negatively influence females' science motivational beliefs. On the other hand, Latinas often outperform Latinos in school, which suggest females' motivational beliefs might be higher than males (Plunkett and Bámaca-Gómez, 2003). Only a couple studies to our knowledge addressed this question directly and showed that Latinos reported slightly higher science (general), biology, chemistry, and physics ability self-concepts than Latinas, but there was no gender difference in how much the adolescents valued science, biology, or chemistry (Else-Quest et al., 2013; Simpkins et al., 2015b).

Previous studies have predominately focused on mean-level differences between males' and females' science motivational beliefs at one time point; gender differences in the *changes* in students' science motivational beliefs have largely been unexplored. The observation that females reported lower science motivational beliefs than males at end of high school could result from lower motivational beliefs for females at start of high school (9th grade), faster declines among females, or a combination of both. Longitudinal data would allow us to empirically address these possibilities. Findings regarding changes in United States students' math motivational beliefs suggested a gendered rate of change where females' math motivational beliefs declined slower than that of males and actually helped close the gender gap from kindergarten through high school (Fredricks and Eccles, 2002; Jacobs et al., 2002). The current study examined not only how gender predicted Latino/as' science motivational beliefs at the beginning of high school, but also how gender predicted the rate of change of these beliefs over time. Given that females are underrepresented in physics and chemistry but overrepresented in biology (National Science Foundation [NSF], 2017), we expected Latinas' physics and chemistry motivational beliefs would be lower at 9th grade and decline at a faster rate than Latinos' beliefs. In contrast, we expected Latinas' biology motivational beliefs to be higher at 9th grade and decline at a slower rate than Latinos' beliefs.

Immigrant Generational Status Differences in the Trajectories of Students' Motivational Beliefs

In the expectancy-value theory, individuals' motivational beliefs are shaped by the social and cultural contexts. An important indicator of Latino/a students' social and cultural context and predictor of their academic achievement is immigrant generational status. Findings on the immigrant paradox in education suggests that third generation Latino/as (i.e., both parents and youth were born in the United States) underperform their first and second generation counterparts (i.e., parents were born outside of the United States) in school after controlling for family socioeconomic status (Palacios et al., 2008; García Coll et al., 2012; Greenman, 2013; Aretakis et al., 2015; Feliciano and Lanuza, 2017). This phenomenon is coined as paradoxical because despite potential language barriers and other burdens to adapt, Latino/a students who immigrated or whose parents immigrated outperform or show 'super-achievement' compared with youth who and whose parents were born in the United States (Harris et al., 2008). Possible mechanisms underlying the immigrant paradox include differences across generations in cultural identity and resources in the proximal community such that earlier generations of immigrants have more support from their community and identify more closely with their ethnic culture, both of which function as protective factors (Perreira et al., 2010; Aretakis et al., 2015). Although prior studies suggest that third generation Latino/as have lower academic achievement and attainment than their first and second generation counterparts, it has not been empirically shown whether students' motivational beliefs follow the same trend. Taken together, immigrant generational status is a relevant predictor to examine, and it is hypothesized that third generation students would have lower science motivational beliefs in 9th grade and evidence faster declines over time than their first and second generation counterparts.

Perceived Support as a Protective Factor

Unlike studies on ethnic minorities that tried to identify risk factors, we took a strength-based perspective and examined whether adolescents' perceived level of science support from parents, teachers, friends, and siblings/cousins in 9th grade could help buffer them against the typical declines in students' motivational beliefs. According to the expectancy-value theory, perceived support from key social agents is positively associated with adolescents' motivational beliefs (Eccles et al., 1983). For high school students, parents, teachers, friends, and siblings jointly create the proximal social contexts for motivation development (Plunkett and Bámaca-Gómez, 2003; Alfaro and Umaña-Taylor, 2015; Lazarides et al., 2017). Perceived support from some or all of these social agents positively predicts adolescents' concurrent motivational beliefs (Bouchey and Harter, 2005; Alfaro et al., 2006; Plunkett et al., 2008). Extending the previous literature, the current study examined how perceived science support from parents, teachers, friends, and siblings/cousins at the beginning of high school was associated with both concurrent levels of science motivational beliefs and *changes* in science motivational beliefs over the next 2 years. A relevant study by Alfaro and Umaña-Taylor (2015) showed that adolescents' perceived support at the beginning of high school predicted their overall academic motivation in the same year, but not change over time. The current study is different from Alfaro and Umaña-Taylor (2015) in two ways. First, we examined science-specific support and motivational beliefs instead of academic support in general. Secondly, we included perceived support from friends, which Alfaro and Umaña-Taylor also argued to be salient particularly during adolescence but did not include in their study. We considered perceived support from parents, teacher, friends, and sibling to capture the fact the adolescents are embedded in and interact with multiple social agents.

It is important to examine support jointly as well as the independent roles of parents, teachers, siblings, and friends in predicting the trajectories of students' motivational beliefs. Prior studies that examined support from specific social agents among predominantly white high school students showed positive associations between the students' science motivational beliefs and their perceived support from parents, teachers, and friends (Adya and Kaiser, 2005; Leaper et al., 2012; Lazarides and Ittel, 2013; Rice et al., 2013). Additionally, most studies that focused on Latino/a adolescents also replicated such positive associations between adolescents' overall academic or science motivational beliefs and their perceived support from their parents (Simpkins et al., 2015b), teachers, (Lewis et al., 2012; Riconscente, 2014), and siblings (Alfaro and Umaña-Taylor, 2010). Taken together, we examined adolescents' perceived science support from parents, teacher, sibling, and friends not only jointly as a composite, but also (in follow up analyses) as individual predictors.

Current Study

Based on both theory and empirical research, the current study examined changes in Latino/a high school students' science motivational beliefs from 9th to 11th grade. We also took into account individual and contextual factors including gender, immigrant generation status, and perceived science support from key socializers. Specifically, three research questions were tested.

Research Question 1 asked how students' science motivational beliefs changed from 9th to 11th grade. We first tested overall science across the three subjects and then examined biology, chemistry, and physics separately. Aligning with results from prior studies (e.g., Gottfried et al., 2009; Wang et al., 2017), we hypothesized that students' motivational beliefs would decline over time.

Research Question 2 examined whether there were demographic differences in students' motivational beliefs at 9th grade and in the changes over time. First, we examined gender differences. Given that physics and chemistry are stereotypically masculine subjects (Su et al., 2009; Sax et al., 2015), we expected females to have lower motivational beliefs in these two subjects at 9th grade and to evidence faster declines than males. In contrast, given that the biological sciences are now overrepresented by females (National Science Foundation [NSF], 2017), we expected females to have higher biology motivational beliefs at 9th grade and to evidence slower declines than males. Next, we examined whether there were differences based on students' immigrant generation status. Specifically, we tested whether being third generation immigrant versus being first or second generation immigrant predicted differences in adolescents' science motivational beliefs at 9th grade and how they changed from 9th to 11th grade. In accordance with the Latino/a immigrant paradox in education (Feliciano and Lanuza, 2017), we expected first and second generation immigrants to report higher 9th grade motivational beliefs and evidence slower declines compared with their third generation counterparts.

Lastly, Research Question 3 asked whether adolescents' perceived science support in 9th grade predicted their concurrent science motivational beliefs and changes in those beliefs over time. With a strength-based perspective and supported by prior studies (e.g., Bouchey and Harter, 2005; Alfaro and Umaña-Taylor, 2015), perceived science support was hypothesized to serve as a buffer such that higher perceived science support would predict higher beliefs at 9th grade and slower declines over time. Research Questions 2 and 3 were tested in the same models with family income, parent education level, the adolescent's school as covariates.

MATERIALS AND METHODS

Sample

Participants of the current study were 104 Latino/a adolescents (40% female) from three high schools in Arizona, United States. When the data were collected, Arizona was one of the six states in the United States with a Latino/a population of two million or more (U. S. Census Bureau, Population Division, 2014). Just a few years prior to collecting these data, a new law (SB-1070) was debated and passed which allowed law enforcement to inquire about people's immigration status. Scholars have documented the potential negative effects of the climate for the Latino/a ethnic community in Arizona, such as ethnic-based microaggressions impacting the daily experiences of adolescents and their families (e.g., Lin et al., 2016).

The participants were first recruited when they were in 9th grade (M_{age} = 14.50, SD = 0.52) during the 2012–2013 school year and were recontacted 1 and 2 years later (i.e., 10th and 11th grade). As part of a larger study, participants were recruited with one of their parents and an older sibling or cousin. Written consents/assents were obtained from all participants, and the study was approved by the Arizona State University institutional review board ethics committee. Data utilized in this study include adolescent-reported data at each time point and parent-reported data in 9th grade. Data collection in 9th and 10th grade mostly happened in participants' homes with a few exceptions at the university or local library; data collection in 11th grade took place in adolescents' schools. Adolescent participants and their parents separately completed questionnaires with trained, bilingual Latino/a interviewers. All questionnaires were available in both Spanish and English. All but one student completed their surveys in English; 43% of parents completed the survey in English. Surveys were translated by bilingual research assistants and checked for accuracy using a forward-translation and panel method approach (Knight et al., 2009). Adolescents and parents each received $50 as compensation each year.

In terms of academics, 9% of the adolescents took honors science classes in 9th grade. The majority of the adolescents were either first or second generation (61%) and Mexican-origin[1] (82%). Around 62% of the adolescents came from 2-parent, married families and 4.8% of parents had a science-related job that required a college degree. Average maternal education was high school. The median household income in 9th grade was $40,000–$49,000.

Measures

In 9th, 10th, and 11th grade, students self-reported three motivational beliefs in biology, chemistry, and physics. Ability self-concept was measured using four items (e.g., 'how good at [biology/chemistry/physics] are you?'; α = 0.89–0.93). Interest was measured with two items [e.g., 'how much do you like [biology/chemistry/physics]?'; Spearman's rho = 0.74–0.88 ($p < 0.001$)] and utility was measured with three items (e.g., 'how useful is what you learn in [biology/chemistry/physics]?'; α = 0.88–0.94). All nine measures (i.e., three motivational beliefs by three subjects) were on 7-point Likert scales. The scales were used in prior studies in accordance to the expectancy-value theory (e.g., Simpkins et al., 2015a). These scales are invariant across Latina, Latino, white male, and white female high school students (Simpkins et al., 2015b). See **Supplementary Table S1** for the list of all motivational beliefs items.

Perceived science support was a composite of 15 items (see **Supplementary Table S1** for a list of all items), each asked in regard to support provided by their parents, older sibling or cousin, science teacher, and same-grade school friends on a 5-point scale (1 = *never*, 5 = *always*). The scale has been used in prior studies (Bouchey and Harter, 2005; Simpkins et al., 2015b, 2018). For each of the 15 items, participants' responses were recoded to 1 if they rated "sometimes" (i.e., the mid-point of the scale) or higher from at least one of the four social agents. That is, participants' responses were recoded to 0 if they did not rate "sometimes" or more from any of the four social agents. The 15 items were then summed to create a scale of perceived science support ranging from 0 to 15, with higher scores indicating more perceived support from at least one of their social agents (α = 0.89). We also conducted a series of follow up analyses examining perceived science support from each *specific* social agent, for which participants' responses on each item were also recoded using the "sometimes" cutoff into 0 and 1. The 15 items were then summed within each social agent, yielding four scales that all ranged from 0 to 15.

A range of demographic variables were included in the current study. Adolescents reported their gender and whether they were born in the United States. Parents also reported their own birth country, family income (categorically on the scale of $10,000; 0 = "*less than $10,000*", 10 = "*over 100,000*"), and level of education (six categories ranging from less than high school to professional degree). Adolescents were coded as third generation if they and at least one of their parents was born in the United States, otherwise adolescents were coded as first or second generation. Two dummy variables were generated to control for the three schools that participating adolescents attended in 9th grade.

Plan of Analysis

Among the 104 high school students recruited in the study, 86 had complete data on all focal variables from 9th to 11th grade. We examined if students with and without complete data across the three time points differed on family income, maternal education, gender, immigrant generational status, 9th grade science grade, 9th grade science motivational beliefs, and 9th grade perceived support. Results indicated that there were no significant differences between the two groups on family income [$t(101) = -0.24$, $p = 0.81$, Cohen's $d = 0.06$], gender ([$X^2(1)$] = 0.15, Cramer's $V = 0.04$), immigrant generational status ([$X^2(1)$] = 0.19, Cramer's $V = 0.04$), 9th grade science grade [$t(101) = 0.13$, $p = 0.89$, Cohen's $d = 0.02$], and overall perceived support [$t(102) = 0.43$, $p = 0.66$, Cohen's $d = 0.12$]. There was a small difference in maternal education [$t(98) = -1.13$, $p = 0.26$, Cohen's $d = 0.30$], and in two of the nine motivational beliefs. Specifically, adolescents with complete data across the 3 years had higher value of physics utility [$t(100) = 2.66$, $p = 0.01$, Cohen's $d = 0.68$] and were more interested in chemistry at 9th grade [$t(98) = 2.46$, $p = 0.02$, Cohen's $d = 0.63$] than adolescents with missing data.

Research Question 1 tested how science motivational beliefs changed from 9th to 11th grade in HLM version 7.3 (Raudenbush et al., 2017). Twelve linear growth curve models (i.e., time nested within person) were estimated for each of the three motivational beliefs (i.e., ability self-concept, interest, and utility) in science overall and separately in each of the three specific science subjects (i.e., biology, chemistry, and physics).

Research Questions 2 and 3 examined whether students' science motivational beliefs differed by gender, immigrant generation status, and perceived science support, respectively. Gender, immigrant generation status, and perceived science support were added as predictors of the intercept and slope that evidenced significant variance from the HLMs described under Research Question 1. A significant predictor of the intercept suggests that adolescents' motivational beliefs at 9th

[1]Other countries of origin include: El Salvador, Cuba, Guatemala, Puerto Rica, and Peru.

grade differed based on that predictor, whereas a significant predictor of the slope suggests that the changes in students' motivational beliefs differed based on that predictor. Family income level, parent education level, and the school that the adolescents attended were added as covariates.

TABLE 1 | Descriptive statistics.

	9th grade		10th grade		11th grade	
	Mean	***SD***	**Mean**	***SD***	**Mean**	***SD***
Biology						
Ability self-concept	4.22	1.10	4.76	0.95	4.66	1.05
Interest	4.81	1.32	4.91	1.42	4.97	1.28
Utility	4.98	1.17	4.91	1.16	4.67	1.25
Chemistry						
Ability self-concept	4.25	1.05	4.26	1.00	4.18	1.18
Interest	4.41	1.34	4.35	1.43	4.41	1.49
Utility	4.69	1.13	4.50	1.38	4.47	1.39
Physics						
Ability self-concept	4.38	1.14	4.19	1.03	4.03	0.95
Interest	4.56	1.31	4.38	1.23	4.39	1.23
Utility	4.79	1.15	4.60	1.32	4.42	1.30
Predictors						
Female	0.40	0.49				
Third generation	0.39	0.49				
Perceived science support	12.69	3.29				

SD, standard deviation.

RESULTS

Table 1 and **Supplementary Table S2** presents the descriptive statistics and bivariate correlations of all variables used in this study.

Changes in Adolescents' Science Motivational Beliefs

Results from hierarchical linear models are presented in **Figure 1** and **Table 2**. Results suggest that the mean of adolescents' 9th grade motivational beliefs fell between 4.19 and 4.90, which corresponded to just above the mid-point on our 7-point Likert scale. The mean of the slope indicates the change in adolescents' motivational beliefs where significant positive values suggest increases, significant negative values suggest decreases, and non-significant values suggest stability (i.e., no change). The findings for utility value and interest were consistent across the three science subjects though the patterns of change varied across the two types of beliefs. Adolescents' utility values decreased for all science subjects from 9th to 11th grade (β = −0.16, −0.16, −0.14, −0.21, respectively, for science, biology, chemistry, and physics, $p < 0.05$). In contrast, adolescents' interest remained stable for all science subjects (β = −0.02, 0.07, −0.02, −0.11, *ns*). The patterns for adolescents' ability self-concepts varied by science subject. Adolescents' ability self-concepts in science overall remained stable over time (β = −0.02, *ns*). However, adolescents' ability self-concept decreased for physics (β = −0.19, $p < 0.01$), remained stable for chemistry (β = −0.05, *ns*), and increased for biology (β = 0.20, $p < 0.001$). Our results showed that if we simply

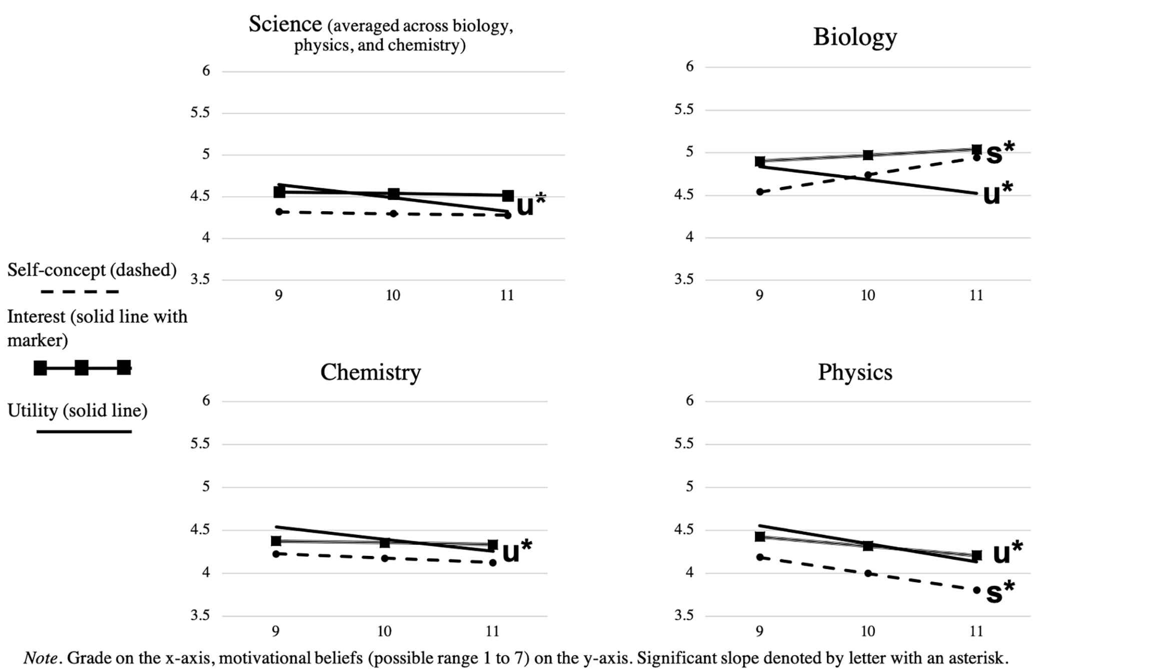


FIGURE 1 | Changes in Science, Biology, Chemistry, and Physics Motivational Beliefs. Grade on the *x*-axis, motivational beliefs (possible range 1–7) on the *y*-axis. Significant slope denoted by letter with an asterisk.

grouped physics, chemistry, and biology together as 'science,' the opposing changes in adolescents' ability self-concept by subject would have gone undetected.

Variances of intercept and slope denote if there were interindividual differences among adolescents in terms of their motivational beliefs at 9th grade and in the changes of their motivational beliefs over time (**Table 2**). There were significant interindividual differences in adolescents' motivational beliefs in all science subjects in 9th grade and in the rate of change from 9th to 11th grade in most subjects. The variance of the slope was not statistically significant for biology ability self-concepts and chemistry interest, suggesting that adolescents did not differ significantly from each other in the rate of change on these two beliefs. Biology ability self-concept and chemistry interest, therefore, were not included in the subsequent predictive analyses. Although some slopes had non-significant means (e.g., chemistry ability self-concept, biology interest, and physics interest), meaning that the adolescents *on average* did not increase or decrease in those motivational beliefs over time, that should not be confused with the presence of significant *variance* in slope, which means that there were significant differences *between adolescents* in their change in the motivational beliefs.

Predictors of Students' Motivational Beliefs at 9th Grade and Change Over Time

We tested gender, immigrant generational status, and perceived science support as predictors to examine differences among

TABLE 2 | Estimates from hierarchical linear models.

	Mean at 9th grade (intercept)	Intercept variance	Change from 9th to 11th grade (slope)	Slope variance
Science				
Ability self-concept	4.32 (0.07)***	0.49***	−0.02 (0.05)	0.11***
Interest	4.56 (0.10)***	0.79***	−0.02 (0.05)	0.09***
Utility	4.65 (0.10)***	0.90***	−0.16 (0.05)**	0.09***
Biology				
Ability self-concept	4.54 (0.08)***	0.39***	0.20 (0.06)***	0.02
Interest	4.90 (0.11)***	0.97***	0.07 (0.07)	0.16***
Utility	4.84 (0.10)***	0.84***	−0.16 (0.06)*	0.18***
Chemistry				
Ability self-concept	4.23 (0.08)***	0.52***	−0.05 (0.07)	0.24***
Interest	4.38 (0.11)***	0.91***	−0.02 (0.08)	0.13
Utility	4.54 (0.11)***	1.05***	−0.14 (0.07)*	0.14**
Physics				
Ability self-concept	4.19 (.09) ***	0.65 ***	−0.19 (0.05)***	0.09**
Interest	4.43 (.10) ***	0.91 ***	−0.11 (0.07)	0.19***
Utility	4.56 (.11) ***	1.10 ***	−0.21 (0.06) ***	0.12***

*Coefficients are unstandardized. Science refers to the average of biology, chemistry, and physics. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.*

TABLE 3 | Gender and perceived support predicting changes in science motivational beliefs.

		Science			Biology			Chemistry			Physics		
		Ability self-concept	Interest	Utility	Ability self-concept	Interest	Utility	Ability self-concept	Interest	Utility	Ability self-concept	Interest	Utility
Female	On intercept (*SE*)	−0.17 (0.15)	−0.46 (0.19)*	−0.20 (0.20)	0.06 (0.17)	−0.14 (0.23)	0.17 (0.20)	−0.29 (0.17)	−0.62 (0.23)**	−0.30 (0.22)	−0.23 (0.17)	−0.63 (0.21)**	−0.49 (0.22)*
	On slope (*SE*)	−0.11 (0.09)	−0.15 (0.11)	0.03 (0.11)	–	−0.21 (0.16)	−0.05 (0.13)	−0.09 (0.14)	–	0.15 (0.15)	−0.07 (0.11)	−0.24 (0.15)	0.02 (0.12)
Perceived support	On intercept (*SE*)	0.04 (0.02)*	0.07 (0.03)**	0.10 (0.02)***	0.05 (0.02)**	0.08 (0.03)**	0.10 (0.02)***	0.03 (0.02)	0.05 (0.03)	0.08 (0.03)**	0.04 (0.02)	0.07 (0.03)**	0.10 (0.03)***
	On slope (*SE*)	−0.01 (0.01)	−0.02 (0.01)	−0.00 (0.01)	–	−0.03 (0.02)	−0.01 (0.02)	−0.01 (0.02)	–	0.00 (0.02)	−0.02 (0.01)	−0.02 (0.02)	0.00 (0.02)

*School (2 dummy variables), family income, and parent educational level were included as covariates. Immigrant generation status was also included as predictor, but the results are not shown. SE, standard error. – means insignificant variance to predict interindividual differences. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.*

Latino/a students in the trajectories of their science motivational beliefs (**Table 3**). Contrary to our expectations, our findings suggested minimal gender differences in students' science motivational beliefs at 9th grade and changes over time. Latinas reported lower physics interest and utility (β = −0.62, −0.49, $p < 0.05$) in addition to lower chemistry interest (β = −0.62, $p < 0.01$) than Latinos at the start of high school (i.e., 9th grade). There were no gender differences at 9th grade for the other motivational beliefs (β ranged −0.30 to 0.17, *ns*). Latinas also did not differ from Latinos in how much their science motivational beliefs changed from 9th to 11th grade (β ranged −0.25 to 0.02, *ns*). Although most of the coefficients were negative, meaning that Latinas on average reported faster declines than Latinos, those differences were not statistically significant.

We expected third generation Latino/as would report lower science motivational beliefs in 9th grade and faster declines than first and second generation Latino/as. However, those differences were not statistically significant. Third generation students did not differ on the level of their science motivational beliefs at 9th grade (β ranged −0.21 to 0.50, *ns*), nor did they differ on the changes in these beliefs from 9th to 11th grade compared with first and second generation students (β ranged −0.27 to 0.18, *ns*).

We expected adolescents who perceived higher science support to have higher science motivational beliefs at 9th grade and evidence slower declines over the next 2 years. When support from multiple social agents was examined as a composite, adolescents' perceived science support at 9th grade positively predicted concurrent science motivational beliefs, but not the changes in their beliefs over time. Adolescents who perceived higher science support from key social agents in 9th grade also had higher motivational beliefs for science overall and for all three science subjects (β ranged 0.04 to 0.10, $p < 0.05$), except chemistry interest (β = 0.05, $p = 0.06$) as well as chemistry and physics ability self-concept (β = 0.03, 0.04, *ns*; **Table 3**). Adolescents' perceived science support in 9th grade, however, did not predict rate of change in any of the motivational beliefs tested (β ranged −0.03 to 0.01, *ns*).

In addition to examining perceived science support as a composite across the four social agents, we conducted follow up analyses to investigate the association between perceived support from each social agent and students' science motivational beliefs. Specifically, perceived science support from parents, teachers, friends, and siblings/cousins were analyzed in separate models to predict adolescents' science motivational beliefs while holding the same set of demographic covariates constant. Results (**Supplementary Table S3**) largely aligned with that when perceived science support was operationalized as a composite of the four social agents. Specifically, most science motivational beliefs in 9th grade were positively associated with concurrent perceived science support from parents (β ranged 0.05 to 0.07, $p < 0.05$; except ability self-concept and interest in chemistry and physics), siblings/cousins (β ranged 0.03 to 0.08, $p < 0.05$; except physics ability self-concept and interest in biology and physics), and friends (β ranged 0.04 to 0.07, $p < 0.05$). Perceived science support from teachers positively predicted utility in all three science subjects (β = 0.07, $p < 0.05$) and biology ability self-concept (β = 0.04, $p < 0.01$). Also aligning with results when perceived science support was operationalized as composite of the four social agents, how much the adolescents' science motivational beliefs changed over time was largely not predicted by perceived science support from parents, siblings/cousin, friends, or teachers. The only exceptions were that higher perceived support from parents and friends predicted faster declines in adolescents' physics ability self-concept, and that higher perceived support from teacher predicted faster decline in biology interest.

DISCUSSION

Students' motivational beliefs are malleable and change over time. Thus, it is essential to examine them with a developmental perspective instead of in a snapshot (i.e., single timepoint measurement). The current study examined the stability and change in United States Latino/a adolescents' science motivational beliefs during the first 3 years of high school. Results suggested that students' utility declined for chemistry, physics, and biology, whereas interest in all three subjects remained stable. Students' ability self-concepts decreased for physics, remained stable for chemistry, and increased for biology from 9th to 11th grade. Our findings mostly align with prior work that found decline and stability in adolescents' motivational beliefs in physical sciences (Wang et al., 2017), and we expand the literature by including biology. These findings highlight the need to examine specific science subjects and specific science motivational beliefs. Combining physics, chemistry, and biology simply as science would have masked the divergent changes in students' ability self-concepts over time. In addition to differentiating between subjects within the sciences, our study also differentiated motivational beliefs in terms of interest, utility value, and ability self-concept. Although these motivational beliefs are often correlated with one another, they are theoretically distinct and serve different functions (Wigfield et al., 2015).

The minimal gender differences in the changes of motivational beliefs was unexpected given the gender differences found in prior work (e.g., Simpkins et al., 2015b). In the current study, Latinos reported higher motivational beliefs at the beginning of high school than Latinas only for chemistry interest, physics interest, and physics utility, but not for the other motivational beliefs. Our findings aligned better with Hyde's gender similarities hypothesis that demonstrated more within-gender than between-gender differences regarding science and math (Hyde and Linn, 2006). The gender similarities hypothesis points out that although gender difference might occur at the mean level, the female and male *distributions* of science and math achievement overlap more than they differ (Hyde, 2005). For example, the patterns of United States high school student's math motivational beliefs showed no less difference within gender (by race/ethnicities) than between gender (Hsieh and Simpkins, 2018). The same study also showed that gender differences in math motivational beliefs among white students did not replicate for ethnic minority students such as Latino/as.

Another unexpected finding was that immigrant generation status did not predict students' science motivational beliefs at 9th grade or the changes over time. Prior studies have pointed out that the immigrant paradox is more likely to emerge under certain conditions. For example, the paradox is more pronounced in schools with a more negative school climate, such that, third generation ethnic minority youth are more susceptible to negative peer influences (Greenman, 2013). Because positive school climate supports high school student's motivational beliefs (Fan and Williams, 2018), a future direction is to bridge these two bodies of literature and examine whether differences by immigrant generation status emerge for the association between school climate and adolescent's motivational beliefs.

Lastly, in regard to perceived science support, we found that students who perceived higher science support in 9th grade also tended to report higher science motivational beliefs at the same time. This finding aligns with theory and the body of literature suggesting that support from parents, teachers, friends, siblings, and other significant social agents are assets for Latino/a students' educational trajectories (Bouchey and Harter, 2005; Plunkett et al., 2008). Students' perceived science support at 9th grade, however, did not consistently predict *changes* in their science motivational beliefs from 9th to 11th grade. Our finding that perceived science support at 9th grade predicted initial level but not change in science motivational beliefs aligned with Alfaro and Umaña-Taylor (2015) findings on general academic support and motivational beliefs. Perhaps it is not enough to only look at support at one time point. It might be the case that changes in support over time would predict changes in motivational beliefs. However, we are limited by the sample size to explore this possibility. We did a series of follow up analyses to examine the associations between adolescents' science motivational beliefs and their perceived science support from each of the social agents separately. Those results largely aligned with the conclusions when perceived science support was operationalized as a composite from the four social agents.

Potential Implications

Overall, the results of the current study have potential research and practical implications. In terms of research implications, we provided empirical evidence for the need to distinguish among students' motivational beliefs (i.e., ability self-concept, interest, and utility) and among specific science subjects (i.e., chemistry, physics, and biology). Additionally, we tested the expectancy-value theory within an often understudied ethnic population and showed that some, but not all, of the hypothesized mechanisms emerged for our Latino/a sample. Therefore, scholars should continue to interrogate when expected patterns generalize to other groups and under what circumstances they do not generalize and why that might be.

In terms of practical implications, our results suggest that the importance Latino/a high school students attach to biology, chemistry, and physics declines from 9th to 11th grade as does their beliefs about their physics abilities. These patterns mostly emerged regardless of adolescents' gender and immigrant generation status. We envision two immediate implications for schools. First, schools should consider how to bolster the importance students attribute to these areas of science. Harackiewicz et al. (2012) bolstered parents' and high school students' value of science by providing informational resources which had real impacts on the courses students took and their subsequent college majors. Second, the lack of differences based on adolescents' gender and immigration generation challenges prevalent stereotypes of who pursues science (National Science Foundation [NSF], 2017). Though it is possible that our analyses were underpowered to detect these differences and our findings need to be replicated, it is also possible that the typical patterns might not hold true for a specific and often understudied subgroups (e.g., Hsieh and Simpkins, 2018). These divergent findings are a cautionary message for both schools and researchers. Schools are likely to be more effective if they tailor to the lived experiences of their students in their specific communities instead of relying on the aggregated experiences of students from the national level. Relatedly, more *within*-group studies are needed before we use findings on the average population to problematize a specific ethnic, racial, immigrant, or gender groups and deepen stereotypical images of them.

Adolescents' perceived science support from parents, siblings/cousins, teachers, and friends was positively related to their concurrent science motivational beliefs at the beginning of high school– highlighting the importance of encouraging support from those social agents. Though people may experience challenges in supporting adolescents if they did not take the same subjects or if adolescents' schooling surpasses that of the person providing the support, it is important to note that there are many ways people can be supportive. Our measure of perceived support includes some aspects that do not necessarily require much prior knowledge of science (e.g., help you feel better when science is hard). Thus, parents and other social agents who have varying levels of science knowledge and educational capital can be empowered to serve as positive social agents. Although the socio-political context often poses barriers for Latino/a adolescents and their families (e.g., Lin et al., 2016), our study showed with a strength-based perspective that they have the potential to succeed with support from the people around them.

Limitations and Future Directions

Although one strength of the current study is its longitudinal nature, studies spanning longer time frames would provide a more complete description of development. The current study speaks to high school, which has been shown to be a crucial developmental period in regard to STEM pathways and science motivational beliefs (Sadler et al., 2012), but future studies could expand to include earlier and subsequent developmental periods (e.g., Jacobs et al., 2002; Robinson and Lubienski, 2011) and crucial *transitional periods* such as the transition from middle to high school or the post high school transition to college or work.

Although our within-group focus is a strength and extends the current literature, the Latino/a adolescents in our study were diverse on some indicators but more homogeneous on other indicators (e.g., United States region). We examined heterogeneity among the participants in terms of gender, immigrant generational status, and perceived science support. Future studies could incorporate other factors that are relevant to the Latino/a population in the United States, such as socioeconomic status, country of origin, which United States region they live in, language ability, sense of family obligation, school ethnic composition, legal status, or the political climate (Goldsmith, 2004; Suárez-Orozco and Carhill, 2008; Greenman and Hall, 2013). For example, although prior studies mostly showed positive association between socioeconomic status and academic achievement for the Latino/a population (Altschul, 2012), the association with motivational beliefs might not be as strong (St-Hilaire, 2002) and associations between socioeconomic status and motivational beliefs of *specific science subjects* remain underexplored. Another example is that Latino/a youth's educational experiences differ depending on where they live in the United States. While Latino/a youth in regions with a smaller Latino/a population and new receiving communities tend to face more discrimination than those in minority-majority regions, their identification with the local Latino/a community could promote their academic motivational beliefs (Perreira et al., 2010). Taken together, future studies could build on ours by examining other mechanisms that contribute to the diversity among Latino/as in the United States. The current sample size was modest and might have rendered some of our analyses under-powered. Studies that address a variety of indicators will require larger sample sizes than the current one to have adequate power to detect differences among multiple indicators of within-group heterogeneity.

We found that perceived science support positively predicted some aspects of science motivational beliefs when support was operationalized as the joint support from multiple social agents. The follow up analyses that examined perceived science support from each social agent in separate models also mostly point to the same conclusion. We presented the former as main analyses and the latter as follow up because we believe perceived science support as jointly from multiple social agents better reflect the experiences of adolescents as they are simultaneously interacting with the multiple social agents. To expand on the conceptualization of perceived science support, future studies could go into the nuances by differentiating the *patterns* of support from multiple social agents (Furrer and Skinner, 2003; Simpkins et al., 2018). Relatedly, perceived science support as operationalized by the current study consisted of both instructional support (e.g., help enroll the adolescent in science lessons, workshops, or tutoring programs outside of class) and motivational support (e.g., help the adolescent feel better when science is hard), but future studies could examine whether one kind of support is more predictive of science motivational beliefs than the other. Lastly, although we distinguished motivational beliefs by specific science subjects (i.e., chemistry, physics, and biology), our measure of perceived science support focused on science instead of specific science subject. This could have weakened the relations between perceived support and adolescents' motivational beliefs, particularly when considering changes over time as students take different science subjects over time. People's support may vary based on the science subject.

Also regarding our measures, the first two waves of data collection largely happened in adolescents' homes whereas the last wave of data collection happened in their school. Though the means for the 11th grade motivational beliefs compared to 9th or 10th grade motivational beliefs do not suggest there was a large impact of context on the data, the change in context is confounded with the adolescents progressing from 10th to 11th grade. Theoretically, the context in which data are collected could impact students in positive or negative ways depending on the climate of their home and school contexts. Future studies could examine the impact of context by comparing the same data collected in various natural or experimental contexts.

CONCLUSION

For our sample of 104 Latino/a adolescents, their motivational beliefs (i.e., interest, utility, and ability self-concept) in physics, chemistry, and biology either decreased or remained stable from 9th to 11th grade, except the increases in their biology ability self-concept. Our findings highlighted the need to differentiate both among science subjects and among motivational beliefs. Adolescents' science motivational beliefs at 9th grade and the changes in those beliefs over time largely did not differ by gender or generation status. Adolescents' perceived science support explained some differences in their science motivational beliefs at 9th grade. Overall, our study contributes to the literature by examining subject-specific motivational beliefs and within-group differences of an often understudied ethnic group. Additionally, we incorporated both demographic (i.e., gender, generation status) and contextual factors (i.e., perceived science support) that are theorized to be salient determinants of students' science motivational beliefs.

AUTHOR CONTRIBUTIONS

T-yH, YL, and SS contributed to conceptualizing the study. T-yH wrote the first draft. YL and SS co-authored this study and helped to revise all drafts. SS conceived of the study from which the current data were drawn.

FUNDING

This research was supported by the National Science Foundation under grants DRL-1054798 and DRL-1560889 to SS.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00380/full#supplementary-material

REFERENCES

Adya, M., and Kaiser, K. M. (2005). Early determinants of women in the IT workforce: a model of girls' career choices. *Inf. Technol. People* 18, 230–259. doi: 10.1108/09593840510615860

Alfaro, E. C., and Umaña-Taylor, A. J. (2010). Latino adolescents' academic motivation: the role of siblings. *Hisp. J. Behav. Sci.* 32, 549–570. doi: 10.1177/0739986310383165

Alfaro, E. C., and Umaña-Taylor, A. J. (2015). The longitudinal relation between academic support and Latino adolescents' academic motivation. *Hisp. J. Behav. Sci.* 37, 219–241. doi: 10.1177/0739986315586565

Alfaro, E. C., Umaña-Taylor, A. J., and Bámaca, M. Y. (2006). The influence of academic support on Latino adolescents' academic motivation. *Fam. Relat.* 55, 279–291. doi: 10.1111/j.1741-3729.2006.00402.x

Altschul, I. (2012). Linking socioeconomic status to Mexican American youth's academic achievement through parent involvement in education. *J. Soc. Soc. Work Res.* 3, 13–30. doi: 10.5243/jsswr.2012.2

Aretakis, M. T., Ceballo, R., Suarez, G. A., and Camacho, T. C. (2015). Investigating the immigrant paradox and Latino adolescents' academic attitudes. *J. Lat. Psychol.* 3, 56–69. doi: 10.1037/lat0000031

Bennett, J., and Hogarth, S. (2009). Would you want to talk to a scientist at a party? High school students' attitudes to school science and to science. *Int. J. Sci. Educ.* 31, 1975–1998. doi: 10.1080/09500690802425581

Bouchey, H. A., and Harter, S. (2005). Reflected appraisals, academic self-perceptions, and math/science performance during early adolescence. *J. Educ. Psychol.* 97, 673–686. doi: 10.1037/0022-0663.97.4.673

Britner, S. L. (2008). Motivation in high school science students: a comparison of gender differences in life, physical, and earth science classes. *J. Res. Sci. Teach.* 45, 955–970. doi: 10.1002/tea.20249

Buday, S. K., Stake, J. E., and Peterson, Z. D. (2012). Gender and the choice of a science career: The impact of social support and possible selves. *Sex Roles* 66, 197–209. doi: 10.1007/s11199-011-0015-4

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). "Expectancies, values, and academic behaviors," in *Achievement and Achievement Motivation*, ed. J. T. Spence (San Francisco, CA: W. H. Freeman), 75–146.

Else-Quest, N., Mineo, C. C., and Higgins, A. (2013). Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychol. Women Q.* 37, 293–309. doi: 10.1177/0361684313480694

Fan, W., and Williams, C. (2018). The mediating role of student motivation in the linking of perceived school climate and achievement in reading and mathematics. *Front. Educ.* 3:50. doi: 10.3389/feduc.2018.00050

Feliciano, C., and Lanuza, Y. R. (2017). An immigrant paradox? Contextual attainment and intergenerational educational mobility. *Am. Sociol. Rev.* 82, 211–241. doi: 10.1177/0003122416684777

Fredricks, J., and Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: growth trajectories in two male sex-typed domains. *Dev. Psychol.* 38, 519–533. doi: 10.1037/0012-1649.38.4.519

Fry, R., and Gonzales, F. (2008). *One-in-Five and Growing Fast: A Profile of Hispanic Public School Students*. Washington, DC: Pew Hispanic Center.

Furrer, C., and Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *J. Educ. Psychol.* 95, 148–162. doi: 10.1037/0022-0663.95.1.148

García Coll, C., Patton, F., Marks, A. K., Dimitrova, R., Yang, H., Suarez-Aviles, G., et al. (2012). "Understanding the immigrant paradox in youth: developmental considerations," in *Realizing the Potential of Immigrant Youth*, eds A. S. Masten, K. Liebkind, and D. G. Hernandez (Cambridge: Cambridge University Press), 159–180. doi: 10.1017/CBO9781139094696.009

Goldsmith, P. A. (2004). Schools' racial mix, students' optimism, and the black-white and latino-white achievement gaps. *Soc. Educ.* 77, 121–147. doi: 10.1177/003804070407700202

Gottfried, A. E., Fleming, J. S., and Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: a longitudinal study. *J. Educ. Psychol.* 93, 3–13. doi: 10.1037/0022-0663.93.1.3

Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., and Oliver, P. H. (2009). A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation. *J. Educ. Psychol.* 101, 729–739. doi: 10.1037/a0015084

Greenman, E. (2013). Educational attitudes, school peer context, and the "immigrant paradox" in education. *Soc. Sci. Res.* 42, 698–714. doi: 10.1016/j.ssresearch.2012.12.014

Greenman, E., and Hall, M. (2013). Legal status and educational transitions for Mexican and Central American immigrant youth. *Soc. Forces* 91, 1475–1498. doi: 10.1093/sf/sot040

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., and Hyde, J. S. (2012). Helping parents to motivate adolescents in mathematics and science: an experimental test of a utility-value intervention. *Psychol. Sci.* 23, 899–906. doi: 10.1177/0956797611435530

Harris, A. L., Jamison, K. M., and Trujillo, M. H. (2008). Disparities in the educational success of immigrants: an assessment of the immigrant effect for Asians and Latinos. *Ann. Am. Acad. Pol. Soc. Sci.* 620, 90–114. doi: 10.1177/0002716208322765

Hsieh, T., and Simpkins, S. D. (2018). A pattern-centered and intersectional approaches to the gender and racial differences in high school math motivation: associations with math achievement and engagement. Paper Presented at the Gender & STEM Network Conference, Eugene, OR.

Hyde, J. S. (2005). The gender similarities hypothesis. *Am. Psychol.* 60, 581–592. doi: 10.1037/0003-066X.60.6.581

Hyde, J. S., and Linn, M. C. (2006). Gender similarities in mathematics and science. *Science* 314, 599–600. doi: 10.1126/science.1132154

Jacobs, J. E., Davis-Kean, P., Bleeker, M., Eccles, J. S., and Malanchuk, O. (2005). ""I can, but I don't want to": the impact of parents, interests, and activities on gender differences in math," in *Gender Differences in Mathematics: An Integrative Psychological Approach*, eds A. M. Gallagher and J. C. Kaufman (New York, NY: Cambridge University Press), 246–263.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., and Wigfield, A. (2002). Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Dev.* 73, 509–527. doi: 10.1111/1467-8624.00421

Knight, G. P., Gonzales, N. A., Saenz, D. S., Bonds, D., Germán, M., Deardorff, J., et al. (2010). The mexican american cultural values scale for adolescents and adults. *J. Early Adolesc.* 30, 444–481. doi: 10.1177/0272431609338178

Knight, G. P., Roosa, M. W., and Umaña-Taylor, A. J. (2009). *Studying ethnic Minority and Economically Disadvantaged Populations: Methodological Challenges and Best Practices*. Washington, D.C: American Psychological Association, doi: 10.1037/11887-000

Kurtz-Costes, B., Rowley, S. J., Harris-Britt, A., and Woods, T. A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *Merrill Palmer Q.* 54, 386–409. doi: 10.1353/mpq.0.0001

Lau, S., and Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educ. Assessment* 8, 139–162. doi: 10.1207/S15326977EA0802_04

Lazarides, R., and Ittel, A. (2013). Mathematics interest and achievement: what role do perceived parent and teacher support play? A longitudinal analysis. *Int. J. Gender Sci. Technol.* 5, 207–231.

Lazarides, R., and Raufelder, D. (2017). Longitudinal effects of student-perceived classroom support on motivation: a latent change model. *Front. Psychol.* 8:417. doi: 10.3389/fpsyg.2017.00417

Lazarides, R., Rubach, C., and Ittel, A. (2017). Adolescents' perceptions of socializers' beliefs, career-related conversations, and motivation in mathematics. *Dev. Psychol.* 53, 525–539. doi: 10.1037/dev0000270

Lazarides, R., and Watt, H. M. G. (2017). Student-perceived mothers' and fathers' beliefs, mathematics and English motivations, and career choices. *J. Res. Adolesc.* 27, 826–841. doi: 10.1111/jora.12317

Leaper, C., Farkas, T., and Brown, C. S. (2012). Adolescent girls' experiences and gender-related beliefs in relation to their motivation in math/science and English. *J. Youth Adolesc.* 41, 268–282. doi: 10.1007/s10964-011-9693-z

Lepper, M. R., Corpus, J. H., and Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: age differences and academic correlates. *J. Educ. Psychol.* 97, 184–196. doi: 10.1037/0022-0663.97.2.184

Lepper, M. R., and Henderlong, J. (2000). "Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation," in *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance*, eds C. Sansone and J. M. Harackiewicz (San Diego, CA: Academic Press), 257–307. doi: 10.1016/B978-012619070-0/50032-5

Lewis, J., Ream, R. K., Bocian, K. M., Cardullo, R. A., Hammond, K. A., and Fast, L. A. (2012). Con cariño: teacher caring, math self-efficacy, and math achievement among Hispanic English learners. *Teach. Coll. Rec.* 114, 1–42.

Lin, A. R., Menjivar, C., Ettekal, A. V., Simpkins, S. D., Gaskin, E. R., and Pesch, A. (2016). "They will post a law about playing soccer" and other ethnic/racial microaggressions in organized activities experienced by Mexican-origan families. *J. Adolesc. Res.* 31, 557–581. doi: 10.1177/0743558415620670

Maltese, A. V., and Tai, R. H. (2011). Pipeline persistence: examining the association of educational experiences with earned degrees in STEM among U.S. students. *Sci. Educ.* 95, 877–907. doi: 10.1002/sce.20441

Nagengast, B., and Marsh, H. W. (2012). Big fish in little ponds aspire more: mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *J. Educ. Psychol.* 104, 1033–1053. doi: 10.1037/a0027697

National Science Foundation [NSF] (2017). *National Center for Science and Engineering Statistics. 2017. Women, Minorities, and Persons with Disabilities in Science and Engineering: 2017*. Special Report NSF 17-310, Arlington, VA: National Science Foundation.

Osborne, J., Simon, S., and Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *Int. J. Sci. Educ.* 25, 1049–1079. doi: 10.1080/0950069032000032199

Palacios, N., Guttmannova, K., and Chase-Lansdale, P. L. (2008). Early reading achievement of children in immigrant families: is there an immigrant paradox? *Dev. Psychol.* 44, 1381–1395. doi: 10.1037/a0012863

Perreira, K. M., Fuligni, A., and Potochnick, S. (2010). Fitting in: the roles of social acceptance and discrimination in shaping the academic motivations of Latino youth in the U.S. Southeast. *J. Soc. Issues* 66, 131–153. doi: 10.1111/j.1540-4560.2009.01637.x

Plunkett, S. W., and Bámaca-Gómez, M. Y. (2003). The relationship between parenting, acculturation, and adolescent academics in Mexican-origin immigrant families in Los Angeles. *Hisp. J. Behav. Sci.* 25, 222–239. doi: 10.1177/0739986303025002005

Plunkett, S. W., Henry, C. S., Houltberg, B. J., Sands, T., and Abarca-Mortensen, S. (2008). Academic support by significant others and educational resilience in Mexican-origin ninth grade students from intact families. *J. Early Adolesc.* 28, 333–355. doi: 10.1177/0272431608314660

Raudenbush, S. W., Bryk, A. S., and Congdon, R. (2017). *HLM 7.03 for Windows [Computer software]*. Skokie, IL: Scientific Software International, Inc.

Rice, L., Barth, J. M., Guadagno, R. E., Smith, G. P., and McCallum, D. M. (2013). The role of social support in students' perceived abilities and attitudes toward math and science. *J. Youth Adolesc.* 42, 1028–1040. doi: 10.1007/s10964-012-9801-8

Riconscente, M. M. (2014). Effects of perceived teacher practices on Latino high school students' interest, self-efficacy, and achievement in mathematics. *J. Exp. Educ.* 82, 51–73. doi: 10.1080/00220973.2013.813358

Robinson, J. P., and Lubienski, S. T. (2011). The development of gender achievement gaps in mathematics and reading during elementary and middle school: examining direct cognitive assessments and teacher ratings. *Am. Educ. Res. J.* 48, 268–302. doi: 10.3102/0002831210372249

Sadler, P. M., Sonnert, G., Hazari, Z., and Tai, R. (2012). Stability and volatility of STEM career interest in high school: a gender study. *Sci. Educ.* 96, 411–427. doi: 10.1002/sce.21007

Sax, L., Kanny, M., Riggers-Piehl, T., Whang, H., and Paulson, L. (2015). 'But I'm not good at math': the changing salience of mathematical self-concept in shaping women's and men's STEM aspirations. *Res. High. Educ.* 56, 813–842. doi: 10.1007/s11162-015-9375-x

Simpkins, S., Liu, Y., Hsieh, T., and Estrella, G. (2018). *Supporting Latino High School Students' Science Motivation: Examining the Unique and Collective Contributions of Family, Teachers, and Friends*. Washington, DC: AERA.

Simpkins, S. D., Davis-Kean, P. E., and Eccles, J. S. (2006). Math and science motivation: a longitudinal examination of the links between choices and beliefs. *Dev. Psychol.* 42, 70–83. doi: 10.1037/0012-1649.42.1.70

Simpkins, S. D., Fredricks, J., and Eccles, J. S. (2015a). The role of parents in the ontogeny of achievement-related motivation and behavioral choices. *Monogr. Soc. Res. Child Dev.* 80, 1–151. doi: 10.1111/mono.12156

Simpkins, S. D., Price, C. D., and Garcia, K. (2015b). Parental support and high school students' motivation in biology, chemistry, and physics: understanding differences among Latino and Caucasian boys and girls. *J. Res. Sci. Teach.* 52, 1386–1407. doi: 10.1002/tea.21246

Singh, K., Granville, M., and Dika, S. (2002). Mathematics and science achievement: effects of motivation, interest, and academic engagement. *J. Educ. Res.* 95, 323–332. doi: 10.1080/00220670209596607

Stake, J. E., and Nickens, S. D. (2005). Adolescent girls' and boys' science peer relationships and perceptions of the possible self as scientist. *Sex Roles* 52, 1–11. doi: 10.1007/s11199-005-1189-4

St-Hilaire, A. (2002). The social adaptation of children of mexican immigrants: educational aspirations beyond junior high school. *Soc. Sci. Q.* 83, 1026–1043. doi: 10.1111/1540-6237.00131

Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychol. Bull.* 135, 859–884. doi: 10.1037/a0017364

Suárez-Orozco, C., and Carhill, A. (2008). Afterword: new directions in research with immigrant families and their children. *New Dir. Child Adolesc. Dev.* 121, 87–104. doi: 10.1002/cd.224

U. S. Census Bureau, Population Division (2014). *Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States, States, and Counties: April 1, 2010 to July 1, 2013*. Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkm

Wang, M-T., Chow, A., Degol, J. L., and Eccles, J. S. (2017). Does everyone's motivational beliefs about physical science decline in secondary school?: Heterogeneity of adolescents' achievement motivation trajectories in physics and chemistry. *J. Youth Adolesc.* 46, 1821–1838. doi: 10/1007/s10964-016-0620-1

Wigfield, A., and Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., and Schiefele, U. (2015). "Development of achievement motivation and engagement," in *Handbook of Child Psychology and Developmental Science*, 7th Edn, Vol. 3, eds R. M. Lerner and M. Lamb (Hoboken, NJ: John Wiley & Sons Inc.), 657–700. doi: 10.1002/9781118963418.childpsy316

Wigfield, A., and Wagner, A. L. (2005). "Competence, motivation, and identity development during adolescence," in *Handbook of Competence and Motivation*, eds A. J. Elliot and C. S. Dweck (New York, NY: Guilford Publications), 222–239.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 05 March 2019
doi: 10.3389/feduc.2019.00014

Mothers and Fathers—Who Matters for STEM Performance? Gender-Specific Associations Between STEM Performance, Parental Pressure, and Support During Adolescence

Frances Hoferichter and Diana Raufelder**

Department School Pedagogics, Institute of Educational Science, Greifswald University, Greifswald, Germany

OPEN ACCESS

Edited by:
Manuela Paechter,
University of Graz, Austria

Reviewed by:
Sarah E. Martiny,
UiT The Arctic University of Norway, Norway
Thomas K. F. Chiu,
Faculty of Education, University of Hong Kong, Hong Kong

****Correspondence:***
Frances Hoferichter
frances.hoferichter@uni-greifswald.de
Diana Raufelder
diana.raufelder@uni-greifswald.de

Specialty section:
This article was submitted to STEM Education, a section of the journal Frontiers in Education

Received: *14 December 2018*
Accepted: *13 February 2019*
Published: *05 March 2019*

Citation:
Hoferichter F and Raufelder D (2019) Mothers and Fathers—Who Matters for STEM Performance? Gender-Specific Associations Between STEM Performance, Parental Pressure, and Support During Adolescence. Front. Educ. 4:14. doi: 10.3389/feduc.2019.00014

Research has shown that parental pressure is negatively whereas parental support is positively associated with various scholastic outcomes, such as school engagement, motivation, and achievement. However, only few studies investigate boys' and girls' perception of mother and father pressure/support in detail. This might be particularly essential when it comes to girls' and boys' achievement in STEM subjects, as girls and boys might profit differently from parental pressure/support regarding their achievement in STEM and vice versa. This study aims to shed light on this topic and explores potential within—and over time associations between students' perception of parental pressure/support and grades in mathematics and biology. Using self-report data from 1,088 8th grade students at T1 ($M_{age} = 13.70$, $SD = 0.53$, 54% girls) from Brandenburg, Germany, multigroup cross-lagged models were conceptualized with Mplus. The results indicate that there are gender differences in the interplay of students' grades in mathematics, biology, and their perception of parental pressure and support: Whereas, mother support plays a central beneficial role for girls' achievement in STEM subjects as well as for the other parental variables over time, for boys mother support is negatively associated with math performance over time. Within-time associations further show that boys—in contrast to girls—do not benefit from any parental support regarding their performance in mathematics or biology. Finally, results suggest that the relationship between adolescents' STEM achievement and parental pressure/support is rather mono-directional than bi-directional over time.

Keywords: STEM performance, gender, parental support, parental pressure, secondary school

INTRODUCTION

The need for specialized labor in the field of science, technology, engineering, and mathematics (STEM) is constantly increasing as technology accompanies daily life. Despite this demand and efforts to inspire youth to follow a STEM career, girls are significantly under-represented in STEM subjects. According to the Trends in International Mathematics and Science Study (TIMSS), in most countries, boys represent the majority of students enrolling in advanced STEM courses in secondary education (Mullis et al., 2016). According to a recent report of the United Nations Educational Scientific Cultural Organization (2017), the gender gap becomes more apparent as the educational level increases and electives are available. In fact, at the age of 10–11 years, students are almost equally interested in STEM subjects, where by the age of 18, only 33% of all boys and 19% of all girls who participated in this UK-based study were engaged in STEM (Kerney and YourLife Campaign, 2016). However, the gender gap regarding the attainment of STEM bachelor degrees after 4 years of college has been narrowing since 1977, where only 25 percent of all STEM degrees were awarded to women compared to the year 2000 in which 40 percent of all STEM degrees were obtained by women. Particularly in the fields of biology and agricultural science sex parity has been reached since the 1990s, whereas the gender gap is still striking in engineering, physical science, and math (National Science Board, 2014). Despite this trend of women receiving an equal number of degrees in the field of biology and agriculture, women prefer non-STEM degrees (Mann and DiPrete, 2013). As an explanation for the small number of girls involved in STEM, researcher mention varying interests of girls and boys (Su and Rounds, 2015), while girls excel in both language and math subjects compared to their male counterpart (Ceci et al., 2009), girls are more interested in tasks that promote symmetrical, quantitative, and verbal abilities (Su and Rounds, 2015). In this sense, the breadth-based model (Lubinski et al., 2001; Valla and Ceci, 2014) indicates that girls have broader career choices and choose careers in which they can apply people-related skills and verbal abilities, although having equal abilities to pursue a STEM career. In contrast, research on gender-specific socialization takes a different view, stating that girls' socialization presents the main factor for girls feeling inferior and less confident in STEM subjects, which are believed to represent masculine topics (Archer et al., 2013). Furthermore, stereotypes associated with STEM professions, e.g., working in STEM fields means being socially isolated, drive girls away from perusing a STEM career, as girls are socialized to interact with others, being social and pleasant (Reinking and Martin, 2018). In fact, children's believes about themselves, their ability and their attitude toward STEM education are strongly impacted by their parents as primary agent of socialization. This reasoning has been included in the General Expectancy-Value Model of Achievement Choices in which Eccles (2014) states that parents' specific beliefs and perceptions (e.g., affective reactions to child's performance, activity choice, competence and interest, parents' expectations for child's success, parents' perceptions of importance of activities and skills) as well as their specific actions and behaviors toward the child (e.g., advice, providing certain equipment, toys, and experiences for the child), impact children's motivation, activity choices, affect, interest, etc. toward STEM. This model provides a theoretical framework for a gendered bias emerging in STEM fields, despite the fact that boys and girls perform equally well in science. In line with this model, empirical work found that children had higher ability self-concepts and assigned a higher value to STEM subjects, when their parents showed positivity, co-activity, and school focused behaviors (Simpkins et al., 2015). In a qualitative study, in which Halim et al. (2018) interviewed parents of children who chose to pursue STEM education, the researchers found parental support and academic expectation to be common features of these parents. Hence, the investigated children who enrolled in STEM education had parents who supported their children in choosing STEM education, assisted them in science subjects, joined them in science-related activities and were concerned about their STEM related academic performance. Furthermore, parental emotional support and stimulating learning settings at home were mentioned to be relevant in choosing a career in STEM and develop an identity as scientist starting early in childhood (Buschor et al., 2014). In fact, support from parents varied considerably among students who maintained their interest in STEM throughout high school compared to those who lost interest in STEM (Aschbacher et al., 2010).

Although the role of parental support, expectation, and pressure have been identified as key factors for students to pursue a STEM career and to maintain interest in STEM (Dabney et al., 2013), only few studies take a more detailed view on the role of parents by examining the impact of mothers and fathers for girls and boys separately: Research shows that especially mother's beliefs about their daughter's ability in mathematics and science impact performance and career choices of girls (Gunderson et al., 2012; Rozek et al., 2015). Past research has predominantly investigated the role of mothers for the development of sons and daughters, while excluding the unique role of fathers. The reason for the lack of research on fathers' role regarding the upbringing of children may be its negative connotation and limitation to financial support (Hawkins and Dollahite, 1997; Marks and Palkovitz, 2004; Saracho and Spodek, 2008). The reason why fathers' role recedes in the background is fuelled by research findings indicating that it is the mother-child relationship which is characterized by low psychological distress (Mallers et al., 2010) and that mother's evaluate their behavior as more supportive toward their children compared to fathers (Fthenakis and Minsel, 2002). Examining the role of mothers and fathers for boys' and girls' career choices, Paa and Hawley McWhirter (2011) indicate, that in comparison to boys, girls perceived more positive feedback and autonomy support from their mother. In contrast, girls and boys equally perceived positive feedback and autonomy support from their father (Paa and Hawley McWhirter, 2011). Additionally, Fthenakis and Minsel (2002) found that fathers' of girls spend more time with their offspring compared to fathers' of boys, while fathers feel less disturbed by conflicts with girls compared to boys. Although research is limited, the few studies that focus on the role of mothers and fathers separately for boys and girls indicate that mother's and father's school related behavior are perceived differently by boys and girls. Fthenakis and Minsel (2002) found that boys compared to girls reported to receive more control and punishment concerning school issues. Similarly, in their study, Levpušček and Zupančič (2009) found that boys perceive significantly more father pressure than girls, while particularly fathers tend to differentiate between the upbringing of girls and boys (Lytton and Romney, 1991). Furthermore, past studies have mainly looked at how parental support/pressures impacts

STEM performance, but not vice versa. Hence, it is not clear, how students' STEM performance impacts their perception of parental support and pressure (bi-directional).

Based on these gender specific findings, the recent study takes a detailed view on the role of mother's and father's support and pressure for boys' and girls' academic performance in STEM subjects such as mathematics and biology. Mathematics was chosen as there is a sex disparity in the fields of engineering, physical science and math. In contrast, biology was chosen as bachelor's degrees are equally attained by male and female students. By choosing a domain that is over-presented by males (math) and a domain in which males and females are involved equally (biology), the study aims at shedding light on the differential role of mothers and fathers for girls' and boys' STEM performance and vice versa (bi-directional) by using a cross-lagged-panel design to identify potential factors that contribute or hinder academic success in the field of STEM and in turn serve as template for further research and intervention involving both mothers and fathers.

HYPOTHESIS

In order to test how mother's and father's pressure and support at Time 1 and 2 as well as STEM performance at Time 1 and 2 would relate to each other within and over time for boys and girls, a multigroup cross-lagged model was designed. In particular, the following hypotheses were tested:

Hypothesis I: Adolescent girls and boys differ regarding their perceived parental pressure/support associated with their grades in two STEM subjects (i.e., mathematics and biology) at the beginning of 8th grade and 1.5 years later at the end of 9th grade (within-time associations).

Hypothesis II: Adolescent girls and boys differ regarding the associations between their perceived parental pressure/support and their grades in two STEM subjects (i.e., mathematics and biology) and vice versa during the beginning of 8th to the end of 9th grade (over time associations).

METHODS

Participants and Procedure

This two-wave study is based on data from 1,088 8th grade students at Time 1 (T1) (M_{age} = 13.70, aged 12–15 years, *SD* = 0.53, 54% girls), who were at the end of 9th grade at Time 2 (T2) (*N* = 845; M_{age} = 14.86, aged 13–17 years, *SD* = 0.57, 55% girls). The participants were recruited from 23 randomly selected public secondary schools out of a pool of 124 public secondary schools in the federal state of Brandenburg, Germany. In order to provide a representative sample for the federal state of Brandenburg, five of the 23 schools were located in the biggest cities of the state (Potsdam, Cottbus, Frankfurt Oder, Brandenburg, and Prenzlau), while the other 18 were located in rural areas. The data collection took place in the autumn term 2011 (T1) and the spring term 2013 (T2). From T1 to T2 the dropout rate amounts 22.33% of participating students. The study focuses on this specific age group, as some studies indicate an achievement drop in school during this time period (Dohn, 1991; Wijsman et al., 2016). Initially, (1) written permission of the ethical committee of the Hoorn, Youth, and Sport (MBJS) of Brandenburg, (2) school consensus, and (3) both parents' and students' written and informed consensus was obtained. Before students filled in the paper-pencil questionnaire, experienced research instructors informed about voluntary participation and confidential treatment of responses. As there is only a small amount of ethnic diversity in Brandenburg (2.6%), data on ethnicity was not collected. The German law prohibits collecting data from a third party (i.e., asking students about their parents income or school graduation), students' socio-economic status could not be assessed.

Measures

All self-report measures used in this study are well-established instruments for German-speaking students. The reported reliability values are based on the current sample.

STEM grades were addressed by students' self-reported grades on their two most recent report cards in Mathematics ($\alpha T1_{\text{girls}}$ = 0.87; $\alpha T2_{\text{girls}}$ = 0.87; $\alpha T1_{\text{boys}}$ = 0.84; $\alpha T2_{\text{boys}}$ = 0.84) and Biology ($\alpha T1_{\text{girls}}$ = 0.83; $\alpha T2_{\text{girls}}$ = 0.83; $\alpha T1_{\text{boys}}$ = 0.84; $\alpha T2_{\text{boys}}$ = 0.84). Grades range from 1 ("very good") to 6 ("insufficient") in the German school system. For the sake of clarity, all grades were reverse-coded in the present study, such as a high score represents high achievement.

Parental Pressure was measured with the "Zurich Questionnaire of Educational Behavior" [Züricher Kurzfragebogen zum Erziehungsverhalten] (ZKE) developed by Reitzle et al. (2001) on a 4-point Likert scale from 1 ("not true at all") to 4 ("totally true"). The subscales father pressure ($\alpha T1_{\text{girls}}$ = 0.79; $\alpha T2_{\text{girls}}$ = 0.83; $\alpha T1_{\text{boys}}$ = 0.80; $\alpha T2_{\text{boys}}$ = 0.80) and mother pressure ($\alpha T1_{\text{girls}}$ = 0.77; $\alpha T2_{\text{girls}}$ = 0.80; $\alpha T1_{\text{boys}}$ = 0.78; $\alpha T2_{\text{boys}}$ = 0.79) consist of six items each (e.g., "My mother/father pushes me to work harder in school" or "My mother/father expects that I do well in school").

Parental Support was also measured with the "Zurich Questionnaire of Educational Behavior" (ZKE) (Reitzle et al., 2001). The subscales father support ($\alpha T1_{\text{girls}}$ = 0.92; $\alpha T2_{\text{girls}}$ = 0.92; $\alpha T1_{\text{boys}}$ = 0.91; $\alpha T2_{\text{boys}}$ = 0.90) and mother support ($\alpha T1_{\text{girls}}$ = 0.90; $\alpha T2_{\text{girls}}$ = 0.91; $\alpha T1_{\text{boys}}$ = 0.88; $\alpha T2_{\text{boys}}$ = 0.88) consist of 10 items each (e.g., "If I do not understand something, my mother/father explains it to me" or "My mother/father is studying with me").

Statistical Analyses

Initially, random parcels for each latent variable were built due to the large amount of single items per variable, which is a common procedure in psychological research (Nasser and Wisenbaker, 2003). Accordingly, each latent variable in the present study consists of three parcels. Little et al. (2002, 2013) list various reasons why parceling can be beneficial compared to using single items regarding psychometrics, model estimation, and fit characteristics. In contrast to item-level data, parcels show higher reliability, higher ratio of common-to-unique factor variance, greater communality, lower likelihood of distributional violations as well as more, tighter, and more-equal intervals, fewer parameter estimates, reduced sources of sampling error, lower indicator-to-sample size ratio as well as lower likelihood of correlated residuals and dual factor loadings.

Furthermore, measurement invariance over time as a precondition of cross-lagged panel design was tested for all variables used in this study. In the next step, three multigroup cross-lagged models were conceptualized: A less-restricted model (free parameters across girls and boys), a semi-restricted model (equal factor loadings, free thresholds, and free regression coefficients among girls and boys) and a more-restricted model (equal factor loadings, equal thresholds and equal regression coefficients across both groups). The less-restricted model was compared to the semi-restricted model to test for measurement invariance between the groups (i.e., girls and boys) by using χ^2-difference test (Yuan and Bentler, 2004). Subsequently, the semi-restricted model (with gender differences) was compared to a more-restricted model (considering no gender differences) in order to test which model would fit the data best.

All analyses were conducted with the "type is complex" command in Mplus to account for the nested structure of the data (students nested in classes) (Asparouhov, 2005). Four primary fit indices were used to determine model fit (Hu and Bentler, 1999): Chi-Square Test of Model Fit (χ^2), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI) and Standardized Root Mean Square Residuals (SRMR). Little's MCAR test ($\chi^2 = 268.07$; $df = 233$; $p > 0.05$) revealed that missing data was completely at random, which allows using full-information maximum likelihood (FIML) estimation.

RESULTS

Bivariate Correlations and Descriptive Statistics

Table 1 presents the gender-specific bivariate correlations and descriptive statistics calculated with IBM SPSS software (see **Table 1**).

Multigroup Cross-Lagged Panel Design

Initially, measurement invariance for the variables of interest was tested stepwise over time (see **Table 2**). As shown in **Table 2** strong factorial invariance is held for all latent variables, which is a precondition for cross-lagged panel design.

To test our hypotheses, three multi-group cross-lagged models (less-restricted model, semi-restricted model, more-restricted model) were conceptualized with Mplus (Muthén and Muthén, 1998-2013). All models included (a) autoregressive paths between the same variable at T1 and T2, (b) cross-lagged paths between a variable and each other variables over time, and (c) within-time covariances between all variables at T1 as well as within-time covariances between all variables at T2. To test for measurement invariance between the groups (i.e., girls and boys) a less-restricted model with all free parameters was conceptualized in a first step [$\chi^2_{(564)} = 1658.82$, $p < 0.001$; CFI $= 0.92$; RMSEA $= 0.06$ (0.06–0.06), SRMR $= 0.06$]. This model was compared with the semi-restricted model [$\chi^2_{(596)} = 1660.37$, $p < 0.001$; CFI $= 0.92$; RMSEA $= 0.06$ (0.05–0.06), SRMR $= 0.07$] with equal factor loadings, free thresholds, and free regression coefficients among boys and girls by using the χ^2-difference test [$\chi^2_{(32)} = 5.71$, $p > 0.05$] (Satorra and Bentler, 2001). The test indicated that the semi-restricted model was favored to the less-restricted model, which confirms measurement invariance between girls and boys. In a next step, a more restricted model [$\chi^2_{(641)} = 1731.80$, $p < 0.001$; CFI $= 0.92$; RMSEA $= 0.06$ (0.05–0.06), SRMR $= 0.08$] with equal factor loadings, equal thresholds and equal regression coefficients across both groups was conceptualized. Again, the χ^2-difference test [$\chi^2_{(45)} = 76.59$, $p < 0.05$] (Satorra and Bentler, 2001) between the semi-restricted and the more restricted model was conducted, implying that the semi-restricted model reflects the data better than the more restricted model. This means that different patterns for girls and boys are exhibited in the cross-lagged model.

Model Girls

Auto-Regressive Effects Over Time

All auto-regressive paths between each variable at T1 and T2 were found to be significant, which supports the stability of the constructs over time (see **Figure 1**).

Cross-Lagged Effects Over Time

Eight cross-lagged effects were found to be significant over time: Mother pressure at T1 negatively predicts the grades in mathematics at T2, whereas mother support at T1 negatively predicts mother pressure at T2 and positively predicts father support at T2, as well as girls' grades in mathematics and biology at T2. Father support at T1 positively predicts mother pressure at T2. Furthermore, the grades in mathematics at T1 positively predict the girls' grades in biology at T2 and vice versa (see **Figure 1**).

Covariances Within-Time

At T1 all covariances except for the association between mother support and father pressure were found to be significant: There was a positive association between mother support and father support ($r = 0.19$, $p < 0.001$), between mother support and girls' grades in mathematics ($r = 0.06$, $p < 0.01$) and biology ($r = 0.07$, $p < 0.001$). In turn, there was a negative association between mother pressure and mother support ($r = -0.08$, $p < 0.001$), father support and mother pressure ($r = -0.05$, $p < 0.05$) as well as between mother pressure and girls' grades in both mathematics ($r = -0.10$, $p < 0.01$) and biology ($r = -0.11$, $p < 0.001$). The association between mother pressure and father pressure ($r = 0.34$, $p < 0.001$) was positive. Furthermore, there was a positive relation between father support and both girls' grades in mathematics ($r = 0.07$, $p < 0.01$) and biology ($r = 0.07$, $p < 0.01$), whereas the association between father pressure and girls' grades in mathematics ($r = -0.10$, $p < 0.01$) and biology ($r = -0.08$, $p < 0.01$) were negative. Finally, the relation between father support and father pressure was positive ($r = 0.09$, $p < 0.01$) as well as the relation between girls' grades in mathematics and biology ($r = 0.31$, $p < 0.001$).

At T2, only eight covariances were found to be significant: The association between mother pressure and father pressure

TABLE 1 | Gender-specific means, standard deviations, and intercorrelations of the constructs.

	MP T2	FP T1	FP T2	MS T1	MS T2	FS T1	FS T2	Ma T1	Ma T2	Bio T1	Bio T2	*R*	*M*	*SD*
GIRLS														
MP T1	0.57**	0.51**	0.28**	−0.11**	−0.10*	−0.00	−0.07	−0.12**	−0.20**	−0.13**	−0.19**	1–4	2.50	0.66
MP T2	–	0.30**	0.49**	−0.08	−0.06	0.05	0.04	−0.06	−0.21**	−0.10*	−0.19**	1–4	2.35	0.68
FP T1		–	0.54**	0.06	−0.02	0.22**	0.13**	−0.14**	−0.18**	−0.10*	−0.09	1–4	2.18	0.66
FP T2			–	−0.05	−0.06	0.10*	0.18**	−0.15**	−0.28**	−0.13**	−0.26**	1–4	2.14	0.72
MS T1				–	60**	0.52**	0.42**	0.12**	0.21**	0.17**	0.19**	1–4	2.96	0.64
MS T2					–	0.36**	0.50**	0.12**	0.25**	0.11*	0.16**	1–4	2.85	0.68
FS T1						–	0.64**	0.13**	0.13**	0.16**	0.11*	1–4	2.91	0.73
FS T2							–	0.12**	0.14**	0.13**	0.11*	1–4	2.77	0.75
Ma T1								–	0.66**	0.53**	0.39**	1–6	4.30	0.79
Ma T2									–	0.47**	0.52**	1–6	4.13	0.87
Bio T1										–	0.57**	1–6	4.75	0.74
Bio T2											–	1–6	4.71	0.78
BOYS														
MPT1	0.56**	0.51**	0.40**	0.15**	0.13*	0.14**	0.17**	−0.14**	−0.19**	−0.15**	−0.14**	1–4	2.67	0.66
MP T2	–	0.33**	0.58**	0.07	0.16**	0.07	0.15**	−0.17**	−0.23**	−0.14**	−0.12*	1–4	2.52	0.68
FP T1		–	0.51**	0.15**	0.11*	0.37**	0.19**	−0.17**	−0.17**	−0.13**	−0.16**	1–4	2.41	0.69
FP T2			–	−0.00	0.14**	0.12*	0.29**	−0.25**	−0.23**	−0.21**	−0.17**	1–4	2.34	0.75
MS T1				–	53**	0.53**	0.27**	−0.08	−0.13*	0.09	0.11*	1–4	2.97	0.59
MS T2					–	0.29**	0.52**	−0.02	−0.03	0.09	0.09	1–4	2.75	0.61
FS T1						–	0.52**	0.01	−0.03	0.14**	0.07	1–4	2.95	0.67
FS T2							–	−0.02	0.01	0.09	0.02	1–4	2.76	0.67
Ma T1								–	0.65**	0.49**	0.35**	1–6	4.48	0.81
Ma T2									–	0.38	0.41**	1–6	4.11	0.88
Bio T1										–	0.62**	1–6	4.55	0.85
Bio T2											–	1–6	4.46	0.83

*Correlations are standardized coefficients; *p < 0.05, **p < 0.01; MP, Mother Pressure; FP, Father Pressure; MS, Mother Support; FS, Father Support; Ma, math grade; Bio, biology grade; Grades in German school system range from 1 = "very good" to 6 = "insufficient"—for sake of clarity, all grades were reverse-coded. T1, Time 1 (8th grade); T2, Time 2 (9th grade); R, Range; M, Mean.*

TABLE 2 | Model fit indices for measurement invariance testing over time and results of χ^2-difference test with scaling correction using MLR estimator and "type is complex" in Mplus.

Model	χ^2	df	*p*	RMSEA	90%CI	CFI	SRMR	$\Delta\chi^2$	*p*	Δdf
Model 0	1444.79	391	<0.001	0.05	0.05–0.05	0.93	0.06	–	–	–
Model 1	1453.98	401	<0.001	0.05	0.05–0.05	0.93	0.06	6.09	>0.05	10
Model 2	**1465.14**	**411**	**<0.001**	**0.05**	**0.05–0.05**	**0.93**	**0.06**	**13.59**	**>0.05**	**10**
Model 3	1482.93	417	<0.001	0.05	0.05–0.05	0.93	0.06	15.38	<0.05	6

Model 0, no constraints but configural invariance; Model 1, loadings invariant across time (weak invariance); Model 2, loadings and intercepts invariant across time (strong invariance); Model 3, measurement model including time invariance restriction (strict invariance). Bold values indicate the best model fit.

($r = 0.20$, $p < 0.001$) was positive, whereas the relation between mother pressure and girls' grades in mathematics ($r = -0.04$, $p < 0.05$) was negative. The association between father support and father pressure ($r = 0.05$, $p < 0.05$) and between father support and mother support ($r = 0.08$, $p < 0.001$) were positive. Furthermore, both the relation between girls' grades in mathematics and father pressure ($r = -0.09$, $p < 0.001$) were negative as well as the association between girls' grades in biology and father pressure ($r = -0.10$, $p < 0.001$). In turn, the association between girls' grades in mathematics and mother support ($r = 0.04$, $p < 0.01$) was positive. Finally, the relation between girls' grades in mathematics and biology ($r = 0.11$, $p < 0.001$) was still positive.

The association between mother support and mother pressure, between father support and mother pressure, between girls' grades in mathematics and father support, between girls' grades in biology and mother pressure, between girls' grades in biology and mother support as well as between girls' grades in biology and father support were no longer significant. Finally, the relation between mother support and father pressure was still not significant.

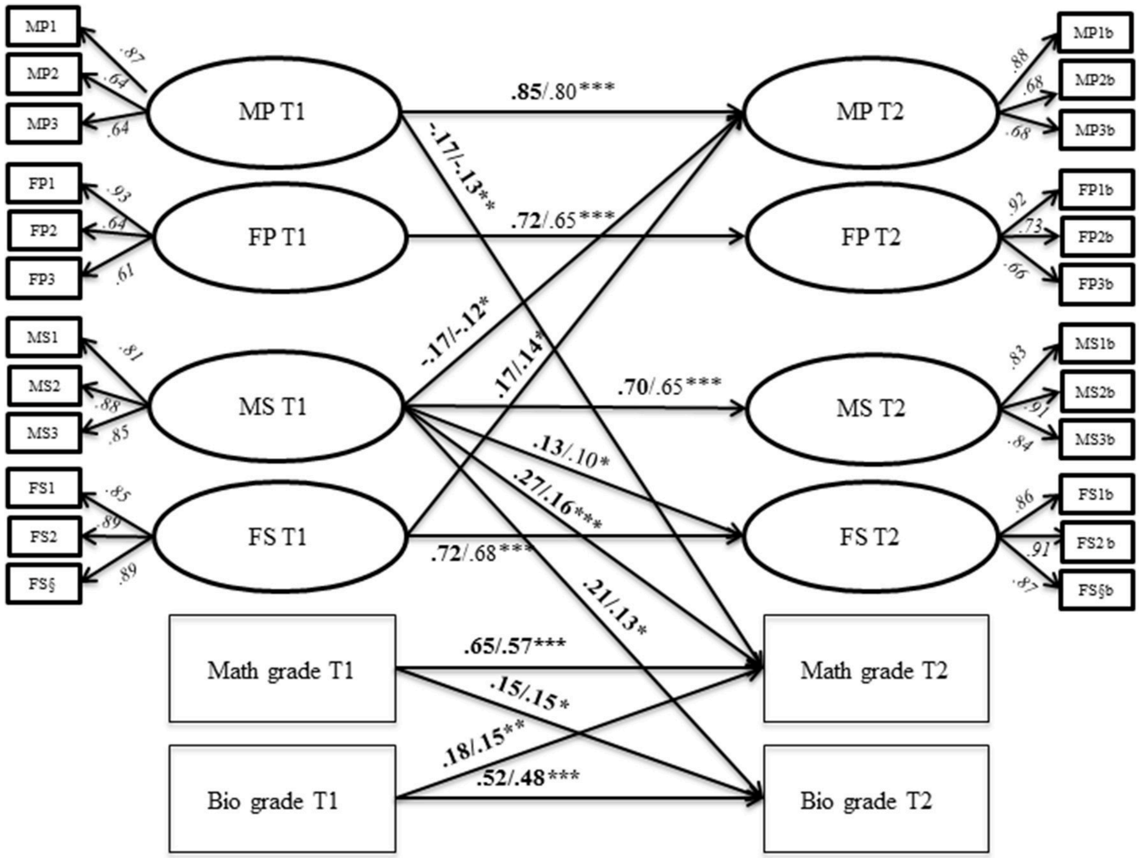


FIGURE 1 | Multigroup cross-lagged model for girls. MP, Mother Pressure; FP, Father Pressure; MS, Mother Support; FS, Father Support; T1, Time 1 (8th grade); T2, Time 2 (9th grade); Factor loadings are shown as standardized coefficients. For sake of clarity, solely significant autoregressive and cross-lagged paths are shown in the figure: First position indicates unstandardized coefficients (B), second position standardized coefficients (β). In order to obtain a clearly arranged figure, the covariances between all variables at T1 and between all variables at T2 are not shown in the figure, but reported in the manuscript.

Model Boys

Auto-Regressive Effects Over Time

All auto-regressive paths between each variable at T1 and T2 were found to be significant, which support the stability of the constructs over time (**Figure 2**).

Cross-Lagged Effects Over Time

Three cross-lagged effects were found to be significant over time: Mother pressure at T1 positively predicts father pressure at T2, whereas mother support negatively predicts boys' grades in mathematics. Finally, the grades in biology at T1 positively predict boys' grades in mathematics at T2, but not vice versa (**Figure 2**).

Covariances Within-Time

At T1 all covariances except for four associations (between mother support and both boys' grades in mathematics and biology as well as between father support and both boys' grades in mathematics and biology) were found to be significant: There was a positive association between mother support and father support ($r = 0.16$, $p < 0.001$), between mother support and father pressure ($r = 0.05$, $p < 0.05$)—which was not significant for girls—and between mother pressure and father support ($r = 0.06$, $p < 0.05$) as well as between mother pressure and mother support ($r = 0.04$, $p < 0.05$), whereas the latter association was negative for girls. Furthermore, the relation between mother pressure and both boys' grades in both mathematics ($r = -0.15$, $p < 0.001$) and biology ($r = -0.11$, $p < 0.001$) were negative. The association between mother pressure and father pressure ($r = 0.36$, $p < 0.001$) was positively. Furthermore, the association between father pressure and boys' grades in mathematics ($r = -0.11$, $p < 0.001$) and biology ($r = -0.11$, $p < 0.001$) were negatively. Finally, the relation between father support and father pressure was positive ($r = 0.17$, $p < 0.001$) as well as the relation between boys' grades in mathematics and biology ($r = 0.34$, $p < 0.001$).

At T2, eight covariances were found to be significant: In contrast to the girls, the association between mother support and mother pressure ($r = 0.04$, $p < 0.05$) as well as between mother support and father pressure ($r = 0.07$, $p < 0.001$) was still positively significant. Furthermore, the relation between mother pressure and father pressure ($r = 0.22$, $p < 0.001$), as well as the relation between father support and father pressure ($r = 0.12$, $p < 0.001$) and between father support and mother support ($r = 0.12$,

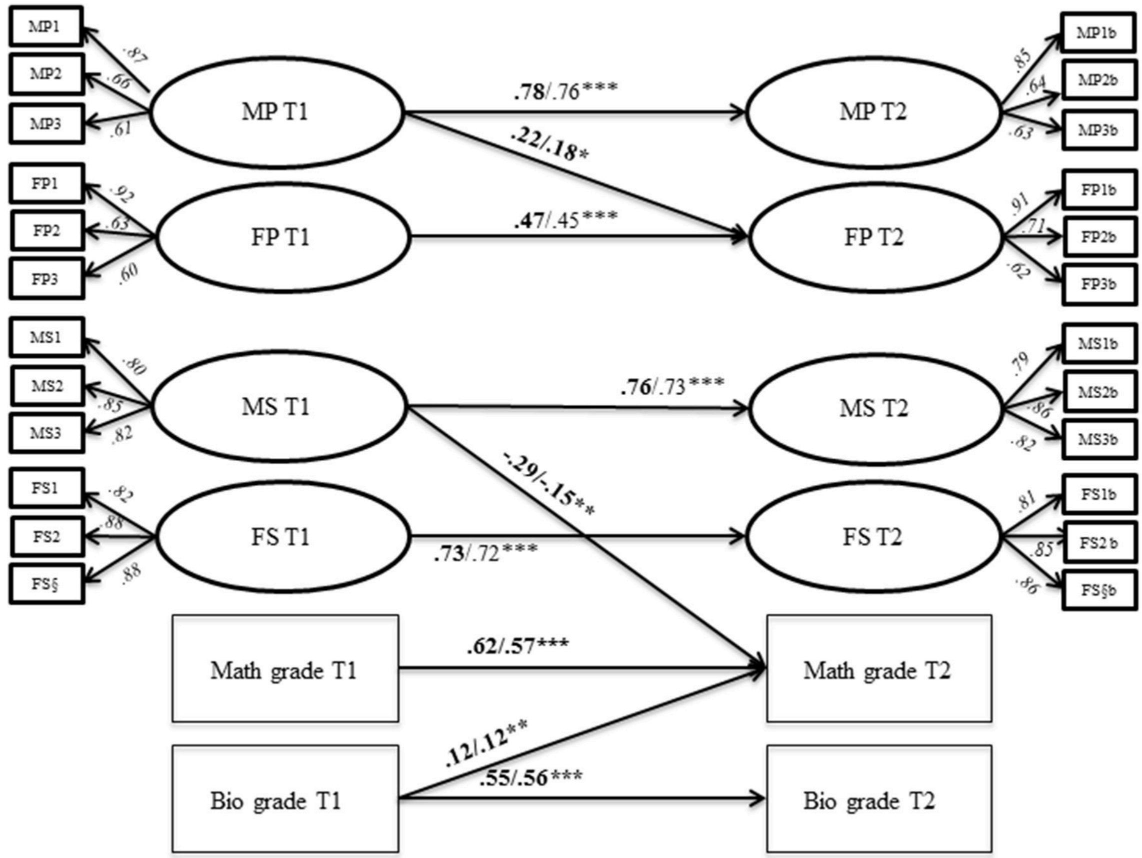


FIGURE 2 | Multigroup cross-lagged model for boys. MP, Mother Pressure; FP, Father Pressure; MS, Mother Support; FS, Father Support; T1, Time 1 (8th grade); T2, Time 2 (9th grade); Factor loadings are shown as standardized coefficients. For sake of clarity, solely significant autoregressive and cross-lagged paths are shown in the figure: First position represents unstandardized coefficients (B), second position standardized coefficients (β). In order to obtain a clearly arranged figure, the covariances between all variables at T1 and between all variables at T2 are not shown in the figure, but reported in the manuscript.

$p < 0.001$) was positively significant. In turn, the relation between boys' grades in mathematics and mother pressure was negative ($r = -0.06$, $p < 0.05$). Finally, boys' grades in mathematics and biology were positively associated ($r = 0.10$, $p < 0.01$).

The association between father support and mother pressure, between boys' grades in mathematics and father support, boys' grades in mathematics and pressure, boys' grades in mathematics and mother support as well as between boys' grades in biology and all parental variables were not significant.

DISCUSSION

This study aimed to shed light on potential gender-specific differences in the within- and over time associations between adolescents' perception of parental pressure/support and their grades in mathematics and biology.

In line with hypothesis I we found that girls and boys differ regarding their perceived parental pressure/support associated with their grades in mathematics and biology at the beginning of 8th grade and 1.5 years later at the end of 9th grade (within-time associations). While for girls, mother and father support in 8th grade were associated with better grades in math and biology, for boys, neither mother nor father support were significantly related to their STEM performance. Hence, boys do not profit from mother/father support regarding their academic performance in math and biology but may be interested in STEM independent of parental behavior (Su and Rounds, 2015). Also, it can be assumed that boys are impacted by their peer group, rather than their parents. Various studies show that students' motivation and involvement in school are influenced by their peers (Raufelder et al., 2013; van Hoorn et al., 2014). Robnett and Leaper (2013) found that students were more likely to be interested in pursuing a STEM career if their peer group valued STEM, even after controlling for individual grades, values, and expectations.

In contrast to boys, girls' performance in math and biology was related to parental support. Hence, the results indicate that (1) parental behavior is perceived differently by girls and boys and/or (2) parents act differently toward their male vs. female offspring. In support of these arguments, Fthenakis and Minsel (2002) found that fathers' of daughters spend more time with their offspring compared to fathers' of sons, which could explain why father support is related to better STEM performance among girls, but not among boys. Additionally, Paa and Hawley

McWhirter (2011) found that particularly girls received more positive feedback and autonomy from their mother.

Compared to the role of mothers, the impact of fathers on their offspring seems much harder to grasp. In past research the father figure was described to be limited to financial support, inadequate, or absent fathering (Hawkins and Dollahite, 1997; Marks and Palkovitz, 2004). The father figure implies rather negative connotations, or is insufficiently conceptualized, which becomes clear in the work "Fathers: the "invisible" parents" (Saracho and Spodek, 2008). However, while the gender gap between mothers and fathers is quite persistent over time regarding housework and child rearing (Kan et al., 2011), more recent work characterizes the father as the "working caring-dad" whose role comprises more than a bread-winning function as he is ready to sacrifice—at least in part and among older fathers—his career for children (Fthenakis and Minsel, 2002). In the current study, the father comes into play, when boys and girls perceive his pressure, which is related to lower STEM performance in math and biology for girls in 8 and 9th grade and for boys in grade 8. The challenge of characterizing the role of fathers is indicated by the finding that both girls and boys report the more father support they receive also the more pressure they perceive, which is consistent across time. Hence, father support cannot be described as solely positive while father pressure cannot be described as solely negative for boys' and girls' STEM performance. In fact, although parental support and pressure are separate concepts, children perceive them as overlapping as part of general parenting behavior. While pressure is described as behavior indicating expectations that are high, unlikely or even impossible to attain, this nominal definition depends on the child's perception of parental pressure and support (Leff and Hoyle, 1995). Besides the positive association between father support and pressure, mother support and pressure was also positively associated for boys in both grades 8 and 9. Contrary, girls who receive mother support, receive less mother pressure in 8th grade. Interestingly, boys who receive mother support also receive less father pressure, while for girls there was no such significant association. In the case of boys, the mother might compensate for the father pressure by supporting her son. It was found that fathers feel more disturbed with respect to conflicts and trouble by their sons, compared to daughters (Fthenakis and Minsel, 2002). While, a diary study indicates that the relationship with mothers is related to lower psychological distress for both boys and girls (Mallers et al., 2010). In line with this finding, according to a self-report study, on average, mothers evaluate themselves as more supportive toward their children compared to fathers (Fthenakis and Minsel, 2002), while fathers show more control and punishment regarding schoolwork toward their sons compared to daughters. In general, the findings show that for both boys and girls, the association between parental support /pressure and STEM performance is higher in grade 8 compared to grade 9. This finding indicates that parental pressure and support has more impact at the age of about 13 compared to the age of about 15, as parents become gradually less important across the development of youth (Erikson, 1993; Eccles, 2007) while peers and their values, activities, and attitudes become more important (Leff and Hoyle, 1995; van Hoorn et al., 2014).

Investigating over time associations between mother/father support/pressure, math and biology performance, all auto-regressions were significant from grade 8 to grade 9, indicating the reliability of the constructs. Furthermore, hypothesis II was partially confirmed, as girls and boys varied in their perception of mother/father support/pressure related to their STEM performance in math and biology over time, but not vice versa. In other words, the relation between parental support/pressure and student's STEM performance seems to be rather mono-directional, such as parental support/pressure predicts STEM performance, but not vice versa. Future longitudinal studies with more than two waves are necessary to test a potential causal ordering of the variables.

For girls, mother support in grade 8 was related to better math and biology performance in grade 9, while mother pressure in grade 8 was related to lower math performance in grade 9, but not to biology performance. Hence, girls' biology performance seems independent of mother pressure, indicating that girls may have stable interests in biology, which is not *per se* a masculine STEM subject, and their peer group may value or engage in biology at school. In fact, Leaper et al. (2011) found that adolescent girls' motivation in STEM courses was positively related with peer support over the school years. These interpretations may give an explanation why females and males are equally engaged in biology and agricultural studies.

For boys, only mother support in grade 8 was related to low math performance in grade 9. Hence, while girls perceive mother support as helpful for better STEM performance, boys perceive mother support as debilitating their performance in math. This finding indicates that boys may receive mother support as pressure related to high expectations, which in turn hinder boys' STEM performance.

Contrary to the impact of mothers, father support or pressure was not related to girls' and boys' STEM performance over time. These findings underline the impact of mothers for students' performance in STEM, which have been investigated in various studies (Paa and Hawley McWhirter, 2011; Gunderson et al., 2012; Rozek et al., 2015). In fact, compared to fathers, it is mostly the mother who is involved in and concerned with school and family issues (Winquist Nord and West, 2001), which might be the reason why mothers' school related behavior impacts both boys' and girls' STEM performance.

As boys' STEM performance is barely impacted by parental behavior, their academic performance might be related to other factors outside the family, such as interest (Su and Rounds, 2015), or boys might just live up to the stereotype that STEM subjects are masculine, confirming their interest even more (Archer et al., 2013). In fact, compared to boys, it is much harder to involve and maintain the interest of girls in STEM. In this sense, the current study indicates that girls profit from mother support and in contrast suffer from mother pressure regarding their STEM performance. Hence, mothers should be aware of their school-related behavior, particularly exerting pressure as it inhibits girl's STEM performance as well as giving support to their male offspring, which in fact is perceived as pressure. Furthermore, the results of the study show that boys' and girls' STEM performance in grade 8 does not impact the school related behavior of mothers and fathers 1.5 years later. Hence, it is the differentiated parental behavior that impacts STEM performance and not vice versa, i.e., STEM performance in grade 8 does not impact mothers' and fathers' support/pressure in 9th grade.

Overall, the results adhere to the General Expectancy-Value Model of Achievement Choices (Eccles, 2014), as results indicate that father and mother behavior impact boys' and girls' performance in STEM. However, this model does not take into account the specific and different role of mothers and fathers regarding the STEM performance of boys and girls. To further analyze the impact of maternal and paternal behavior, future studies should include information on parental time spent with children, gender attitudes or distribution of domestic tasks and child care. In sum, this study indicates that mother support plays an essential but different role for boys' and girls' STEM performance. While mother support should further be encouraged for girls, the support directed toward boys should be reconsidered as mother support seems to be accompanied by expectations or pressure. Furthermore, mother's pressure inhibits STEM performance among girls and should therefore be reduced or eliminated in the school-related behavior of mothers. The results also show that the father does not have a long lasting effect (from grade 8 to grade 9) on neither boys' nor girls' STEM performance. However, cross-sectionally father's pressure is related to low STEM performance for both boys and girls. Therefore, similarly to mother's pressure, the father should also be urged to eliminate his pressure toward his offspring regarding school performance. Additionally, fathers should be encouraged to support their sons and daughters in school related issues and engage in school work, conversations about science, and in whatever concerns their offspring in order to build confidence in the father-child relationship and support boys' and girls' development positively. Parent-child interventions may help parents to reflect and find their role regarding their scholastic behavior as well as give parents the opportunity to communicate with their children about the child's needs, fears, and hopes regarding their academic performance in STEM. Furthermore, gender-specific parent-child activities in school could be added to existing models, such as the American model of family–school partnerships (Epstein, 1995), which helps parents to recognize the value of their contributions to schooling practices and foster students' academic involvement (Nawrotzki, 2012).

Strength, Limitations, and Future Directions

Strength: Girls and boys perception of pressure and support from both mothers and fathers was considered separately. Furthermore, results are based on data from a large sample with two waves (beginning of 8th grade and end of 9th grade) grasping a longer period during adolescence.

Limitations: Self-reported grades are rather a weak indicator of students' achievement in STEM. However, we used self-report measures as we were particularly interested in students' perception of their parents' support and pressure. Future studies are warranted, which use more detailed instruments of girls' and boys' STEM achievement and additionally consider associated concepts, such as subject-related interest and motivation.

Future Directions: Results suggest that the relationship between adolescents' STEM achievement and parental pressure/support is rather mono-directional than bi-directional over time. Future longitudinal studies with several measurement points might identify a potential causal order of the variables. In addition, variables such as parental time spent with children, gender attitudes or distribution of domestic tasks and child care should be included as moderators in future studies.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

FH wrote the main part of the paper. DR did the statistical analyses and wrote the method and results section.

ACKNOWLEDGMENTS

The research reported in this paper was supported by a grant from The Volkswagen Foundation (Schumpeter Fellowship II/84 452). The authors would like to thank the principals, teachers, and students for their cooperation in making these studies possible.

REFERENCES

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., and Wong, B. (2013).'Not girly, not sexy, not glamorous': primary school girls'and parents'constructions of science aspirations. *Pedag. Cult. Soc.* 21, 171–194. doi: 10.1080/14681366.2012.748676

Aschbacher, P., Li, E., and Roth, E. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *J. Res. Sci. Teach.* 47, 564–582. doi: 10.1002/tea.20353

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Struct. Equat. Model.* 12, 37–41. doi: 10.1073/pnas. 0910967107

Buschor, C. B., Berweger, S., Keck Frei, A., and Kappler, C. (2014). Majoring in STEM – What accounts for women's career decision making? A mixed methods study. *J. Educ. Res.* 107, 167–176. doi: 10.1080/00220671.2013.788989

Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412

Dabney, K. P., Chakraverty, D., and Tai, R. H. (2013). The association of family influence and initial interest in science. *Sci. Educ.* 97, 395–409, doi: 10.1002/sce.21060

Dohn, H. (1991). Dropout in the Danish high school (gymnasium). An investigation of psychological, sociological and pedagogical factors. *Int. Rev. Educ.* 37, 415–428. doi: 10.1007/BF00597619

Eccles, J. S. (2007). "Families, schools, and developing achievement-related motivations and engagement," in *Handbook of Socialization: Theory and Research,* eds J. E. Grusec and P. D. Hastings (New York, NY: Guilford Press), 665–691.

Eccles, J. S. (2014). Gendered socialization of STEM interests in the family. *Int. J. Gender Sci. Technol.* 7, 116–132.

Epstein, J. L. (1995). School/family/community partnerships: caring for the children we share. *Phi Delta Kappan* 76, 702–712. doi: 10.1177/003172171009200326

Erikson, E. H. (1993). *Identität und Lebenszyklus: drei Aufsätze [Identity and Lifecycle: Three Essays]*. Translated by K. Hügel, Vol. 1. Frankfurt Main: Suhrkamp.

Fthenakis, W. E., and Minsel, B. (2002). *Die Rolle des Vaters in der Familie. [The Role of Fathers Within the Family]. Band 213. Schriftenreihe des Bundesministeriums für Familie, Senioren, Frauen und Jugend*. Stuttgart; Berlin; Köln: Kohlhammer.

Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles* 66, 153–166. doi: 10.1007/s11199-011-9996-2.144

Halim, L., Rahmann, N. A., Zamri, R., and Mohtar, L. (2018). The roles of parents in cultivating children's interest towards science learning and careers. *Kasetsart J. Soc. Sci.* 39, 190–116. doi: 10.1016/j.kjss.2017.05.001

Hawkins, A. J., and Dollahite, D. C. (1997). *Generative Fathering: Beyond the Deficit Perspective*. Thousand Oaks, CA: Sage.

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: conventional criteria versus new alternatives. *Struct. Equat. Model.* 6, 51–55. doi: 10.1080/10705519909540118

Kan, M. Y., Sullivan, O., and Gershuny, J. (2011). Gender convergence in domestic work: discerning the effects of interactional and institutional barriers from large-scale data. *Sociology* 45, 234–225.doi: 10.1177/0038038510394014

Kerney, A. T., and YourLife Campaign (2016). *Tough Choices: The Real Reasons A-level Students are Steering Clear of Science and Maths*. Available online at: https://www.atkearney.com/documents/10192/7390617/Tough+Choices.pdf

Leaper, C., Farkas, T., and Brown, C. S. (2011). Adolescent girls' experiences and gender-related beliefs in relation to their motivation in math/science and English. *J. Youth Adolesc.* 41, 268–282. doi: 10.1007/s10964-011-9693-z

Leff, S. S., and Hoyle, R. H. (1995). Young athletes' perceptions of parental support and pressure. *J. Youth Adolesc.* 24, 187–203. doi: 10.1007/BF015 37149

Levpušček, M. P., and Zupančič, M. (2009). Math achievement in early adolescence: the role of parental involvement, teacher's behavior, and students' motivational beliefs about math. *J. Early Adolesc.* 29, 541–570. doi: 10.1177/0272431608324189

Little, T. D., Cunningham, W. A., Shahar, G., and Widaman, K. F. (2002). To parcel or not to parcel: exploring the question and weighing the merits. *Struct. Equat. Model.* 9, 151–173. doi: 10.1207/S15328007SEM0902_1.

Little, T. D., Rhemtulla, M., Gibson, K., and Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychol. Methods* 18, 285–300. doi: 10.1037/a0033266

Lubinski, D., Benbow, C. P., Shea, D. L., Eftekhari-Sanjani, H., and Halvorson, M. B. (2001). Men and women at promise for scientific excellence: similarity not dissimilarity. *Psychol. Sci.* 12, 309–317. doi: 10.1111/1467-9280.00357

Lytton, H., and Romney, D. M. (1991). Parents' differential socialization of boys and girls: a meta-analysis. *Psychol. Bull.* 109, 267–296. doi: 10.1037/0033-2909.109.2.267

Mallers, M. H., Charles, S. T., Neupert, S. D., and Almeida, D. M. (2010). Perceptions of childhood relationships with mother and father: daily emotional and stressor experiences in adulthood. *Dev. Psychol.* 46, 1651–1661.doi: 10.1037/a0021020

Mann, A., and DiPrete, T. A. (2013). Trends in gender segregation in the choice of science and engineering majors. *Soc. Sci. Res.* 42, 1519–1541. doi: 10.1016/j.ssresearch.2013.07.002

Marks, L., and Palkovitz, R. (2004). American fatherhood types: the good, the bad, and the uninterested. *Fathering* 2, 113–129. doi: 10.3149/fth.0202.113

Mullis, I. V. S., Martin, M. O., Foy, P., and Hooper, M. (2016). *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. Boston College, TIMSS & PIRLS International Study Center. Available online at: http://timssandpirls.bc.edu/timss2015/international-results/advanced/

Muthén, L. K., and Muthén, B. O. (1998-2013). *Mplus User's Guide, 7th Edn*. Los Angeles, CA: Muthén & Muthén.

Nasser, F., and Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educ. Psychol. Measure.* 63, 729–757. doi: 10.1177/0013164403258228

National Science Board (2014). *Science and Engineering Indicators 2014*. Arlington VA: National Science Foundation (NSB 14–01).

Nawrotzki, K. D. (2012). "Parent–school relations in England and the USA: Partnership, problematized," in *The Politicization of Parenthood: Shifting Private and Public Responsibilities in Education and Child Rearing*, eds M. Richter & S. Andresen (Dordrecht: Springer), 69–83.

Paa, H. K., and Hawley McWhirter, E. (2011). Perceived influences on high school students' current career expectations. *Career Dev. Q.* 49, 29–44. doi: 10.1002/j.2161-0045.2000.tb00749.x

Raufelder, D., Jagenow, D., Drury, K., and Hoferichter, F. (2013). Social Relationships and Motivation in Secondary School: 4 different motivation types. *Learn. Individ. Diff.* 24, 89–95. doi: 10.1016/j.lindif.2012.12.002

Reinking, A., and Martin, B. (2018). The gender gap in STEM fields: theory, movements, and ideas to engage girls in STEM. *J. N. Approaches Educ. Res.* 7, 148–153. doi: 10.7821/naer.2018.7.271

Reitzle, M., Metzke, C. W., and Steinhausen, H.-C. (2001). Eltern und kinder: der Zürcher kurzfragebogen zum erziehungsverhalten (ZKE). [Parents and children: the Zürich questionnaire of educational behavior (ZKE)]. *Diagnostica* 47, 196–207. doi: 10. 1026//0012-1924.47.4.196.

Robnett, R. D., and Leaper, C. (2013). Friendship groups, personal motivation, and gender in relation to high school students' STEM career interest. *J. Res. Adolesc.* 23, 652–664. doi: 10.1111/jora.12013

Rozek, C. S., Hyde, J. S., Svoboda, R. C., Hulleman, C. S., and Harackiewicz, J. M. (2015). Gender differences in the effects of a utility-value intervention to help parents motivate adolescents in mathematics and science. *J. Educ. Psychol.* 107, 195–206. doi: 10.1037/a0036981

Saracho, O. L., and Spodek, B. (2008). Fathers: the 'invisible' parents. *Early Child Dev. Care* 178, 821–836, doi: 10.1080/03004430802352244

Satorra, A., and Bentler, P. M. (2001). A scaled difference Chi square test statistic for moment structure analysis. *Psychometrika* 66, 507–514. doi: 10.1007/BF02296192.

Simpkins, S. D., Price, C. D., and Garcia, K. (2015). Parental support and high school students' motivation in biology, chemistry, and physics: understanding differences among Latino and Caucasian boys and girls. *J. Res. Sci. Teach.* 52, 1386–1407. doi: 10.1002/tea.21246

Su, R., and Rounds, J. (2015). All STEM fields are not created equal: people and things interests explain gender disparities across STEM fields. *Front. Psychol.* 6:189. doi: 10.3389/fpsyg.2015.00189

United Nations Educational Scientific and Cultural Organization (2017). *Cracking the Code: Girls' and Women's Education in Science, Technology, Engineering and Mathematics (STEM). Education 2030.* Paris. Available online at: http://unesdoc.unesco.org/images/0025/002534/253479E.pdf

Valla, J. M., and Ceci, S. J. (2014). Breadth-based models of women's underrepresentation in STEM fields: an integrative commentary on Schmidt (2011) and Nye et al. (2012). *Perspect. Psychol. Sci.* 9, 219–224. doi: 10.1177/1745691614522067

van Hoorn, J., van Dijk, E., Meuwese, R., Rieffe, C., and Crone, E. A. (2014). Peer influence on prosocial behavior in adolescence. *J. Res. Adolesc.* 26, 90–100.doi: 10.1111/jora.12173

Wijsman, L. A., Warrens, M. J., Saab, N., van Driel, J. H., and Westenberg, P. M. (2016). Declining trends in student performance in lower secondary education. *Eur. J. Psychol. Educ.* 31, 595–612. doi: 10.1007/s10212-015-0277-2

Winquist Nord, C., and West,. J. (2001). *Fathers' and Mothers' Involvement in Their Children's Schools by Family Type and Resident Status, NCES 2001–032*. Washington, DC: U.S. Department of Education. National Center for Education Statistics. Available online at: https://nces.ed.gov/pubs2001/2001032.pdf

Yuan, K.-H., and Bentler, P. M. (2004). On Chi square difference and z-tests in mean and covariance structure analysis when the base model is misspecified. *Educ. Psychol. Measure.* 64, 737–757. doi: 10.1177/0013164404264853

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 26 March 2019
doi: 10.3389/fpsyg.2019.00603

Self-Concept and Achievement in Math Among Australian Primary Students: Gender and Culture Issues

*Feifei Han**

Griffith University, Mt Gravatt, QLD, Australia

While gender stereotype on math learning and achievement is consistently reported among existing research, these studies predominantly focus on mainstream students with Western cultural backgrounds. There is a dearth of study, which investigates gender effect among Australian Indigenous students. To fill this gap, the present study adopted a multiple-indicator-multiple-indicator-cause approach to structural equation modeling to investigate effects of gender, culture (Indigenous vs. non-Indigenous), and the interaction of the two on students' self-concept of competence and affect in math, as well as math achievement among Australian primary school students. We found gender stereotype effect not only on students' self-perceptions of their competence in math but also their actual math performance reflected in their math achievement scores in a standard math test. Boys had higher ratings on math competence and scored more highly on math test than girls. However, the gender stereotype was not found for self-concept of affect. Instead, culture was significantly impacted on self-concept of math affect, indicating that Indigenous students had less enjoyment toward learning math compared with their non-Indigenous peers. Furthermore, significant interaction effects between gender and culture were observed on both self-concept of math competence and math affect. In practice, to enhance Indigenous students' interest and enjoyment in math learning, educators are suggested to incorporate Indigenous students' values, beliefs, and traditions when delivering new math knowledge.

Keywords: gender, Australian Indigenous culture, self-concept of competence and affect, math achievement, primary school students

OPEN ACCESS

Edited by:
Silke Luttenberger,
University of Teacher Education Styria, Austria

Reviewed by:
Josip Burusic,
Institut Društvenih Znanosti Ivo Pilar, Croatia
Lu Wang,
Ball State University, United States

Correspondence:
Feifei Han
feifei.han@griffith.edu.au

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: 15 December 2018
Accepted: 04 March 2019
Published: 26 March 2019

Citation:
Han F (2019) Self-Concept and Achievement in Math Among Australian Primary Students: Gender and Culture Issues. Front. Psychol. 10:603. doi: 10.3389/fpsyg.2019.00603

INTRODUCTION

Indigenous Australians are the first peoples of Australia (Craven et al., 2013) and are one of the most disadvantaged Indigenous populations in the world (Cooke et al., 2007). All Australian governments in the last two decades have acknowledged that Indigenous Australians are disadvantaged in a number of socioeconomic indicators, including education (e.g., Commonwealth of Australia, 2006). Indigenous students participate significantly less in education and have significantly higher attrition rates compared to other Australian populations. For instance, the retention rate for Indigenous students from year 7 (the first year of secondary school) to year 12 was 55%, whereas the retention rate for non-Indigenous students during this same time period was 82% (Australian Bureau of Statistics, 2012). Craven et al. (2013) have called for paying a special attention to the disparities in educational outcomes between Indigenous students and their peers and taking immediate actions to make Indigenous Australians' full potential flourish. Thus, the first aim of the present study is to gain a thorough understanding on how Indigenous Australian students differ from non-Indigenous Australian students on self-concept and achievement in math subject in order to inform the development of effective intervention programs to help close the educational gap for Indigenous students.

Moreover, past research has revealed that there may be a gender stereotype in academic self-concept, and achievement in math (Meece et al., 2006; Yeung et al., 2012a), in particular, gender differences tend to become observable from early primary school (Eccles et al., 1993; Usher and Pajares, 2008). However, existing research regarding students' self-concept in math predominantly focused only on the perceptions of one's competence, neglecting the perceptions of one's affect in math (Midgley et al., 2001; Yeung and Han, 2017). Furthermore, whether the commonly found gender stereotype in self-concept and achievement in math is also observable in Australian indigenous student population is unknown. As Australian Indigenous culture, values, and perspectives differ from the dominant Western culture, we cannot legitimately apply the gender stereotype in math self-concept and achievement from the mainstream Western culture to Australian Indigenous students. Therefore, the present study examines both the gender and culture effects and the interaction of the two, on self-concept of competence and affect, as well as achievement in math subject among Australian primary school students from Indigenous and non-Indigenous backgrounds.

LITERATURE REVIEW AND RESEARCH QUESTIONS

Self-Concept and Its Impact

Positive self-beliefs are at the heart of the positive psychology movement (Marsh and Craven, 2006) and enhancing self-concept is enshrined in educational policies internationally. For example, the *Melbourne Declaration on Educational Goals for Young Australians* (MCEETYA, 2008) emphasizes that students should "have a sense of self-worth, self-awareness, and personal identity that enables them to manage their emotional, mental, spiritual, and physical wellbeing" (p. 9). Positive self-concept has been demonstrated "to impact on a wide range of critical wellbeing outcomes and serve as an influential platform for enabling full human potential" (Craven and Marsh, 2008, p. 104). Interventions specifically addressing domain-specific self-concept have been shown to result in domain-specific gains in a range of achievement outcomes (Craven and Yeung, 2008). Numerous studies have identified strong relations between self-concept and outcomes such as well-being, coursework selection, rate of school completion, adaptive academic behaviors, coping mechanisms, enhanced academic achievement, and reduced mental health problems (e.g., Marsh and Craven, 2006; Craven and Yeung, 2008). Self-concept and achievement are also known to be reciprocally related whereby they share a dynamic causal relation (Marsh and Craven, 2006). In the school context, academic self-concepts in different school subjects have been consistently demonstrated that they are not only causes for cognitive outcomes; but are also triggers of desirable psychological outcomes (Marsh and Craven, 2006; Craven and Yeung, 2008).

The structure of self-concept has been empirically demonstrated as multidimensional and domain specific (Arens et al., 2011). Marsh (1990), for example, found distinct self-concepts in a number of school subjects, including verbal, math, physical, art, music, and religion, with a general academic self-concept as an overarching construct. Traditionally, researchers either conflated the competence and affect aspects of academic concept (e.g., Jansen et al., 2014) or they have placed more emphasis on self-concept of competence over affect. Hence, academic self-concept has been consistently measured by either combining the competence and affect aspects or predominantly using competence aspect alone (Pinxten et al., 2013; Seaton et al., 2014).

However, in recent years, academic self-concept it has been empirically demonstrated that self-concept of competence (in relation to cognition) and self-concept of affect (in relation to emotion) are clearly distinguishable (e.g., Arens et al., 2011; Pinxten et al., 2013). While the competence component is concerned with the extent to which students perceive themselves to have capabilities in a specific school subject (e.g., I am good at math.), the affect component is about the extent to which an individual enjoys participating in a subject (e.g., I like math.). Thus, we will examine both the competence and affect aspects of self-concept in math.

Gender and Culture Issues in Self-Concept and Achievement

Gender differences may be observable as early as in elementary school when self-beliefs and perceptions begin to form (Eccles et al., 1993). Past self-concept research focusing only on competence has indicated that in general boys tend to have higher competence beliefs than girls (Midgley et al., 2001). Boys sometimes overestimate their competence whereas girls tend to underestimate theirs (Metallidou and Vlachou, 2007), even though such self-perceptions of abilities may not match their real ability (Yeung et al., 2012a,b). However, boys and girls may exhibit very different levels of competence beliefs in different curriculum areas. Research has consistently reported that boys tend to have higher perceptions of competence in math and science-related subjects (Marsh and Yeung, 1998; Klapp Lekholm and Cliffordson, 2009), whereas girls show higher self-concepts in language and verbal-related subjects (Marsh, 1993; Kurtz-Costes et al., 2008).

Gender has also been found to interact with other factors, including students' ability and cultural backgrounds (Dai, 2001; Chiu and Klassen, 2010). For instance, Dai (2001) observed that for average-ability students, girls reported a higher verbal self-concept and lower math self-concept than boys. However, for gifted students, girls were found to have a comparable math self-concept to boys. Chiu and Klassen (2010) found that for students from a culture which tolerates more uncertainty, math self-concept had a stronger relation with math achievement for boys than for girls. Extending a focus on self-concept of competence to affect beliefs, Yeung et al. (2012a,b) found an interaction effect between culture and gender—the competence and affect differences between Asian and Anglo Australian students were more pronounced for boys than for girls.

With regard to academic achievement, interaction effects between culture and gender have also been demonstrated. Lai (2010), for instance, showed that Chinese girls performed better than boys in both primary and middle schools. But for American students, this pattern was not consistent. American girls achieved better than boys in elementary school, but boys gradually caught up in math and science in middle schools.

The Current Study and Research Questions

Although the above studies demonstrated that gender effects on academic self-concept and achievement may be partly influenced by culture, there is a lack of research on how gender effect may be interacted with Australian Indigenous culture. In this investigation, we will test gender, Indigenous culture, and the interaction of the two on the self-concept of math competence and affect, as well as math achievement among Australian primary school students.

Three research questions were addressed in the study:

1. Do boys and girls differ in the self-concept of math competence, affect, and math achievement?
2. Do Indigenous and non-Indigenous students differ in the self-concept of math competence, affect, and math achievement?
3. Are there interaction effects between gender and Indigenous culture on the self-concept of math competence, affect, and math achievement?

MATERIALS AND METHODS

Participants

The study was conducted with 566 boys (44.6%) and 702 girls (55.4%), who studied in urban and rural areas of Australian primary schools. The students were from grade 3 (average age around 10) to grade 6 (average age around 13). Among them, 496 (39.1%) self-identified as Indigenous background in the demographic information, whereas 772 (60.9%) were self-identified as non-Indigenous students.

Materials

The materials used for data collection were a self-report questionnaire and students' scores in a standard math achievement test. The questionnaire started with a section on demographic information including age, gender, grade, and cultural background followed by items on self-concept of math competence and affect, which are explained in detail below.

Math Competence and Math Affect Scales

To measure students' self-concept of math competence, we used the four positive items from Marsh (1992) Self Description Questionnaire I (SDQI). These items are: "I learn things quickly in math," "Work in math is easy for me," "I get good marks in math," and "I am good at math." The items which evaluated students' self-perceptions of affect toward math were also from SDQI, including: "I like math," "I am interested in math," "I look forward to math," and "I enjoy doing work in math." We excluded the two negatively worded items from the original math competence and affect scales because past research showed that incorporating negative items resulted in negative item bias and reduced the reliability of scales (Marsh, 1986; Arens et al., 2011). The reason for using SDQI is that it is the most widely used instrument for measuring self-concept of students from diverse cultures, including Australian Indigenous students (e.g., Worrell et al., 2008; Bodkin-Andrews et al., 2010). All the items were on a 5-point scale with 1 representing strongly disagree and 5 indicating strongly agree.

Math Achievement Scores

The math achievement scores were obtained using a state-wide standardized test organized by Department of Education and Training. The test lasted about 20 min and had different items for different grades (see the **Appendix** for the sample items).

Data Collection Procedure

The data collection strictly followed the ethics requirements. Before administering the study, the written consent from the participants and their parents for voluntary participation was obtained. The effort was made to ensure the anonymity of the participants by assigning a code to each participating student. The questionnaire was administered in groups by research assistants, who read each item aloud to the students to minimize potential problems arising from reading difficulties.

Data Analysis

We started data analysis by constructing a CFA model with four items of self-concept of math competence, four items of self-concept of math affect, and the math achievement scores using Mplus version 7. Because the math achievement scores were a single-item indicator, the measurement error of scores was fixed with a perfect reliability estimate (Marsh and Yeung, 1997).

The criteria for the evaluation of CFA models followed the general procedures proposed by Kline (2005) and Jöreskog and Sörbom (2005). We considered four fit statistics as our primary indicators of model fit, namely the Comparative Fit Index (*CFI*, Bentler, 1990), the Tucker-Lewis Index (*TLI*, Tucker and Lewis, 1973), the root mean squared residual (*SRMR*, Bentler, 1995), and the root mean square error of approximation (*RMSEA*, Browne and Cudeck, 1993). The values of *TLI* and *CFI* higher than 0.950, *SRMR* less than 0.080, and *RMSEA* below 0.060 are generally considered as good fit between the hypothesized model and the observed data (Bentler, 1990; Browne and Cudeck, 1993; Hu and Bentler, 1999). Besides these fit statistics, support for the fit of CFA models also requires: (1) acceptable reliability for each scale (i.e., α = 0.700 or above); (2) factor loadings of items above 0.300 on the corresponding scales, and (3) appropriate correlations among the latent factors to ensure that they are distinguishable from each other (*r*s below 0.900).

When the CFA model was established (model 1), we then conducted a series of measurement invariance tests to determine if the CFA model was equivalent across female and male students (i.e., gender), and across Indigenous and non-Indigenous students (i.e., culture). The invariance tests involve evaluating various levels of restricted models and proceed in a stepwise manner from loosest to tightest. Therefore, the invariance models are nested because the imposed constraints are progressively added (Brown, 2006; Byrne, 2016). We first constructed three models (models 2A–2C) to examine measurement invariance by gender. We followed Brown's recommended procedure for performing invariance tests by starting from a configural CFA (model 2A), which is the least restricted model, tests whether the factor structures are identical across groups. Following the configural model, we tested whether factor loadings were equal in the metric model (model 2B). We then constrained intercepts to be equal, referred to as the scalar model (model 2C). Similarly, the next three successive models (models 3A–3C) were constructed to test whether measurement equivalence could be attained across the two cultural groups. Model 3A was a configural CFA model that examines whether the factor structures are identical across Indigenous and non-Indigenous students. Model 3B was the metric model, in which all the factor loadings were constrained to be equal across the two cultural groups of students, and model 3C was the scalar model, in which the equal constraints were put on both the factor loadings and the intercepts.

To examine the effect of gender, culture, and the interaction of the two on students' self-concept of math competence, math affect, and math achievement, we adopted a multiple-indicator-multiple-indicator-cause (MIMIC) approach to structural equation modeling (Aiken and West, 1991; Jöreskog and Sörbom, 2005; Marsh et al., 2005). The advantage of a MIMIC model is that measurement errors of latent variables are corrected (Marsh et al., 2006; Yeung et al., 2012b). The MIMIC model (model 4) examined the multiple causes of the three discrete grouping variables (1) gender (1 = boy, 2 = girl), (2) culture (1 = Indigenous, 2 = non-Indigenous), and (3) gender × culture interaction to students' self-concept of math competence, math affect, and math achievement.

RESULTS

CFA of Model 1

The CFA of model 1 produced a good fit to the data: χ^2 (25) = 188.039, *CFI* = 0.980, *TLI* = 0.970, *SRMR* = 0.019, *RMSEA* = 0.072 (**Table 1**). All factor loadings of items on their corresponding scales were above 0.750, and both self-concept of math competence and math affect scales were highly reliable (0.891 and 0.924, respectively). The correlations between math competence, affect, and math achievement scores are presented in **Table 2**, which shows that all the correlations are significant and positive—math competence and math affect: $r = 0.815$, $p < 0.010$; math competence and math achievement: $r = 0.105$, $p < 0.010$; math affect and math achievement: $r = 0.057$, $p < 0.050$. The results of correlations showed that the relation between self-concept of math competence and math affect was substantial, whereas both the relations between self-perceptions of math competence and achievement and between perceptions of math affect and achievement were weak.

Factorial Invariance Across Groups

The results of a series of invariance tests are summarized in **Table 1**. **Table 1** shows that the baseline model (model 2A) resulted in a good fit: χ^2 (50) = 245.129, *CFI* = 0.977, *TLI* = 0.967, *SRMR* = 0.022, *RMSEA* = 0.078. Both model 2B: χ^2 (56) = 248.364, *CFI* = 0.977, *TLI* = 0.971, *SRMR* = 0.024, *RMSEA* = 0.074; and model 2C: χ^2 (62) = 250.971, *CFI* = 0.978, *TLI* = 0.974, *SRMR* = 0.025, *RMSEA* = 0.069, produced similar fits to model 2A, providing evidence for the equivalence of the measurement structure across the boy and the girl groups (Cheung and Rensvold, 2002). Following the same procedure, for the invariance of culture groups, we found that the baseline model (model 3A) yielded an appropriate fit: χ^2 (50) = 239.216, *CFI* = 0.978, *TLI* = 0.968, *SRMR* = 0.021, *RMSEA* = 0.077. Across the two culture groups, the fit statistics of model 3B: χ^2 (56) = 244.244, *CFI* = 0.978, *TLI* = 0.971, *SRMR* = 0.025, *RMSEA* = 0.073; and 3C: χ^2 (62) = 254.357, *CFI* = 0.977, *TLI* = 0.974, *SRMR* = 0.027, *RMSEA* = 0.070, were comparable to those of model 3A, supporting the factorial invariance across Indigenous and non-Indigenous groups.

Paths of the MIMIC Model

The invariance of measurement across groups allowed us to proceed with MIMIC model (model 4), which displayed a

TABLE 1 | Goodness of fit of models.

Models	χ^2	*df*	*CFI*	*TLI*	*SRMR*	*RMSEA*
Model 1 (CFA of latent factors)	188.039	25	0.980	0.970	0.019	0.072
Model 2A (gender invariance-configural)	245.129	50	0.977	0.967	0.022	0.078
Model 2B (gender invariance-metric)	248.364	56	0.977	0.971	0.024	0.074
Model 2C (gender invariance-scalar)	250.971	62	0.978	0.974	0.025	0.069
Model 3A (culture invariance-configural)	239.216	50	0.978	0.968	0.021	0.077
Model 3B (culture invariance-metric)	244.244	56	0.978	0.971	0.025	0.073
Model 3C (culture invariance-scalar)	254.357	62	0.977	0.974	0.027	0.070
Model 4 (MIMIC model)	214.924	43	0.980	0.970	0.017	0.056

CFI = comparative fit index; TLI = Tucker-Lewis index; SRMR = standardized root mean square residual.

TABLE 2 | Correlation between math competence, math affect, and math achievement.

Variables	Math affect	Math achievement
Math competence	0.815**	0.105**
Math affect	—	0.057*

*p < 0.050; **p < 0.010.

TABLE 3 | Solution of model 4.

Variables	Math competence	Math affect	Math achievement
Factor loadings			
Item 1	0.753**	0.818**	1.000
Item 2	0.834**	0.866**	—
Item 3	0.830**	0.897**	—
Item 4	0.853**	0.891**	—
Uniqueness			
Item 1	0.433**	0.331**	0.000
Item 2	0.304**	0.251**	—
Item 3	0.311**	0.195**	—
Item 4	0.272**	0.207**	—
Gender	−0.083**	−0.012	−0.066*
Culture	0.032	0.073*	0.019
Interaction	0.132**	0.078**	0.009

*p < 0.050; **p < 0.010.

good fit to the data: χ^2 (43) = 214.924, *TLI* = 0.980, *CFI* = 0.970, *SRMR* = 0.017, *RMSEA* = 0.056. The factor loadings and paths of model 4 are displayed in **Table 3** and the MIMIC model is also displayed in **Figure 1**. The descriptive statistics of self-concept of math competence, math affect, and math achievement by gender and culture are presented in **Table 4**.

From **Table 3**, we can see that the main effect of gender was significant and negative for self-concept of math competence ($\beta = -0.083$, $p < 0.010$) and math achievement ($\beta = -0.066$, $p < 0.050$). The significant and negative path coefficients suggested that boys not only had higher ratings on their perceptions of abilities in math compared with girls (boys: $M = 4.092$, $SD = 1.119$; girls: $M = 3.894$, $SD = 1.131$), they also obtained higher scores in the math achievement test (boys: $M = 54.695$, $SD = 9.998$; girls: $M = 54.190$, $SD = 10.175$). The coefficients of the paths from culture were only significant for self-concept of math affect ($\beta = 0.073$, $p < 0.050$). The positive path suggested that Indigenous students had less enjoyment toward learning math compared with their non-Indigenous peers (Indigenous students: $M = 3.817$, $SD = 1.367$; non-Indigenous students: $M = 4.036$, $SD = 1.208$).

Statistically significant interaction effects between gender and culture were also found on both self-concept of math competence ($\beta = 0.132$, $p < 0.010$) and math affect ($\beta = 0.078$, $p < 0.010$). For self-concept of math competence, while non-Indigenous boys ($M = 4.154$, $SD = 1.064$) had higher ratings than non-Indigenous girls ($M = 3.935$, $SD = 1.063$), there was no significant difference between Indigenous boys ($M = 3.992$, $SD = 1.200$) and Indigenous girls ($M = 3.833$, $SD = 1.226$). For self-concept of math affect, non-Indigenous boys ($M = 4.067$, $SD = 1.223$) had higher ratings than Indigenous girls ($M = 3.803$, $SD = 1.350$), whereas there was no significant difference between non-Indigenous girls ($M = 4.010$, $SD = 1.193$) and Indigenous boys ($M = 3.834$, $SD = 1.393$). *Post hoc* power analyses indicated that with the sample size of the study, the power to detect obtained effects at the 0.050 level was 0.999 in prediction of math competence, was 0.874 in prediction of math affect, and was 0.867 in prediction of math achievement.

The present study investigated the effect of gender, culture, and the interaction of the two on Australian primary school students' self-concept of math competence and affect, as well as their math achievement. Two separate sets of measurement invariance tests on examination of invariant factor structure, factor loadings, and intercepts across both gender and culture demonstrated that female and male students, Indigenous and non-Indigenous students shared the same interpretation of the items with regard to their self-evaluation and perceptions of their capabilities and liking toward the math subject. Because of the identical pattern of factor-indicator relationships, factor loadings, and intercepts, the factor scores from the four subgroups of the sample (i.e., Indigenous boys, non-Indigenous boys, Indigenous girls, and non-Indigenous girls) can be legitimately compared. The examination of the measurement invariance ensured the potential use of the PBLEQ in various academic disciplines in higher education. The MIMIC approach found that gender differences were not only in students' perceived capabilities in the processes of learning math (self-concept of math competence) but also evident in their actual abilities in solving math problems (math achievement). Consistent with the past findings on the gender stereotype, our study also showed that girls' self-perceptions and confidence in evaluating their competence in math was lower than boys (e.g., Eccles et al., 1993; Marsh, 1993; Marsh and Yeung, 1998; Kurtz-Costes et al., 2008; Klapp Lekholm and Cliffordson, 2009). However, when taking the culture effect into consideration, we found such difference only existed among the non-Indigenous students. Such result may suggest that the general finding of the gender stereotype is only applicable among the mainstream students from Western cultural background given the fact that the majority of past studies did not investigate Indigenous population.

By separating the cognitive (competence) and affective (affect) aspects of math self-concept, our findings further extended the examination of gender effect of math competence to math affect. We observed that different from gender stereotype in self-perceptions of math competence, no gender effect was found on students' self-perceptions of liking and enjoyment of studying math subject. This means that boys and girls had similar ratings on their enjoyment of learning math. Instead, gender only had a significant impact on self-concept of math affect when it interacted with cultural backgrounds. Generally speaking, Indigenous students reported less enjoyment in learning math compared with their non-Indigenous counterparts. However, there was no such difference between Indigenous boys and non-Indigenous girls.

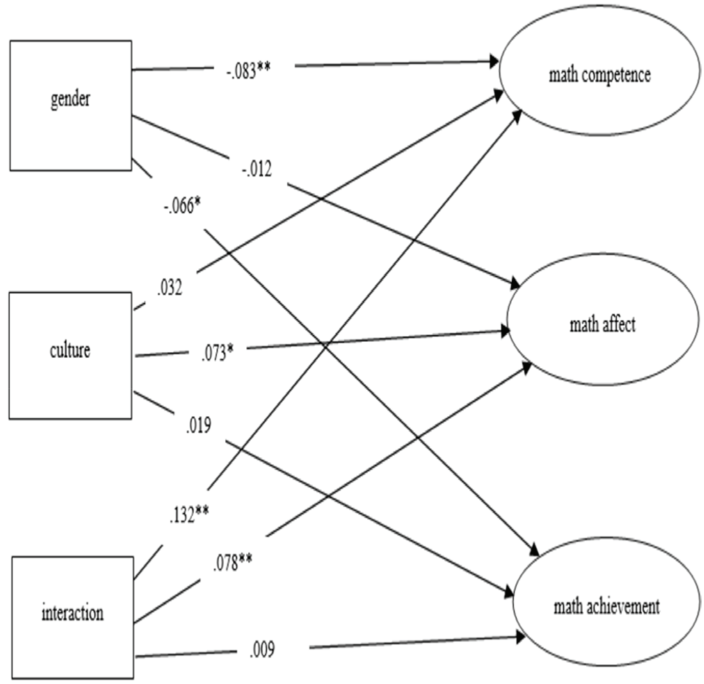


FIGURE 1 | Paths of the MIMIC model. **p < 0.010; **p < 0.050.*

TABLE 4 | Descriptive statistics by gender and culture.

Variables	Indigenous boys (N = 215)	Non-Indigenous boys (N = 351)	Indigenous girls (N = 218)	Non-Indigenous girls (N = 421)
	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**
Math competence	3.992 (1.200)	4.154 (1.064)	3.833 (1.226)	3.935 (1.063)
Math affect	3.834 (1.393)	4.067 (1.223)	3.803 (1.350)	4.010 (1.193)
Math achievement	51.737 (10.580)	56.506 (9.860)	50.591 (8.915)	56.592 (9.970)

Out of the expectation, there was no achievement gap between Indigenous and non-Indigenous students on their math achievement scores. Rather, students' performance on the math achievement test was consistent with how they self-evaluated their competence, as the results showed a gender stereotype of math achievement among Australian primary school students regardless whether they were from an Indigenous or a non-Indigenous background.

Implications for Educational Practice

In educational practice in Australian primary school contexts, teachers may need to make some efforts in boosting Australian female primary school students' beliefs of their competence in math subject. Based on the known reciprocal effects between self-concept and achievement (Marsh and Craven, 2006), students are likely to improve their math performance through interventions that focus on boosting their confidence in math competence. Furthermore, educators need to pay special attention to Indigenous students' lower interest and enjoyment in math learning. Past research has shown that when teaching is designed to incorporate Indigenous students' values, beliefs, and traditions, and when new knowledge is delivered in a way that is culturally appropriate to Indigenous students, the learning tends to be more effective (e.g., Gitari, 2006; Yunkaporta, 2009). Thus, to nurture Indigenous students' interests in math learning, teachers may consider using some materials which are able to foster relevance to Indigenous students' culture.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of "National Statement on Ethical Conduct in Human Research, the Human Research Ethics Committee of Western Sydney University" with written informed consent from all subjects and their parents. All subjects and their parents gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by "the Human Research Ethics Committee of Western Sydney University."

AUTHOR CONTRIBUTIONS

FH contributed to the conception of the work and to the acquisition, analysis, and interpretation of the data; drafted the work and revised it critically for important intellectual content; approved the final version of the paper to be published, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

ACKNOWLEDGMENTS

The author wishes to express gratitude to the participants for their voluntary participation in the study.

REFERENCES

Aiken, L. S., and West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* (Newbury Park, CA: Sage).

Arens, A. K., Yeung, A. S., Craven, R. G., and Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: domain specificity and separation between competence and affect components. *J. Educ. Psychol.* 103, 970–981. doi: 10.1037/a0025047

Australian Bureau of Statistics (2012). *The health and welfare of Australia's Indigenous and Torres Strait Islander peoples.* (Canberra: Australian Institute of Health and Welfare).

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychol. Bull.* 107, 238–246. doi: 10.1037/0033-2909.107.2.238

Bentler, P. M. (1995). *EQS structural equations program manual.* (Encino, CA: Multivariate Software).

Bodkin-Andrews, G. H., Ha, M. T., Craven, R. G., and Yeung, A. S. (2010). Construct validation and latent mean differences for the self-concepts of Indigenous and non-Indigenous students. *Int. J. Test.* 10, 47–79. doi: 10.1080/15305050903352065

Brown, T. (2006). *Confirmatory factor analysis for applied research.* (New York: The Guildford Press).

Browne, M. W., and Cudeck, R. (1993). "Alternative ways of assessing model fit" in *Testing structural equation models.* eds. K. A. Bollen and J. S. Long (Beverly Hills: Sage), 136–162.

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming.* (New York: Routledge).

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Modeling* 9, 233–255. doi: 10.1207/S15328007SEM0902_5

Chiu, M. M., and Klassen, R. M. (2010). Relations of math self-concept and its ca bration with math achievement: cultural differences among fifteen year olds in 34 countries. *Learn. Motiv.* 20, 2–17. doi: 10.1016/j.learninstruc.2008.11.002

Commonwealth of Australia (2006). *The national report to parliament on Indigenous education and training.* (Canberra: Author).

Cooke, M., Mitrou, F., Lawrence, D., Guimond, E., and Beavon, D. (2007). Indigenous well-being in four countries: an application of the UNDP'S human development index to Indigenous peoples in Australia, Canada, New Zealand, and the United States. *BMC Int. Health Hum. Rights* 7, 1–39. doi: 10.1186/1472-698X-7-9

Craven, R. G., and Marsh, H. W. (2008). The centrality of the self-concept construct for psychological wellbeing and unlocking human potential: implications for child and educational psychologists. *Educ. Child Psychol.* 25, 104–118.

Craven, R. G., and Yeung, A. S. (2008). "International best practice in effective educational interventions: why self-concept matters and examples from bullying, peer support, and reading research" in *Research on sociocultural influences on motivation and learning. Vol. 8: Teaching and learning: International best practice.* eds. D. M. McInerney, V. E. Shawn and M. Dowson (Greenwich, CT: Information Age), 267–294.

Craven, R. G., Dillon, A., and Parbury, N. (2013). *In black & white: Australians all at the crossroads.* (Ballan: Connor Court Publishing).

Dai, D. Y. (2001). A comparison of gender differences in academic self-concept and motivation between high-ability and average Chinese adolescents. *J. Sec. Gift. Educ.* 13, 22–32. doi: 10.4219/jsge-2001-361

Eccles, J. S., Wigfield, A., Harold, R., and Blumenfeld, P. (1993). Age and gender differences in children's achievement self-perceptions during the elementary school years. *Child Dev.* 64, 830–847. doi: 10.2307/1131221

Gitari, W. (2006). Everyday objects of learning about health and healing and implications for science education. *J. Res. Sci. Teach.* 43, 172–193. doi: 10.1002/tea.20094

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55.

Jansen, M., Schroeders, U., and Lüdtke, O. (2014). Academic self-concept in science: multidimensionality, relations to achievement measures, and gender differences. *Learn. Individ. Differ.* 30, 11–21. doi: 10.1016/j.lindif.2013.12.003

Jöreskog, K. G., and Sörbom, D. (2005). *LISREL 8.72: Structural equation modelling with SIMPLIS command language.* (Chicago, IL: Scientific Software International).

Klapp Lekholm, A., and Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educ. Res. Eval.* 15, 1–23. doi: 10.1080/13803610802470425

Kline, R. B. (2005). *Principles and practices of structural equation modelling.* 2nd edn. (New York: The Guilford Press).

Kurtz-Costes, B., Rowley, S., Harris-Britt, A., and Woods, T. A. (2008). Gender stereotypes about math and science and self-perceptions of ability in late childhood and early adolescence. *Merrill-Palmer Q.* 54, 386–409. doi: 10.1353/mpq.0.0001

Lai, F. (2010). Are boys left behind? The evolution of the gender achievement gap in Beijing's middle schools. *Econ. Educ. Rev.* 29, 383–399. doi: 10.1016/j.econedurev.2009.07.009

Marsh, H. W. (1986). Negative item bias in rating scales for preadolescent children: a cognitive-developmental phenomenon. *Dev. Psychol.* 22, 37–49. doi: 10.1037/0012-1649.22.1.37

Marsh, H. W. (1990). The structure of academic self-concept: the Marsh/Shavelson model. *J. Educ. Psychol.* 82, 623–636. doi: 10.1037/0022-0663.82.4.623

Marsh, H. W. (1992). *Self Description Questionnaire: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept: A test manual and a research monograph.* (Sydney: SELF Research Centre, University of Western Sydney).

Marsh, H. W. (1993). The multidimensional structure of academic self-concept: invariance over gender and age. *Am. Educ. Res. J.* 30, 841–860.

Marsh, H. W., and Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspective. *Perspect. Psychol. Sci.* 1, 133–163. doi: 10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., Ellis, L., Parada, L., Richards, G., and Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: operationalizing criteria for short form evaluation with new applications of confirmatory factor analyses. *Psychol. Assess.* 17, 81–102. doi: 10.1037/1040-3590.17.1.81

Marsh, H. W., Tracey, D., and Craven, R. G. (2006). Multidimensional self-concept structure for preadolescents with mild intellectual disabilities: a hybrid multigroup-MIMIC approach to factorial invariance and latent mean differences. *Educ. Psychol. Meas.* 66, 705–818. doi: 10.1177/0013164405285910

Marsh, H. W., and Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: structural equation models of longitudinal data. *J. Educ. Psychol.* 89, 41–54. doi: 10.1037/0022-0663.89.1.41

Marsh, H. W., and Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: gender differences in the development of math and English constructs. *Am. Educ. Res. J.* 35, 705–738.

MCEETYA (2008). Melbourne declaration on educational goals for young Australians. Available at: http://www.curriculum.edu.au/verve/_resources/National_Declaration_on_the_Educational_Goals_for_Young_Australians.pdf (Accessed December 9, 2018).

Meece, J. L., Glienke, B. B., and Burg, S. (2006). Gender and motivation. *J. Sch. Psychol.* 44, 351–373. doi: 10.1016/j.jsp.2006.04.004

Metallidou, P., and Vlachou, A. (2007). Motivational beliefs, cognitive engagement, and achievement in language and math in elementary school children. *Int. J. Psychol.* 42, 2–15. doi: 10.1080/00207590500411179

Midgley, C., Kaplan, A., and Middleton, M. (2001). Performance-approach goals: good for what, for whom, under what circumstances, and at what cost? *J. Educ. Psychol.* 93, 77–86. doi: 10.1037/0022-0663.93.1.77

Pinxten, M., De Fraine, B., Van Damme, J., and D'Haenens, E. (2013). Student achievement and academic self-concept among secondary students in Flanders: gender and changes over time. *Irish Educ. Stud.* 32, 157–178. doi: 10.1080/03323315.2012.749058

Seaton, M., Parker, P., Marsh, H. W., Craven, G., and Yeung, A. S. (2014). The reciprocal relations between self-concept, motivation and achievement: juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educ. Psychol.* 34, 49–72. doi: 10.1080/01443410.2013.825232

Tucker, L. R., and Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 38, 1–10. doi: 10.1007/BF02291170

Usher, E. L., and Pajares, F. (2008). Sources of self-efficacy in school: critical review of the literature and future directions. *Rev. Educ. Res.* 78, 751–796. doi: 10.3102/0034654308321456

Worrell, F. C., Watkins, M. W., and Hall, T. E. (2008). Reliability and validity of self-concept scores in secondary school students in Trinidad and Tobago. *Sch. Psychol. Int.* 29, 466–480. doi: 10.1177/0143034308096435

Yeung, A. S., Craven, R. G., and Kaur, G. (2012a). "Gender differences in achievement motivation: grade and cultural considerations" in *Psychology of gender differences.* ed. S. McGeown (New York: Nova Science Publishers), 59–79.

Yeung, A., and Han, F. (2017). "Chinese-background Australian students' academic self-concept, motivational goals, and achievements in math and English" in *Educating Chinese-heritage students in the global-local nexus: Achievement, challenges, and opportunities.* eds. G. Li and W. Ma (New York: Routledge), 141–159.

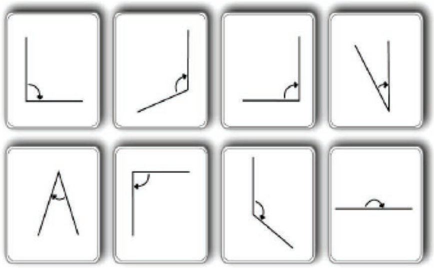
Yeung, A. S., Taylor, P. G., Hui, C., Lam-Chiang, A. C., and Low, E.-L. (2012b). Mandatory use of technology in teaching: who cares and so what? *Br. J. Educ. Technol.* 46, 857–570. doi: 10.1111/j.1467-8535.2011.01253.x

Yunkaporta, T. (2009). Aboriginal pedagogies at the cultural interface. Unpublished doctoral thesis. (Australia: James Cook University).

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX: SAMPLE ITEMS OF MATH ACHIEVEMENT TESTS

Ben put 8 angle cards in a group.



How many cards showing right angles are in this group?

- ⬭ 1
- ⬭ 3
- ⬭ 2
- ⬭ 7

Look at this number pattern.

The numbers follow the same rule.

58, 52, 46, 40, 34, __, __, ? .

Which number goes here?

- ⬭ 6
- ⬭ 22
- ⬭ 16
- ⬭ 28

frontiers in Psychology

ORIGINAL RESEARCH
published: 29 March 2019
doi: 10.3389/fpsyg.2019.00594

Contextual Choices in Online Physics Problems: Promising Insights Into Closing the Gender Gap

Samuel R. Wheeler[1]* and Margaret R. Blanchard[2]

[1]Department of Physics, North Carolina School of Science and Math, Durham, NC, United States, [2]Department of STEM Education, North Carolina State University, Raleigh, NC, United States

Throughout the world, female students are less likely than males to take advanced physics courses. This mixed-methods study uses a concurrent, nested design to study an online homework intervention designed to address choice and achievement. A choice of three different contexts (biological, sports, and traditional) were offered to students for each physics problem, intending to stimulate females' interest and enhance achievement. Informed by aspects of Artino's social-cognitive model of academic motivation and emotion, we investigated: Which context of physics problems do males and females select?; What explanations do students give for their choices?; Are there differences in the achievement of males and females?; and Is there a relationship between student achievement and the context selected? Fifty-two high school physics students from five US states participated. Data included pre- and post-Force Concept Inventory scores, homework context choices and achievement, and rationales for choices. Findings indicate that females were most likely to select biology contexts; males, traditional. All students made more attempts on video questions over word questions, although females did not score as well. For all questions, students generally persisted until they answered them correctly, with females taking fewer attempts on problems. Context choice was mostly driven by interest, for males, and perceptions of difficulty level for females; however, rationales were indistinguishable by gender. On their first homework question attempt, females scored significantly better than the males. Initially, males had significantly higher FCI scores; post homework intervention, females increased their mean scores significantly on the FCI, erasing the initial gender gap, with no growth nor decline in males' scores. Females with FCI growth were equally as likely to choose biology contexts as traditional contexts; males were more likely to choose biology contexts. Findings from this study suggest that modest changes to homework problems that provide choice and make the physics problems more contextually interesting—even without changes in classroom instruction—could increase interest and motivation in students and increase achievement for both male and female students. Recommendations will be discussed.

Keywords: gender, physics, online, choice, motivation

OPEN ACCESS

Edited by:
*Bernhard Ertl,
Universität der Bundeswehr
München, Germany*

Reviewed by:
*Bibiana Regueiro,
University of A Coruña, Spain
Bianca Watzka,
Ludwig Maximilian University of
Munich, Germany*

***Correspondence:**
*Samuel R. Wheeler
srwheele@ncsu.edu*

Specialty section:
*This article was submitted to
Gender, Sex and Sexuality Studies,
a section of the journal
Frontiers in Psychology*

Received: *27 December 2018*
Accepted: *04 March 2019*
Published: *29 March 2019*

Citation:
*Wheeler SR and Blanchard MR
(2019) Contextual Choices in Online
Physics Problems: Promising Insights
Into Closing the Gender Gap.
Front. Psychol. 10:594.
doi: 10.3389/fpsyg.2019.00594*

INTRODUCTION

In the 21st century, the demand for skilled workers in Science, Technology, Engineering, and Mathematics (STEM) is outpacing the rate at which they are produced from universities (Wieman, 2012). The National Research Council (NRC, 2013) identified physics as the ultimate foundation for all the other branches of science, with over 500,000 students a year taking an introductory physics course in the United States (US), but only 1% of college graduates completing a degree in physics. In the early grades, there is no gender gap in interest in STEM subjects for US students; yet, from the time, a young girl enters kindergarten until the time she begins her senior year of high school, chances are that she will have lost much of her interest in STEM subjects as compared with her male peers (Baram-Tsabari and Yarden, 2011). This drop-off in interest begins before students go to college (Heilbronner, 2013).

The problem with fewer female students choosing to take advanced physics courses has been documented throughout the world, including Ghana (Buabeng et al., 2012), Scotland (Reid and Skryabina, 2003), Australia (Oliver et al., 2017), England, Singapore, Spain, and Mexico (Oon and Subramaniam, 2010). In the US, approximately 36% of undergraduate STEM degrees and 19% of undergraduate physics degrees were awarded to women in 2015 (American Physical Society, 2015). Similarly, the gender gap for graduate degrees is 23% of masters' degrees (Mulvey and Nicolson, 2014) and 21% of PhDs that are awarded go to women (American Physical Society, 2015).

The gender gap between the enrollment of male and female students in physics and the physical sciences points to three influences that place pressures on both genders to adhere to established stereotypes: cultural, attitudinal, and educational (Baram-Tsabari and Yarden, 2008). Cultural influences stem from established societal views of the "male image of science": parental beliefs that girls are not as interested in science as are boys (particularly in the physical sciences), family responsibilities, and lack of support when in a STEM occupation. Early exposure to STEM activities and family influences have been found to contribute to long-term female student motivation to pursue a professional career in STEM fields (Talley and Martinez Ortiz, 2017).

Among the challenges that young women face in physics and engineering degree programs are microaggressions; brief, but frequent everyday interactions that send subtle but negative messages to them that they cannot be scientists or physicists (Grossman and Porche, 2014). Stereotype threat is a well-studied phenomenon that occurs when "a stereotype about an individual's social or racial group can provide a potential explanation for the person's poor performance" is thought to be a contributing factor to creating the gender gap in mathematics and is believed to be a contributing factor in the observed gender gap in physics (Marchand and Taasoobshirazi, 2012), p. 3051.

Attitudinal influences undermining girls' interest in science include perceptions of the impersonal nature of physical sciences, difficulty with the material, and an image of the physical sciences as a masculine field (Baram-Tsabari and Yarden, 2008). Some assert that the gender gap in STEM is due to female student perceptions of engineers and physicists as being "nerdy" and "reclusive" people who have no time for interactions and relationships (Johnson, 2012). Females' perceptions of educational barriers to learning and doing physics impede their full exploration and immersion in the subjects (Grossman and Porche, 2014). In addition to addressing the classroom environment and traditional pedagogy, researchers recommend making physics more personally relevant to girls (Murphy and Whitelegg, 2006; Baram-Tsabari and Yarden, 2008; Gibson et al., 2015).

Until recently, stereotypical masculine interests and characteristics were widely represented in the images and language used in textbooks with references to male names and traditionally male activities and images (McCullough, 2007). In addition to the textbooks used, validated formal assessments such as the Force Concept Inventory (FCI), one of the most widely used physics concept assessments (Hestenes et al., 1992), is largely dominated by questions from stereotypical male contexts (McCullough, 2004). These contexts lay the foundation for gender biases, which send the message to young female students that they may lack the aptitude to do well in physics or in STEM-related fields (Grossman and Porche, 2014). From the perspective of both male and female students, physics tends to be personified by masculine traits; from the teacher's perspective, physics is perceived as having characteristics from both genders (Makarova and Herzog, 2015).

Interest and positive student motivation toward STEM subjects have been linked to the use of collaborative learning and social modeling in the classroom (Bryan et al., 2011). Sawtelle et al. (2012) found that "vicarious learning experiences," seeing a particular task they are expected to perform modeled for them and comparing their achievement to that of others, positively influenced the development of female students' self-efficacy in physics, a strong factor in perseverance in physics classes.

The gender gap in physics was once attributed to the assumption that the subject was too difficult for females, and programs were developed to address girls' deficiencies (Zohar and Sela, 2003). But the gap is not due to lack of ability; female students who take physics in high school are just as likely to succeed in the course as male students (NRC, 2013). Stereotypically, male students tend to be interested in physics for the sake of physics, while female students tend to report being interested in physics for the sake of what physics can do to help humankind and other social associations (Bøe and Henriksen, 2013). Female role models in physics, such as a female physics teacher or physicist, can positively impact female students' attitudes and interest in physics by providing someone who has a "physics identity" for female students to observe; yet, these role models are few (Hazari et al., 2010). McCullough (2007) recommends the use of specific language in physics examples and problems that involve familiar, relevant contexts for all students, such as cars, food, and school activities. Other researchers suggest tapping into the interests of female students by integrating medical and biological fields into the traditional physics curriculum (Gibson et al., 2006). In the United Kingdom (UK), a study found that female physics students wanted pedagogies that connected the relevance of physics with the greater world and with their own interests, suggesting the creation of a curriculum that relates physics to health applications and the human body (Mitrevski and Treagust, 2011).

The interdisciplinary approach to teaching physics by incorporating life science into the curriculum is on the rise, mainly as a response to the greater demand for students to more fully understand the relevance of physics in relation to biology and chemistry (Crouch and Heller, 2014). Crouch and Heller designed a course for the growing number of life science majors who need physics, to deliver a "coherent view of physics as a discipline" (p. 379). Others have recommended the integration of biology and physics in university courses to begin recognizing the similarities of the two disciplines instead of the differences (Hoskinson et al., 2014). One study found that by incorporating topics and phenomena that students do not encounter in everyday life into the physics curricula, students become more interested in the physics concepts (Badri et al., 2016). The Badri et al. (2016) study found that females were more interested in phenomena that could not be easily explained by high school physics, while males were more interested in traditional phenomena such as mechanical equipment and lasers.

Giving students choice in their assignments or classroom is thought to be a key factor in supporting and fostering intrinsic motivation (Patall et al., 2008). One study investigated giving students three choices for a task and found that participants who were already interested in a concept or topic showed more motivation and better performance on the task when given the opportunity to choose, which did not happen for disinterested students (Patall, 2013). Patall et al. (2010) concluded that performance and engagement stemmed from intrinsic motivation to complete the task. Thus, giving students choice is a key factor in supporting and fostering intrinsic motivation. Others have also found that choice can be a motivating factor when the choice is meaningful, relevant, and enhances the competence of the student (Evans and Boucher, 2015).

Teachers with rich content knowledge and enthusiasm toward teaching can result in positive gains in student motivation in physics (Keller et al., 2017). One recent study showed that female students' motivation to study and do well in physics is linked to several factors, including having a combination of teachers, supportive and knowledgeable teachers, engaging pedagogy, the school's science culture, and social interactions with family and peers (Oliver et al., 2017). In another study, students' motivation in physics was positively related to the task-value they saw in the physics they were doing and interest in the science being studied (Wang et al., 2017). Similar results were found in a Croatian study, which suggested that a key motivational factor for female students was perceptions of its utility value for students (Jugović, 2017).

Another important factor in learning physics is how students comprehend a range of multimedia representations, such as visual pictures (e.g., a bar graph or photo), visual texts (written information), and sound (Schnotz, 2014). Learning occurs when an individual understands what is presented; that is, "when the individual uses external representations in order to construct internal (mental) representations of the learning content in working memory and if he or she stores those in long-term memory" (p. 75). Mayer's generative theory of textbook design (Mayer and Sims, 1994; Mayer et al., 1995) focuses on the relationship between illustrations in textbooks and the corresponding text. Illustrations, photos, drawings, and animations are examples of visualizations, a type of multimedia representation involving spatial relations that communicate information (Scheiter et al., 2009). Mayer et al. (1995) found that students received higher scores when illustrations were accompanied by text in close proximity. The use of pictures and illustrations most enhances student learning when the image and the information from the text are integrated, compared to text only (and the complexity of the diagrams influences the outcome) (Mason et al., 2013; Jian and Wu, 2015). The learning is enhanced when the words and pictures are semantically related, if they are presented close together in space or time, and when the picture appears before the text (Schnotz, 2014). Using their spatial ability helps students to consolidate and clarify ideas, remember ideas, and helps with problem-solving (Baker and Pilburn, 1997).

The use of videos as pedagogical tools was the next logical step from diagrams and photographs and was originally seen as a way to introduce concepts to students that would motivate them to explore the concept further, to understand more, and to examine "what if" questions—therefore allowing them time to bridge the gap between the abstract and the concrete (Zollman and Fuller, 1994). Videos and video analysis technology as pedagogical tools were introduced over 25 years ago to more effectively teach kinematics and help students better understand the physics of motion (Beichner et al., 1989), and has been found to increase student excitement and engagement with the material being presented (Lee and Sharma, 2008).

The development of additional video analysis software (e.g., Vernier's Logger Pro©, Pasco's© commercial versions) and other technologies developed specifically for the physics classroom provide students with the ability to collect real-time data, which can motivate them to want to learn the underlying physics concepts and also provide a way for them to more easily clarify and correct their misconceptions about motion (Beichner and Abbott, 1999). Struck and Yerrick (2010) found that video analysis as the sole lab technology more effectively promoted student comprehension. However, interactivity with a computer and controls on an animation or dynamic displays as well as inaccurate prior knowledge can reduce comprehension (Hegarty, 2014).

THEORETICAL FRAMEWORK

This study was guided by Artino's (2010) social-cognitive model of achievement motivation (**Figure 1**). Based on his work with at-risk students in an online setting, Artino found that students

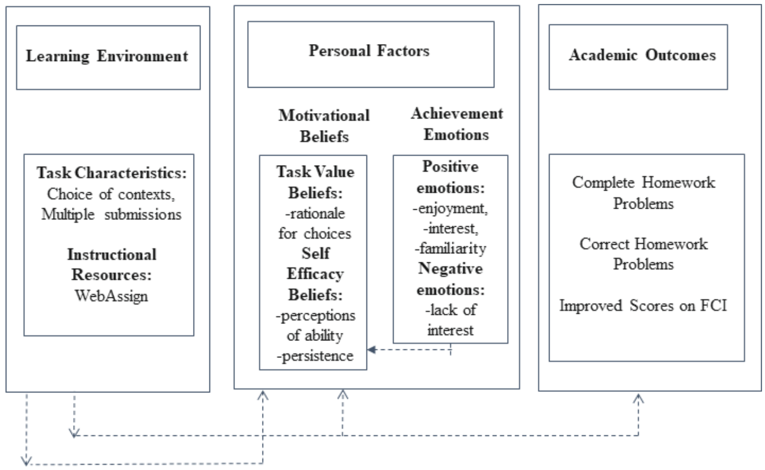


FIGURE 1 | Predicted social-cognitive model of achievement motivation and emotion (adapted from Artino, 2010).

who were more satisfied with their experiences and more confident in their abilities were more likely to prefer taking online courses in the future. In his model, the learning environment and motivational beliefs contribute to (dis) satisfaction and academic outcomes. Students' motivational beliefs are directly linked to student self-efficacy or one's beliefs about the task's interest and significance, which will determine his or her motivation for completing that task (Eccles and Wigfield, 2002). Phillips (2015) used Artino's model to investigate motivational factors of at risk students who worked on self-paced, online modules during summer school to remediate a failed science course. She found that most of the students were satisfied with the opportunity to set their own pace and take control of their learning, and all of the students achieved passing grades for their courses (Phillips, 2015). Achievement emotions describe the feelings (e.g., boredom, enjoyment) that are the direct result of achievement outcomes experienced during learning, which influence self-efficacy beliefs and task value beliefs (Pekrun, 2006). The model predicts that students' motivational beliefs and achievement emotions are linked to their academic outcomes and are influenced by the learning environment.

When translating the context of this study into Artino's (2010) model (see **Figure 1**), the *Learning Environment* includes the instructional resource WebAssign, and the task characteristics of the homework assignment are students' choice of contexts and the ability to re-take the items, use multiple contexts, and take the needed time for both word problems and video problems. Under *Personal Factors*: *Motivational Beliefs*, the students provided a rationale for the homework contexts that they chose, some of which related to perceptions of their ability, and students were allowed to persist up to five attempts on the problem. Under *Personal Factors: Achievement Emotions*, students could report such aspects as enjoyment, interest, familiarity, or lack of interest to explain their choices. The *Academic Outcomes* considered in this manuscript were completion of and success on homework problems and (improved) scores on the Force Concept Inventory (FCI).

RESEARCH QUESTIONS

The main objective for this study, guided by our theoretical framework, was to examine choices students made during a research-designed, online physics homework intervention focused on Newton's Laws and their applications. Therefore, the research questions guiding this study were:

(1) When given problem choices designed around gender stereotypes, which types of physics problems do males and females select? (2) What explanations do students give for the problem contexts they select? (3) Are there differences in the achievement of males and females?, and (4) Is there a relationship between student achievement and the context of physics problems they select?

MATERIALS AND METHODS

Research Design

This mixed-methods study was designed to help us understand more about the role of gender on choice and achievement for high school physics students. Students were given choices of contexts that were designed by the first author, based on research for how to improve females' interest in physics. This study used a concurrent, nested (embedded) design (Creswell et al., 2003); the main data collection was quantitative data, but nested qualitative data was gathered to understand students' rationales for their choices.

Recruitment and Participants

Fifty-two students, 21 females (40.4%) and 31 males (59.6%) representing eight different schools from five states in the United States (US), took part in this study (53.8% white, 32.7% Asian, 9.6% Hispanic, 1.9% Native American, 1.9% other). All of the students were currently enrolled in Honors Physics or AP Physics at their schools. All the students had completed a unit on Newton's Laws and their applications prior to this

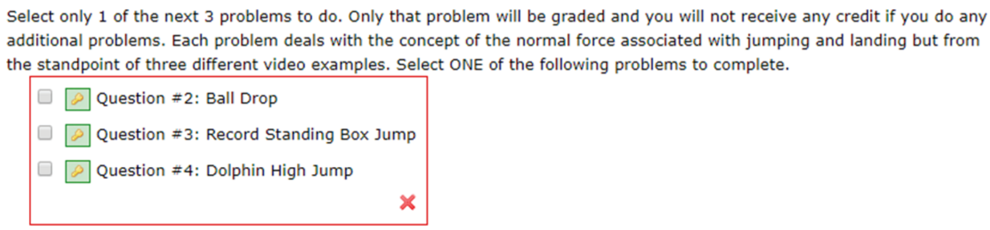


FIGURE 2 | Sample view of context choices in WebAssign.

study. The Institutional Review Board (IRB) at the university approved the proposed study.

Homework Problems

The sets of physics problems used in this study were delivered to the students through the online homework delivery system WebAssign.[1] The WebAssign problem set instrument consisted of a total of 21 problem sets made up of 16 word problems, and 5 video analysis problems, all of which were developed by the first author (for additional detail see Wheeler, 2017). Each set of problems (word and video) contained three questions of equivalent difficulty and covered the same physics concept but using different contexts (i.e., sports, biological, or traditional). The homework questions were validated and vetted by current physics teachers to ensure equivalent difficulty level, content, and consistency of the intended contexts. WebAssign was provided freely to all study participants.

Each problem set consisted of the same fundamental physics problem but from the perspective of three different format contexts: traditional (e.g., ramp, ball, pendulum), sports (e.g., baseball, basketball, extreme sports), or biological (e.g., frog, cat, leopard). WebAssign was programmed so that students were first presented with three different scenarios to choose from, without being able to see the actual question. For example, the student was shown that the question was about the concept of net forces and asked which context they wish to choose to investigate the concept: mass and spring, the high jump, or the baby bird. **Figure 2** shows an actual view in WebAssign that the student would see. WebAssign collected data on the students' context choices, the order of those choices, and whether the responses were correct. Students received immediate feedback from WebAssign on each question they answered as to whether their answers were correct or not (either a green check if the answer was correct, or a red "x" if the answer was incorrect).

Student Rationale

After completing each question, students were asked to give the reason (rationale) for why they chose the particular question context by writing a short answer response in WebAssign. Codes were developed by the first author by reading through the (blinded) student responses and creating categories that corresponded to their choice explanations. Once the initial categories were developed, they were collapsed into similar categories (5). An independent physics education researcher took students' responses and coded them using the researcher-established (first author) five codes: (1) Interest, (2) Familiarity of the problem, (3) Random Choice, (4) No preference, and (5) Easy/straightforward. Differences between the results were discussed and resolved after careful evaluation of the coding scheme in which two of the codes (3—random choice and 4—no preference) were collapsed into one. After the refinements, the final inter-rater reliability was found using Cohen's kappa to be 0.875 (87.5%) agreement.

Force Concept Inventory

Conceptual understanding of Newtonian concepts was measured, before and after the delivery of the problem sets, by using a 29-item Force Concept Inventory (FCI). The FCI was developed in the early 1990s by physics educators who saw the need to more fully understand student misconceptions in order to design more effective introductory physics courses (Hestenes et al., 1992). The Force Concept Inventory was designed specifically around what the authors called "common sense alternatives" to actual Newtonian physics since, as Hestenes et al. (1992) found, many students coming into introductory physics courses fail to grasp Newtonian concepts but instead rely on their own misconceptions and beliefs that do not match scientific explanations. The FCI has been validated and found to be a reliable tool for identifying how much students understand about the physics concepts of Newton's Laws and forces (Hestenes et al., 1992). Although there have been concerns about the gender bias of the items as being dominated by questions from stereotypical male contexts (McCullough, 2004; Grossman and Porche, 2014), as this was the most widely used instrument available that was linked to content focus of the unit, it was selected to measure students' pre and post content knowledge.

FINDINGS

Physics Problems Selected by Males and Females

The first research question investigated the types of physics problems males and females selected. Overall, the females were more likely to choose the biology context (37.6%) over the traditional or sports context (see **Table 1**). In the "Female 1st

[1]www.webassign.net

Choice Combo" category, females who chose to complete combinations (more than one context selected, per question) were more likely to choose to do the biology context (39.7%) over the other contexts. Female students who did not choose to complete a second context were more likely to choose the traditional context (37.6%) over the biology (34.9%) or sports context (25.9%). Females who chose to complete multiple contexts tended to select more traditional contexts throughout the project and fewer sports contexts, while their choices of the biology context remained relatively constant. Females who chose only one context tended to select more biology contexts and fewer traditional contexts across the project, while the number of sports contexts remained constant. The breakdown of each category, by percentage, is recorded in **Table 1**. Twelve (57%) female students chose more than one context to complete. The highest rate of context pair choices for female students was the traditional/sports context pair choices. Females' second highest combination was completing every context (traditional, sports, and biology) in the problem.

Overall, the males were more likely to choose the traditional context (38.9% of the time) as compared with biology context (33.2% of the time) and sports (27% of the time). Males chose the traditional contexts 38.1% of the time (as part of a combo), and 40.5% of the time (traditional without a combo), with approximately the same rate of choice for biology (with combo 33.3%; without combo 32.9%) and sports (with combo 27.4%; without combo 26.2%). Males who selected multiple contexts tended to select more biology contexts and fewer sports questions, while the number of traditional contexts selected remained relatively low. Males who chose only one context tended to select more traditional questions and tended to select slightly fewer sports and biology contexts across the duration of the project. Overall, males were most likely to choose traditional contexts (38.9%) (see **Table 1**).

TABLE 1 | Student choices of question context by gender.

Group	Number of students	Traditional contexts chosen		Biology contexts chosen		Sports contexts chosen	
		Total	%	Total	%	Total	%
Female (overall)	21	154	34.9	166	**37.6**	116	26.3
Male (overall)	31	253	**38.9**	216	33.2	176	27.0
Female 1st choice combo	12	83	32.9	100	**39.7**	67	26.6
Male 1st choice combo	21	168	**38.1**	147	33.3	121	27.4
Female No Combo	9	71	**37.6**	66	34.9	49	25.9
Male No Combo	10	85	**40.5**	69	32.9	55	26.2

Each row may not add up to 100% because some of the options were not selected, although this only represents fewer than 2% of the choices.
Note: Highest values for each group are bolded.

Males were more likely to choose multiple contexts within one problem set than females (males, 72; females, 37); however, there were more male participants than female participants. When looking at the percentage of choices by gender, 57% of females chose multiple contexts and 68% of males did. The highest rate of context pair choices for female students was the traditional/sports context pair choices. Male students' highest rate of choice for context pairs was for biology/sports. Males' second highest combination was completing every context in the problem. Less than 2% of the questions were not answered by each group and did not significantly affect the total.

Students' Explanations for the Question Contexts That They Selected

The second research question addressed students' rationale for the choices that they made. Students' written explanations for problem context choice differed by gender. Males were more likely to select a question context based on interest, such as "I like leopards" (males 45% of the time vs. 34.3% females; $p < 0.001$; Cohen's $d = 1.1$). Females were more likely to choose a problem because they thought it looked easier (32.6% of the time; males 25%; $p = 0.026$; Cohen's $d = 0.306$), with no significant differences between males and females on problem selection due to it being familiar, such as the "type of mechanics I am most familiar with" (17.6% males; 12.3% females) or just a "random" choice (15% males; 15.6% females). As you can see in **Table 2**, the written explanations by male and female students were quite similar.

Differences Between Males and Females on Homework Submissions and Success

The third research question investigated differences in achievement for males and females. First, the homework questions that were completed on WebAssign were examined. For all of the homework questions, students generally persisted until they answered them correctly (up to five attempts allowed), and a majority of the students took advantage of the opportunity to choose different (and therefore, multiple) question contexts (55% of females; 68% of males). Students were more likely to make attempts on video problems (V) versus the word problems (W); on video problems, males averaged 2.52 attempts on traditional (VTr) contexts (vs. 2.3 WTr) and 3.38 attempts on biology (VBio) contexts (vs. 2.52 WBio) (see **Table 3**). Females averaged from 2.78 attempts on video sports contexts (vs. 2.23 WSp) to 3.02 attempts on traditional contexts (vs. 2.24 WTr). Although females were most likely to choose biology contexts, overall (37.6%), and most likely to choose biology as a first choice (39.7%) (**Table 1**), they had the fewest number of *attempts* to answer the biology context questions (1.88; **Table 5**), suggesting that they were more likely to answer these questions correctly with fewer attempts than with other contexts (i.e., 2.24 attempts Tr, 2.23 attempts Sp).

TABLE 2 | Sample male and female responses for choosing video problems (VP) and word problems (WP).

Gender	*Student response*	% of gender who chose question
	5V(Tr)	
	Video problem deals with circular forces (ball in loop, running man in loop, car in loop)	
M	• *The ball in a loop seemed kind of borinVg, and the car in a loop isn't that impressive, but a human running in a perfect loop is pretty cool.* • *I enjoyed seeing the 5 second video and was motivated to actually complete the problem because of it.*	57.1
F	• *Because it's more impressive than the others, and cool to watch and think about.* • *ITS GOT A HUMAN and its cool! I think I did it the correct way and I have looked through my notes to find it.*	42.9
	10V (Sp)	
	Video problem about friction/resistive forces (arrow in gel, plane landing, frog jumping)	
M	• *I choose this because I shoot bow and arrow sometimes and I thought it was cool.* • *The context was interesting, as I have not seen an object slow in its velocity due to a solid.*	62.5
F	• *A jumping frog is easy to picture.* • *I feel more confident calculating parabolas.*	37.5
	3(Bio)	
	Word problem deals with tension in a taut line (spider & thread; mountain climber & rope; elevator & cable)	
M	• *The problem, as I read it, was extremely straightforward and I knew what I needed to do to solve it immediately.* • *I thought that incorporating "spider" and "fly" made it interesting*	75
F	• *This problem was straightforward and easy to understand. The context was interesting.* • *Animals make the problem more relatable and "friendly of sorts".*	25

As shown in **Table 4**, for both males and females, students scored highest (correct response on homework question) on the biology contexts than any other context. This trend follows with the traditional context, which is the second highest score correct for every gender and group, and then the sports contexts (which had the lowest score correct for each gender and group). The number of submissions did not follow such a simple pattern, but the general trend was that questions that had more submissions suggest that the student stayed with the problem longer, while fewer submissions seems to suggest the student either got the problem correct more quickly or gave up on the problem earlier. For both females who did multiple contexts and those who did not, the number of submissions for the biology context is smaller compared with the other contexts. The high score on the biology contexts and the low number of submissions indicate that females had an easier time with the biology context questions than the other contexts.

Male and Female Achievement on the Force Concept Inventory

In order to investigate any differences between achievement on the Force Concept Inventory (FCI), scores of females pre (M = 15.95) and post (M = 18.2) were compared and found to be significantly different in their performance (p = 0.004; Cohen's d = 0.359). There were no significant differences in the pre (M = 20.9) and post (M = 20.5) scores for males (p = 0.298; Cohen's d = 0.055; see **Table 5**).

Next, the FCI scores for the females and males were analyzed using a two-sample t-test assuming unequal variances and summarized (**Table 5**). Males' pre-FCI scores (M = 20.6) were significantly higher than those of the females (M = 16.9) in the pre-FCI, but this gap was closed by the post-FCI (females 18.4; males 20.3; p = 0.155; Cohen's d = 0.354).

In order to further examine the differences between the FCI scores of males and females, these groups were subdivided into a "high" group (FCI scores of 21–30) and a "low" group (FCI scores 0–20) for males and females. The mean pre-FCI scores of the females (M = 13.4) and males (M = 14.4) in the "low" FCI group were not different (p = 0.292; Cohen's d = 0.21); nor were the mean scores of the low post-FCI group (Females: M = 16.5; Males: M = 14.2; p = 0.168; Cohen's d = 0.379). The mean pre-FCI scores of the females (M = 24) and the males (M = 24.9) in the "high" group were not different (p = 0.257; Cohen's d = 0.359); nor were the post-FCI scores of the high group (Females: M = 23.7; Males: M = 25.3; p = 0.065; Cohen's d = 0.656).

Next, the FCI scores of males and females whose FCI scores increased (growth groups) were analyzed in two ways. First, both the pre and post scores of growth males and growth females were compared. As can be seen in **Table 6**, males and females in this group were significantly different from each other, both pre- (p = 0.0115; Cohen's d = 0.985) and post-FI (p = 0.0114; Cohen's d = 0.977). Next, the pre- and post-FCI scores of females were compared, as were the pre- and post-FCI scores of males (see **Table 7**). There were significant increases in the growth for females (76% of females improved), from pre- to post-FCI (Female Pre: M = 14.6; Post: M = 18.1; p = 0.0466; Cohen's d = 0.613). When the pre- to post-FCI scores for the male growth group (35% of males improved) was examined, there were not significant differences (Male Pre: M = 20.9; Post: M = 23.7; p = 0.163; Cohen's d = 0.430).

Student Achievement and the Context of the Selected Physics Problems

For the final research question, the students' success on homework problems were analyzed based on the context of the homework problems they selected. Students could select the same context

TABLE 3 | Average percentage of responses correct and # submissions by category of video questions.

Gender	Traditional contexts video		Biology contexts video		Sports contexts video	
	% correct	# submissions	% correct	# submissions	% correct	# submissions
Females total	66.0	3.02	**69.8**	2.80	45.7	2.78
Males total	71.5	2.52	**73.0**	3.38	53.5	2.99
Females who chose a combination						
Video Qs	**72.6**	2.58	71.1	2.47	55.2	2.17
Word Qs	81.8	2.09	**86.8**	1.78	70.4	1.98
Males who chose a combination						
Video Qs	**71.0**	2.58	67.3	3.34	58.0	2.56
Word Qs	**81.0**	2.23	79.5	2.21	1.90	2.23
Females who did not choose a combination						
Video Qs	57.1	3.57	**73.5**	3.41	29.4	3.82
Word Qs	64.8	3.04	**92.6**	2.41	58.1	3.35
Males who did not choose a combination						
Video Qs	78.0	2.16	**88.2**	3.29	50.0	3.33
Word Qs	**92.4**	1.59	91.3	1.83	85.5	1.84

Note: Highest values for each group are bolded.

TABLE 4 | Average score of all responses and submissions by category.

Gender	Traditional contexts		Biology contexts		Sports contexts	
	% correct	# submissions	% correct	# submissions	% correct	# submissions
Females who chose a combination	73.7	2.24	**83.1**	1.88	69.6	2.23
Males who chose a combination	74.9	2.29	**76.3**	2.52	72.0	2.35
Females who did not choose a combination	59.9	3.17	**86.9**	2.65	52.7	3.63
Males who did not choose a combination	85.1	1.80	**91.3**	2.10	77.4	2.40

Note: Highest values for each group are bolded.

TABLE 5 | Pre/Post FCI comparisons within and between gender.

Group	Pre-FCI Avg	Post-FCI Avg	Δ Avg	t	df	*p* (one-tailed)
Pre/Post FCI comparisons within gender						
Female	15.95	18.2	2.25	−2.986	20	0.004**
Male	20.84	20.48	−0.36	0.535	30	0.298
Pre FCI comparisons female to male						
Female	15.95					
Male	20.84			−2.718	42	0.005**
Post FCI comparisons female to male						
Female		18.2				
Male		20.5		−1.266	47	0.106

$^{**}p \leq 0.01$.

TABLE 6 | Comparisons of FCI growth groups only FCI scores.

Group	Female	Male	T	df	*p* (one-tailed)
Pre FCI	14.6	20.9	−2.47	19	0.0115*
Post FCI	18.1	23.7	−2.47	20	0.0114*

$^{*}p \leq 0.05$.

TABLE 7 | Comparisons of pre and post FCI scores for growth groups only.

Group	Pre FCI	Post FCI	t	df	*p* (one-tailed)
Female	14.6	18.1	−1.73	30	0.0466*
Male	20.9	23.7	−1.01	20	0.163

$^{*}p \leq 0.05$.

more than once or select a different context from the choices of biology, sports, or traditional. The data in **Table 8** display the first, second, and third context choices of males and females, overall, the average number of correct responses, and the mean number of submissions. The means were compared using two-sample *t*-tests, assuming unequal variances. Females who chose multiple contexts per question had significantly more correct responses ($p = 0.013$; Cohen's $d = 0.441$) on their first choice and achieved that with significantly fewer submissions ($p \leq 0.001$; Cohen's $d = 0.885$) than male students who also chose to complete multiple contexts.

Next, females and males were divided by those who chose a combination of different contexts, and those who did not.

TABLE 8 | Average score of responses and submissions by gender.

	First choice		Second choice		Last choice (for 'all' category)	
Gender	**% correct**	**# submissions**	**% correct**	**# submissions**	**% correct**	**# submissions**
Female	85.1*	2.11	40.5	2.30	33.3	2.89
Male	69.5	3.22***	47.2	2.49	25.0	2.55

*$p \leq 0.05$; ***$p \leq 0.001$.*

TABLE 9 | Average score of responses and submissions by category.

	Traditional contexts		Biology contexts		Sports contexts	
Gender	**% correct**	**# submissions**	**% correct**	**# submissions**	**% correct**	**# submissions**
Females who chose a combination	73.7	2.24	**83.1**	1.88	69.6	2.23
Males who chose a combination	74.9	2.292	**76.3**	2.52	72.0	2.35
Females who did not choose a combination	59.9	3.17	**86.9**	2.65	52.7	3.63
Males who did not choose a combination	85.1	1.80	**91.3**	2.10	77.4	2.40

Note: Highest values for each group are bolded.

TABLE 10 | Average percentage of responses correct and # submissions by category of video questions.

	Traditional contexts video		Biology contexts video		Sports contexts video	
Gender	**% correct**	**# submissions**	**% correct**	**# submissions**	**% correct**	**# submissions**
Females total	66.0	3.02	**69.8**	2.80	45.7	2.78
Males total	69.6	2.52	**73.0**	3.38	53.5	2.99
Females who chose a combination	**72.6**	2.58	71.1	2.47	55.2	2.17
Males who chose a combination	**71.0**	2.58	67.3	3.34	58.0	2.56
Females who did not choose a combination	57.1	3.57	**73.5**	3.41	29.4	3.82
Males who did not choose a combination	78.0	2.16	**88.2**	3.29	50.0	3.33

Note: Highest values for each group are bolded.

As you can see in **Table 9**, all students were most likely to answer the biology context problem correctly, followed by the traditional context, and then the sports context. The mean number of submissions ("tries") on each problem indicated how many times a student attempted the problem. Female students had a lower number of submissions for the biology context. The high score on the biology contexts and the low number of submissions suggest that female students had an easier time with the biology context questions than the other contexts. To determine whether these differences were significant, the mean percentages of their correct questions (in **Table 9**) were compared using two-sample t-tests, assuming unequal variances.

Overall, students were most likely to choose a biology context, most likely to get that choice correct, and generally it took the fewest number of submissions to get those homework problems correct. Virtually all of these differences were found to be significantly different from one another. Females scored significantly better (85.1% correct) on their first choice than did males $p \leq 0.05$, (Cohen's $d = 0.441$) regardless of context. Females also had significantly fewer submissions (2.11) on their first choice than did males (3.22) $p \leq 0.001$, (Cohen's $d = 0.885$) regardless of context.

Results on the video and word problems were also investigated to see if there were differences in choice and success on these problems. Overall, males and females were most likely to get the biology context questions correct, although they were less successful on all of the video problems than they were on the word problems, with an average success rate for females 69.6% (2.8 attempts) and males 73.0% (3.38 attempts). For the video questions, the patterns for males and females were similar, although the males tended to make more attempts on problems, and the variation between nearly every choice group (combination/no combo), traditional, biology, sports, and male/female was significant (**Table 10**).

Student achievement on FCI scores were investigated in connection to the context choices made by females and males. Males who had FCI growth chose significantly more biology contexts than did males without growth (p = 0.040). Males who had no growth on the FCI chose significantly more traditional contexts than biology (p = 0.002). Females who showed FCI growth chose as many traditional contexts as biology and had more growth than those females who chose sports contexts. There were no differences on FCI scores based on the number of context choices females selected.

LIMITATIONS

Our findings need to be viewed in light of several limitations. First, we are unable to rule out all potential alternative explanations for the observed differences due to the nature of the research design. Our choice of outcomes and how we decided to measure them provides us with only a limited picture of what the students experienced during the WebAssign physics unit. Our findings, as a result, may have differed if we chose to target different learning outcomes. Second, the number of participants in this study was small and the nature of their experience during the intervention was unique and did not consider socioeconomic status nor students in lower-level courses. Additionally, the number of males and females also was unequal—typical of physics classrooms and part of the rationale for this study. The generalizability of our findings, therefore, might be limited to this single case. Third, we only analyzed one physics homework unit. Students had already learned about Newton's Laws, and they were in eight different classrooms across the US, which were not directly observed. Differences in time were unable to be clearly understood from the data in WebAssign, and we do not know whether students used additional help, although the problems were developed for this study and not available online. We therefore cannot make any claims about their direct experiences or other learning or interest in physics, nor whether these findings would be the case for a different population of students. With these limitations in mind, we will now discuss the findings of this study in light of the available literature.

DISCUSSION

Differences Between Females and Males in Context Selections

The context choices made by male and female students were analyzed to answer the first research question, "When given assignment choices designed around gender stereotypes, which types of physics problems do males and females select?" Females, overall, chose biology contexts more often than the other two contexts, traditional and sports. Male students, overall, were more likely to choose the traditional context over the other two contexts. Although no other studies have conducted a similar online intervention, nor in a physics problem set or providing the choice of a sports context, the findings are resonant with those of Baram-Tsabari and Yarden (2008). In their work, Baram-Tsabari and Yarden studied the interest of females and males, of various ages, toward biology and physics and similarly found that females were significantly less interested in traditional physics than the males, but females were significantly more interested in biology than male students. Baram-Tsabari and Yarden's (2008) study asked children, adolescents, and adults to create sets of self-generated questions that were classified according to interest in biology or physics.

An unexpected but interesting affordance of the WebAssign technology was that students could choose one or more contexts on additional attempts of the same problem. Although students were told at the beginning of the assignment that they would not receive any extra credit for completing extra assignments, most students chose to answer multiple contexts (and thus, complete extra problems). The software allowed the researcher to track the pattern of choices of the students. Those males and females who chose to do multiple contexts within a single question were placed into subcategories for analyses; females with combinations ("combos") and males with combos, regardless of how many multiple contexts they chose to do. Females reported they chose a certain context for two main reasons: because they were interested in the context (34% of the time) or because they thought the context presented looked easy (33% of the time). As a reminder, the students were asked to select the question context prior to seeing the actual physics question. That is, a student might have chosen a "cuddly kitten" (interest) or "the ball drop" (easy) prior to actually seeing and then attempting the problem, which they could view after this selection. However, students then reported their rationale for the choice after completing each problem.

The trend of context choices over time for females with combos shows that they had a constant rate of selecting biology contexts throughout the project, a slightly increasing rate of selecting traditional contexts, and that they chose sports contexts at a decreasing rate from the beginning of the project to the end. Females who chose to complete multiple contexts were more likely to choose to complete traditional-sports combinations or answer all contexts, followed by traditional-biology, and biology-sports combinations, which were selected at equal rates. Females who chose only one context tended to choose slightly more biology questions, fewer sports contexts, and chose traditional contexts at a relatively constant rate.

These context choices were explained by students with written comments such as, "It looked like something we did in class," "I like kittens," or "It looked easy." Female students were equally likely to write that they chose questions because they were interested in the context, or they thought the question was easy. Therefore, the given context of the question led to students' perceptions of these factors. Similar to this study, in which the researcher designed the choice options (biological, traditional, and sports), the Patall et al. (2010) study used teacher-determined written assignments given to high school students in chemistry, biology, and history classes, who chose which assignment to complete. Patall et al. (2010) found when students had a choice of homework, they had higher intrinsic motivation to complete the assignment and felt more competent doing the assignment. Given their continued choices, it seems likely that providing choices was perceived by students as motivating.

It seems that the students' perceptions about the choices that they made were consistent through the end of the problem set, even though several students correctly noted that the different contexts of each question were essentially the same problem, just presented in a different circumstance.

Video Questions

Compared with the written questions, both male and female students were more likely to choose to complete multiple contexts with video questions than they were word problems. The main reason why students chose to complete a video question was because of interest in the context or the "cool" factor. Students described the reasons they selected the video problems with responses such as "I enjoyed seeing the 5 second video and was motivated to actually complete the problem," "the graph made it easy to utilize the values in the problem," and "I found it most interesting."

Students were interested in the video analysis questions enough to try multiple contexts and they were interested in the written questions #3 (which dealt with tension in a taut line and included the choices of a spider and thread, mountain climber and rope, or elevator and cable), and #13 (which dealt with the normal force experienced when something hits the ground and included the choices of a rocket lift off, a cuddly kitten jump, or a standing high jump), seemingly because of the examples used in the contexts. Based on the response rate and student comments, the video format was more engaging to all students than the word format.

Female and Male Achievement Differences

To answer the question, "Are there differences in achievement of males and females and is this related to the types of physics questions they select?" achievement in this study was measured through FCI scores and WebAssign scores. First, FCI scores will be discussed. Females made significant gains on the post-FCI as compared with their pre-FCI, and they closed the gender gap with males by the end of the problem set. Female students tended to choose biology contexts more than male students and most female students chose to complete a second context. Females did better on the biology contexts than any other context in either word or video format. Overall, students were limited to the context choice, the time they spent on the problem, and the number of submissions per problem as the only variables that they could manipulate in this study. McCullough (2004) argues that the FCI is dominated by questions that align with stereotypical male contexts. In the findings of this study, females not only significantly improved from their pre- to their post-FCI score, but their post score was statistically equal to the males' post score. Given these findings and the research of McCullough (2004), it is possible that the intervention may have been even more successful at improving female students' conceptual understanding than first thought.

In this study, all of the choice options afforded to the students were equally difficult. This has not necessarily been the case with other studies in the literature, which have also found that giving limited choice in the type of assignments or homework students can result in greater gains on assessments. Fulton and Schweitzer (2011) found that giving students a choice in the type of final class assignment (one defined as easier, the other more challenging) in a computer science class resulted in lower performing students choosing an easier assignment option and therefore learning less than their peers. However, similar to this study, in a meta-study literature analysis on the effects of choice in a variety of settings, Katz and Assor (2007) found that choice motivates students when the choices are limited and aligned with student interests and goals.

In contrast to the female students, males did not show the same kinds of gains and had no significant differences between their pre- and post-FCI scores, even though males had the highest scores on the traditional context questions of the problem set. The intervention appeared to have had a positive effect on female students' performance on the FCI but not the males' performance. Why might this have happened? Even more so than the females, males were more likely to choose a question to work on because it looked interesting to them (45%). One difference from the females was that males who chose to do multiple contexts tended to choose traditional contexts at a constant rate throughout the project, and they chose fewer sports contexts and slightly more biology contexts throughout the project. Indeed, males who chose to complete only one context tended to choose more traditional contexts from the beginning to the end of the homework set. Males selected slightly fewer sports contexts across the project but chose biology contexts at a constant rate. The most common combination of contexts were biology-sports, then all three contexts chosen together, followed by traditional-biology, and traditional-sports.

The effect of the choice of context on female students' FCI growth can be seen in the feedback loop of Artino's (2010) framework, in which positive emotions such as interest or enjoyment resulted in positive academic outcomes such as achievement and conceptual understanding. For males, the feedback was that they were more likely to get the traditional context problems correct compared with the other contexts, which may have resulted in more confidence in doing the problems and more satisfaction which reinforced their choice of traditional contexts. The homework questions were open response, not multiple choice, so the students needed to complete the problems to get credit instead of just checking a box. It is expected that the more time a student spends on an assignment or the more problems they complete, the more their understanding will grow from exposure to the content.

There was not a significant difference between the pre- and post-FCI scores of males who chose to do combos versus males who chose only one context; however, males who chose one context had a higher absolute score on both measures. Indeed, traditional physics courses have privileged male students (Baram-Tsabari and Yarden, 2008), and the traditional choice also was the one more likely to correspond to an example given in the classes of the teachers who helped with this study (as explained in the Methods section).

Students who chose combos, regardless of gender, tended to do better on their first-choice question context. Interest in the context seemed to be more of a deciding factor for initial choice rather than the perceived difficulty level (e.g., choosing a problem because it seemed "easier"). Even though the questions were essentially the same, except for the context, students reported that one context seemed easier than another, suggesting that interest impacted perceptions of difficulty and perhaps making the problem seem more relevant, and therefore more manageable.

Homework Question Format: Word and Video

The homework problems had an additional "wrinkle." They were either word problems or video problems. The video questions were novel from what the students had seen in their classes (based on communication with the teachers), but they also required students to gather data from the video in order to try to solve the problem. Therefore, there were some additional and unfamiliar steps involved in trying to solve those problems. Significant differences were found between the performance of males on the video questions and the word problems. Males, regardless of multiple context choice or not, performed better and used fewer submissions on the word problems compared to the video problems for each context (biological, traditional, or sports). Although males indicated that they enjoyed the video questions more than the word questions, this did not lead to better performance on those problems. Males who selected multiple contexts scored significantly higher on the biology video questions than did males who chose only one context.

Results from the female students on the video questions were not as clear. Overall, females who completed combos tended to have higher mean FCI scores and higher scores on traditional and sports questions than female students who did not choose a combo, although these differences were not significant. Females who did not choose a combo tended to have higher scores on the biology contexts compared with females who chose to complete multiple contexts.

CONCLUSION

This study adapted research regarding student interests and choice into an online physics assignment designed to investigate their possible role in the attitudes, understanding, and achievement of male and female students. A number of conclusions can be drawn from the findings of this study. From the findings, we see that when given problem choices designed around gender stereotypes, females were more likely to choose questions related to biology contexts, while males were more likely to select traditional physics contexts. Females were more likely to indicate that they selected a problem because they thought it looked easy, while males indicated they selected a problem because they were interested in it. When offered choice, many males and females chose to complete multiple contexts within a problem. Females and males who chose multiple contexts had higher scores on the biology problems than the other contexts. There were significant improvements in female post-FCI scores as compared with pre-FCI scores but no difference between pre- and post-FCI scores were found for the males.

We investigated whether there was a relationship between student achievement and the context of physics problems they selected. First, students made many choices of different contexts, suggesting that they were motivated by the ability to choose. The students who were more likely to take advantage of these combinations of problems were those who selected biology contexts over traditional contexts. Second, males and females were both motivated by interest, but this reason was more likely to be stated by males. Third, students persisted in problems, as shown by the number of submissions made on each problem, that were more challenging or with which they were less successful when they were motivated by interest or novelty, as they were in the video questions. Fourth, the intervention led more females than males to improve their scores and erased the initial gender gap seen on the FCI scores, pre-intervention. Fifth, even a short-term, online homework intervention—with no professional development on the part of the teacher—can positively impact students' engagement, achievement, and motivation in physics.

IMPLICATIONS

The development of the WebAssign intervention is one that could be used by other teachers to try to better engage students, particularly female students, in physics. Given the research design, it was not possible to tease out the effects of choice and the combination of choices of biology, sports, and traditional options. A different design that separates out these variables is needed to know specifically what was most motivating for the students. The use of video analysis questions is an area for future research that wasn't fully explored in this study and how students engage with the problems and its potential needs further investigation. It certainly seems, given the results, that some experience navigating with data gathering in video problems would enhance student achievement on these problems, and the potential for making physics more interesting with video questions ought to be explored. We also want to understand more about why the females had greater gains than the males, despite high interest by the males. We wonder how much could be achieved if this sort of homework set could be used throughout the year, and biology examples were used in class, in addition to traditional examples. We believe that our findings give some promising insights into closing the gender gap in the achievement high school physics students.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The Internal Review Board (IRB) at the university (NCSU) approved the proposed study on February 10th, 2016, Protocol #6552. Once IRB approved this study, a letter was sent out to each school's administration and cooperating teacher. Permission for the research to be conducted in the school was sought from principals and any administration at the district level that was needed. Students were given a consent form and a letter to take home to their parents explaining the research and the student's role in the research inviting them to participate. The letter consisted of the project description and an invitation for both students and parents. The signed written consent form was returned by the student to the teacher. The cooperating teacher returned all the signed forms to the first author. Teachers were also asked to complete a signed consent forms for their part in the study and these were returned to the lead researcher with the student forms. Teachers were provided with self-addressed stamped envelopes to return all the materials. Students were asked to complete a pre and post survey which could have made them feel uncomfortable, but the results of the surveys were collected online and no one else could see the students' responses except the researchers. Academic risk was minimized and potential stress from the automatic grading done by WebAssign was mitigated by giving students five chances to get each question correct. Students were also given the option of getting an extension on the assignment if they wanted more time. These measures were taken to avoid any potential risk or stress due to the academic nature of the assignment. The scores students received on these assignments weren't part of their class grade.

AUTHOR CONTRIBUTIONS

Both authors contributed to this research and are accountable for the content of this work.

FUNDING

The NCSU Physics Department and Dr. Robert Beichner provided funding to compensate students and teachers who participated in this study. WebAssign provided free access to the participants in the study. The NCSU Department of STEM Education provided publication support.

ACKNOWLEDGMENTS

We want to thank the North Carolina School of Science and Math for providing software used in the video editing and creating process. WebAssign graciously provided free student and instructor accounts for use in this study. Thank you again to the NCSU Physics Department and Dr. Robert Beichner for their generosity and support. Finally, thank you to all the teachers and students who participated in this study. Without their help, this study would not have been possible.

REFERENCES

Artino, A. R. (2010). Online or face-to-face learning? Exploring the personal factors that predict students' choice of instructional format. *Internet High. Educ.* 13, 272–276. doi: 10.1016/j.iheduc.2010.07.005

American Physical Society (2015). Fraction of bachelor degrees in STEM disciplines earned by women 1965–2015. [Web log post]. Available at: http://www.aps.org/programs/education/statistics/womenstem.cfm (Accessed February 17, 2019).

Badri, M., Mazroui, K. A., Al Rashedi, A., and Yang, G. (2016). Variation by gender in Abu Dhabi high school students' interests in physics. *J. Sci. Educ. Technol.* 25, 232–243. doi: 10.1007/s10956-015-9589-x

Baker, D. R., and Pilburn, M. D. (1997). *Constructing science in middle and secondary school classrooms*. (Boston, MA: Ally and Bacon).

Baram-Tsabari, A., and Yarden, A. (2008). Girls' biology, boys' physics: evidence from free-choice science learning settings. *Res. Sci. Technol. Educ.* 26, 75–92. doi: 10.1080/02635140701847538

Baram-Tsabari, A., and Yarden, A. (2011). Quantifying the gender gap in science interests. *Int. J. Sci. Math. Educ.* 9, 523–550. doi: 10.1007/s10763-010-9194-7

Beichner, R., DeMarco, M., Ettestad, D., and Gleason, E. (1989). "VideoGraph: a new way to study kinematics" in *Proceedings of the conference on computers in physics instruction*. eds. E. F. Redish and J. S. Risley (Redwood City, CA, USA: Addison Wesley Publishing Company), 244–245.

Beichner, R. J., and Abbott, D. S. (1999). Video-based labs for introductory physics courses. *J. Coll. Sci. Teach.* 29, 101–104.

Bøe, M. V., and Henriksen, E. K. (2013). Love it or leave it: Norwegian students' motivations and expectations for postcompulsory physics. *Sci. Educ.* 97, 550–573. doi: 10.1002/sce.21068

Bryan, R. R., Glynn, S. M., and Kittleson, J. M. (2011). Motivation, achievement, and advanced placement intent of high school students learning science. *Sci. Educ.* 95, 1049–1065. doi: 10.1002/sce.20462

Buabeng, I., Ampiah, J. G., and Quarcoo-Nelson, R. (2012). Senior high school female students' interest in physics as a course of study at the university level in Ghana. *IFE PsychologIA* 20, 369–379.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., and Hanson, W. E. (2003). "Advanced mixed methods research designs" in *Handbook of mixed methods in the behavioral and social sciences*. eds. A. Tashakkori and C. Teddlie (Thousand Oaks, CA: Sage), 209–240.

Crouch, C. H., and Heller, K. (2014). Introductory physics in biological context: An approach to improve introductory physics for life science students. *Am. J. Phys.* 82, 378–386. doi: 10.1119/1.4870079

Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153

Evans, M., and Boucher, A. R. (2015). Optimizing the power of choice: Supporting student autonomy to foster motivation and engagement in learning. *Mind Brain Educ.* 9, 87–91. doi: 10.1111/mbe.12073

Fulton, S. and Schweitzer, D. (2011). Impact of Giving Students a Choice of Homework Assignments in an Introductory Computer Science Class. *Int. J. for the Scholarship of Teach. and Learning* 5. doi: 10.20429/ijsotl.2011.050120

Gibson, A., Cook, E., and Newing, A. (2006). Teaching medical physics. *Phys. Educ.* 41, 301–306. doi: 10.1088/0031-9120/41/4/001

Gibson, V., Jardine-Wright, L., and Bateman, E. (2015). An investigation into the impact of question structure on the performance of first year physics undergraduate students at the University of Cambridge. *Eur. J. Phys.* 36:045014. doi: 10.1088/0143-0807/36/4/045014

Grossman, J. M., and Porche, M. V. (2014). Perceived gender and racial/ethnic barriers to STEM success. *Urban Educ.* 49, 698–727. doi: 10.1177/0042085913481364

Hazari, Z., Sonnert, G., Sadler, P. M., and Shanahan, M. C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: a gender study. *J. Res. Sci. Teach.* 47, 978–1003. doi: 10.1002/tea.20363

Hegarty, M. (2014). "Multimedia learning and the development of mental models" in *The Cambridge Handbook of Multimedia Learning*. ed. R. Mayer (Cambridge: Cambridge University Press), 673–701.

Heilbronner, N. N. (2013). The STEM pathway for women: what has changed? *Gift. Child Q.* 57, 39–55. doi: 10.1177/0016986212460085

Hestenes, D., Wells, M., and Swackhamer, G. (1992). Force concept inventory. *Phys. Teach.* 30, 141–158. doi: 10.1119/1.2343497

Hoskinson, A. M., Couch, B. A., Zwickl, B. M., Hinko, K. A., and Caballero, M. D. (2014). Bridging physics and biology teaching through modeling. *Am. J. Phys.* 82, 434–441. doi: 10.1119/1.4870502

Jian, Y. C., and Wu, C. J. (2015). Using eye tracking to investigate semantic and spatial representations of scientific diagrams during text-diagram integration. *J. Sci. Educ. Technol.* 24, 43–55. doi: 10.1007/s10956-014-9519-3

Johnson, A. T. (2012). The trouble with STEM. *Issues Sci. Technol.* 29, 16–16.

Jugović, I. (2017). Students' gender-related choices and achievement in physics. *Cent. Edu. Policy Studies J.* 7, 71–95.

Katz, I., and Assor, A. (2007). When choice motivates and when it does not. *Educ. Psychol. Rev.* 19, 429–442. doi: 10.1007/s10648-006-9027-y

Keller, M. M., Neumann, K., and Fischer, H. E. (2017). The impact of physics teachers' pedagogical content knowledge and motivation on students' achievement and interest. *J. Res. Sci. Teach.* 54, 586–614. doi: 10.1002/tea.21378

Lee, K. J., and Sharma, M. D. (2008). Incorporating active learning with videos: a case study from physics. *Teach. Sci.* 54, 45–47.

Makarova, E., and Herzog, W. (2015). Trapped in the gender stereotype? The image of science among secondary school students and teachers. *Equal. Divers. Inclusion* 34, 106–123. doi: 10.1108/EDI-11-2013-0097

Mason, L., Pluchino, P., Tornatora, M. C., and Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *J. Exp. Educ.* 81, 356–384. doi: 10.1080/00220973.2012.727885

Mayer, R. E., and Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *J. Educ. Psychol.* 86, 389–401. doi: 10.1037/0022-0663.86.3.389

Mayer, R. E., Steinhoff, K., Bower, G., and Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educ. Technol. Res. Dev.* 43, 31–43. doi: 10.1007/BF02300480

Marchand, G. C., and Taasoobshirazi, G. (2012). Stereotype Threat and Women's Performance in Physics. *Int. J. Sci. Educ.* 35, 3050–3061. doi: 10.1080/09500693.2012.683461

McCullough, L. (2004). Gender, context, and physics assessment. *J. Int. Women's Stud.* 5, 20–30. http://vc.bridgew.edu/jiws/vol5/iss4/2

McCullough, L. (2007). Gender in the physics classroom. *Phys. Teach.* 45, 316–317. doi: 10.1119/1.2731286

Mitrevski, J., and Treagust, D. (2011). Girls and upper school physics: Some optimism and opportunity. *Teach. Sci.* 57, 35–40.

Mulvey, P. J., and Nicolson, S. (2014). *Trends in exiting physics masters*. AIP Focus On, March. Retrieved from https://www.aip.org/sites/default/files/statistics/graduate/trendsmasters-p-12.2.pdf#page=7

Murphy, P., and Whitelegg, E. (2006). Girls and physics: continuing barriers to 'belonging. *Curric. J.* 17, 281–305. doi: 10.1080/09585170600909753

NRC (2013). *Adapting to a changing world—challenges and opportunities in undergraduate physics education*. (Washington, DC: The National Academies Press).

Oliver, M. C., Woods-McConney, A., Maor, D., and McConney, A. (2017). Female senior secondary physics students' engagement in science: a qualitative study of constructive influences. *Int. J. STEM Educ.* 4, 1–15. doi: 10.1186/s40594-017-0060-9

Oon, P. T., and Subramaniam, R. (2010). On the declining interest in physics among students—from the perspective of teachers. *Int. J. Sci. Educ.* 33, 727–746. doi: 10.1080/09500693.2010.500338

Patall, E. A. (2013). Constructing motivation through choice, interest, and interestingness. *J. Educ. Psychol.* 105, 522–534. doi: 10.1037/a0030307

Patall, E. A., Cooper, H., and Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychol. Bull.* 134, 270–300. doi: 10.1037/0033-2909.134.2.270

Patall, E. A., Cooper, H., and Wynn, S. R. (2010). The effectiveness and relative importance of choice in the classroom. *J. Educ. Psychol.* 102, 896–915. doi: 10.1037/a0019545

Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341. doi: 10.1007/s10648-006-9029-9

Phillips, P. P. (2015). The impact of e-education on at risk high school students' science achievement and experiences during summer school credit recovery courses. Doctoral dissertation. Retrieved from: https://repository.lib.ncsu.edu/handle/1840.16/10365

Reid, N., and Skryabina, E. A. (2003). Gender and physics. *Int. J. Sci. Educ.* 25, 509–536. doi: 10.1080/0950069022000017270

Sawtelle, V., Brewe, E., and Kramer, L. H. (2012). Exploring the relationship between self-efficacy and retention in introductory physics. *J. Res. Sci. Teach.* 49, 1096–1121. doi: 10.1002/tea.21050

Scheiter, K., Wiebe, E. N., and Holsanova, J. (2009). "Theoretical and instructional aspects of learning with visualizations" in *Cognitive effects of multimedia learning*. ed. R. Zheng (Hershey, PA: IGl Global/IDEA Publishing).

Schnotz, W. (2014). "Integrated model of text and picture comprehension" in *The Cambridge handbook of multimedia learning*. ed. R. Mayer (Cambridge: Cambridge University Press), 72–103.

Struck, W., and Yerrick, R. (2010). The effect of data acquisition-probeware and digital video analysis on accurate graphical representation of kinetics in a high school physics class. *J. Sci. Educ. Technol.* 19, 199–211. doi: 10.1007/s10956-009-9194-y

Talley, K. G., and Martinez Ortiz, A. (2017). Women's interest development and motivations to persist as college students in STEM: a mixed methods analysis of views and voices from a Hispanic-Serving Institution. *Int. J. STEM Educ.* 4, 1–24. doi: 10.1186/s40594-017-0059-2

Wang, M. T., Chow, A., Degol, J. L., and Eccles, J. S. (2017). Does everyone's motivational beliefs about physical science decline in secondary school?: Heterogeneity of adolescents' achievement motivation trajectories in physics and chemistry. *J. Youth Adolesc.* 46, 1821–1838. doi: 10.1007/s10964-016-0620-1

Wheeler, S. R. (2017). Using Choice to Uncover the Role of Gender Stereotypes in High School Physics Assignments: Examining Students' Interests, Beliefs, Conceptual Understanding and Motivations. (Doctoral Dissertation). North Carolina State University. Retrieved from: http://www.lib.ncsu.edu/resolver/1840.20/34829

Wieman, C. (2012). Applying new research to improve science education. *Issues Sci. Technol.* 29, 25–32.

Zohar, A., and Sela, D. (2003). Her physics, his physics: gender issues in Israeli advanced placement physics classes. *Int. J. Sci. Educ.* 25, 245–268. doi: 10.1080/09500690210126766

Zollman, D. A., and Fuller, R. G. (1994). Teaching and learning physics with interactive video. *Phys. Today* 47, 41–47. doi: 10.1063/1.881428

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Sociology

ORIGINAL RESEARCH
published: 12 April 2019
doi: 10.3389/fsoc.2019.00026

STEM/Non-STEM Divide Structures Undergraduate Beliefs About Gender and Talent in Academia

Kimberlyn A. Bailey[1*†], *David Horacek*[1], *Steven Worthington*[2], *Ampalavanar Nanthakumar*[3], *Scott Preston*[3] and *Carolina C. Ilie*[4*]

[1] Department of Philosophy, State University of New York at Oswego, Oswego, NY, United States, [2] Institute for Quantitative Social Science, Harvard University, Cambridge, MA, United States, [3] Department of Mathematics, State University of New York at Oswego, Oswego, NY, United States, [4] Department of Physics, State University of New York at Oswego, Oswego, NY, United States

OPEN ACCESS

Edited by:
Rebecca Lazarides,
Universität Potsdam, Germany

Reviewed by:
Nicole Farris,
Texas A&M University Commerce,
United States
Eva Cifre,
University of Jaume I, Spain

***Correspondence:**
Kimberlyn A. Bailey
kimberlyn.a.bailey@gmail.com
Carolina C. Ilie
carolina.ilie@oswego.edu

†Present Address:
Kimberlyn A. Bailey,
Technical University of Dresden,
Dresden, Germany

Specialty section:
This article was submitted to
Gender, Sex and Sexuality Studies,
a section of the journal
Frontiers in Sociology

Received: 15 August 2018
Accepted: 19 March 2019
Published: 12 April 2019

Citation:
Bailey KA, Horacek D, Worthington S, Nanthakumar A, Preston S and Ilie CC (2019) STEM/Non-STEM Divide Structures Undergraduate Beliefs About Gender and Talent in Academia. Front. Sociol. 4:26. doi: 10.3389/fsoc.2019.00026

Research and popular debate on female underrepresentation in academia has focused on STEM fields. But recent work has offered a unifying explanation for gender representation across the STEM/non-STEM divide. This proposed explanation, called the field-specific ability beliefs (FAB) hypothesis, postulates that, in combination with pervasive stereotypes that link men but not women with intellectual talent, academics perpetuate female underrepresentation by transmitting to students in earlier stages of education their beliefs about how much intellectual talent is required for success in each academic field. This theory was supported by a nationwide survey of U.S. academics that showed both STEM and non-STEM fields with fewer women are also the fields that academics believe require more brilliance. We test this top-down schema with a nationwide survey of U.S. undergraduates, assessing the extent to which undergraduate beliefs about talent in academia mirror those of academics. We find no evidence that academics transmit their beliefs to undergraduates. We also use a second survey "identical to the first but with each field's gender ratio provided as added information" to explicitly test the relationship between undergraduate beliefs about gender and talent in academia. The results for this second survey suggest that the extent to which undergraduates rate brilliance as essential to success in an academic field is highly sensitive to this added information for non-STEM fields, but not STEM fields. Overall, our study offers evidence that, contrary to FAB hypothesis, the STEM/non-STEM divide principally shapes undergraduate beliefs about both gender and talent in academia.

Keywords: gender bias, women in science, underrepresentation of women, talent, STEM/non-STEM divide, gender stereotypes, STEM

1. INTRODUCTION

Established stereotypes linking men but not women with innate brilliance may hinder women's paths into academia (Bennett, 1996; Tiedemann, 2000; Kirkcaldy et al., 2007; Lecklider, 2013; Leslie et al., 2015). Indeed, a 2015 study by Leslie et al. found that academic disciplines whose members highly value unteachable talent have gender ratios skewed toward men at the doctoral level. Using a nationwide survey of postdoctoral researchers, faculty and graduate students (henceforth: academics) in 30 disciplines, the authors found that the disciplines with the fewest women had practitioners who most strongly considered talent essential to success in their field (Leslie et al., 2015). Based on these findings, Leslie et al. proposed a theory for female underrepresentation in academia: the field-specific ability beliefs hypothesis. A field-specific ability belief (FAB) is the extent to which one believes that success in a given academic field requires talent. The authors propose that FABs are passed down from academics, saturating the general public, and combine with stereotypes about women's intellect to create and perpetuate the academic gender gap.

Leslie et al. build on prior research that suggests people vary in the extent to which they believe unteachable, fixed talent is essential to success in any activity. The FAB hypothesis is grounded in Carol Dweck's work on the "growth" vs. "fixed" mindset. Dweck's work suggests that individuals may be placed on a spectrum, where, on one end, an individual believes talent is innate ("fixed") and, on the other end, an individual believes talent may be cultivated through effort ("growth") (Dweck, 2006). The FAB hypothesis builds on Dweck's distinction, proposing that, rather than focusing on placing people along that "growth" vs. "fixed" spectrum, entire academic fields may be placed along that same spectrum. On this spectrum—so Leslie et al. suggest—success in some academic fields is widely believed to require "fixed" unteachable intellectual talent, while success in other academic fields is widely believed to require hard work (Leslie et al., 2015). They propose that the degree to which academics believe success in a given academic field requires fixed talent – their FABs – strongly influences the extent to which the wider public believes a given academic field requires fixed talent.

The FAB hypothesis also includes specific claims about how FABs play into causal mechanisms responsible for the gender gap. Key to understanding their proposed mechanism is that field-specific ability beliefs are a metric of the extent to which an academic field is believed to require fixed, unteachable brilliance, as opposed to talent that can be developed through hard work. If a field is indeed believed to require unteachable brilliance, then, naturally, success in that field becomes viewed as an insurmountable challenge for anyone who feels he or she lacks that innate intellectual spark. The FAB hypothesis posits, first, that academics hold negative stereotypes about women's innate intellectual ability and thus exhibit biases against them in high-FAB fields (Valian, 1999). The FAB hypothesis posits, second, that women internalize these stereotypes about themselves and/or believe that such pervasive stereotypes render high-FAB fields inhospitable to them and, as a result, decide to not pursue high-FAB fields (Wigfield and Eccles, 2000; Dar-Nimrod and Heine, 2006).

For such mechanisms to explain the gender gap, young women making decisions about entering academia must be able to reliably identify high-FAB "brilliance required" fields as such. This is an explicit prediction of what, in a later follow-up study by the same research group, is called the "extended FAB hypothesis" (Meyer et al., 2015). Within this extended framework, the authors posit that those with exposure to a given academic field will "absorb" the FABs of the practitioners of that discipline (Meyer et al., 2015). It is this tenet that we aim to test. If true, then we should expect undergraduate FABs to strongly reflect those of academics or come to do so as undergraduates spend more time in college. The researchers propose that their "extended FAB hypothesis" includes the general public, such that the public at large holds FABs like those of academics. Using another survey-based study, they found this to be the case and, further, that those with college-level exposure to a field have FABs that more closely reflect those of academics than those without (Meyer et al., 2015). Insofar as undergraduates have constant and direct exposure to academics, they are the population that seems most likely to "absorb" the FABs of academics. Indeed, studies show that the largest drop-off of women in male-dominated fields happens at the undergraduate level (Ceci et al., 2014). It is also predominantly undergraduates who face the choice about whether to attend graduate school, and if they decide to go, which field to enter. If the FAB hypothesis is to explain their academic choices at this crucial juncture, and the gender gap more generally, then undergraduate FABs must become aligned with those of the academics with whom they are in constant contact. We test this top-down inheritance of beliefs using a national survey of undergraduates that mirrors the one used by Leslie et al. to estimate undergraduate FABs and compare these to the FAB scores of academics collected by Leslie et al.

We test an additional key prediction of the FAB hypothesis about the similarity of undergraduate and academic FABs. Current research about the gender gap in academia has largely focused on women's underrepresentation in STEM (Ceci and Williams, 2007). Yet there is considerable variation in female representation on both sides of the STEM/non-STEM divide (Ceci and Williams, 2007, 2011). Less than 20% of all physics Ph.D.'s are awarded to women, while neuroscience programs award around 50% (National Science Foundation, 2011). Similarly, women currently earn more than 70% of all Ph.D.'s in art history, but less than 35% in philosophy (National Science Foundation, 2011). Leslie et al.'s FAB hypothesis is novel in that it offers a unified explanation for variation in female representation across the STEM/non-STEM divide. Indeed, the FABs of academics were predictive of female representation in academia both across the STEM/non-STEM divide and within these subsets of fields. Thus, the FAB hypothesis postulates that undergraduate FABs should be predictive of female representation not only within STEM fields, but also non-STEM fields and across both subsets of fields combined.

As a second part of our study, we also seek to explicitly explore the relationship between undergraduate beliefs about gender and talent. We also use a second survey that, like the first, asks undergraduates to rate the extent to which they believe a given academic field requires talent, with one key difference: The gender ratio for each academic field is given as added information.

Overall, our study serves as a test of some key predictions of the FAB hypothesis about undergraduate beliefs and, more generally, as a study into undergraduate beliefs about gender and talent in academia.

1.1. Summary of Predictions

We sought to test several key predictions of the FAB hypothesis using survey version one (respondents not given the female representation of each field as added information). We thus formed our hypotheses for survey one in accordance to what the FAB hypothesis would predict. In this section we briefly summarize our hypotheses, with more details furnished in the introduction and discussion sections. Just as Leslie et al.'s original study, we used the percent of Ph.D.'s awarded in 2011 within the US to females and African Americans as our metric for female and African American representation in academia, respectively (National Science Foundation, 2011).

The FAB hypothesis claims that the FABs of undergraduates are strongly influenced by the FABs of academics and, thus, the hypothesis predicts that undergraduate FABs, like those of academics, predict female representation. We thus hypothesized that, (1) as Leslie et al. found for academic FABs, there would be an association between average undergraduate FABs for each field and female representation for each field.

The FAB hypothesis claims that FABs play an important explanatory role for female representation in academia across all academic fields. Thus, the FAB hypothesis posits that FABs are predictive of female representation not only for STEM fields, but also non-STEM fields. We hypothesized that, as Leslie et al. found for academic FABs, (2) there would be an association between average undergraduate FABs and female representation for STEM fields alone, non-STEM fields alone and across all fields and (3) undergraduate FABs would remain an important predictor of female representation, even when a field's classification as STEM or non-STEM is taken into account.

The FAB hypothesis claims that FABs, as beliefs about how much innate talent is required for a given academic field, influence representation in academia of any group stereotyped as lacking innate intellectual talent. African Americans are one such group (Steele and Aronson, 1995). We thus hypothesized that, (4) as Leslie et al. found for academic FABs, there would be an association between undergraduate FABs and African American representation.

The FAB hypothesis claims that the FABs of academics influence the FABs of the public at large, such that those with more exposure to academics develop FABs that more closely resemble those of academics. Thus, the FAB hypothesis predicts that undergraduates—who have direct and constant exposure to academics—will have FABs that resemble those of academics or come to do so with more exposure to academics. We thus hypothesized that (5) undergraduate FABs would differ between undergraduate class years (Freshman, Sophomore, Junior or Senior) as a result of differing degrees of exposure to academics.

With survey version two (respondents given the female representation of each academic field as added information), we sought to generally probe the relationship between undergraduate beliefs about gender and talent in academia and to test a specific implication of the FAB hypothesis. The FAB hypothesis builds off of work that suggests are stereotyped as having less innate intellectual talent than men. Informing undergraduates of the gender ratio in each academic field could thus be expected to trigger those stereotypes and strongly influence respondent FABs. The FAB hypothesis thus predicts that (6) respondents provided the gender ratio in each academic field should give FABs that, relative to undergraduates not provided that added information, rank male-dominated fields as more "brilliance-required" fields (higher FABs) and female-dominated fields as less "brilliance-required" (lower FABs).

2. METHODS AND MATERIALS

2.1. Survey Respondents and Administration

The FAB scores of academics—graduate students, postdoctoral students and faculty—were used from Leslie et al.'s original study (Leslie et al., 2015). To compare the FABs of academics and undergraduates, we surveyed undergraduates in this study.

Participants were composed of 1075 U.S. undergraduates from 197 universities across the Country, ranging from the Ivy League to community colleges. The surveys were approved by the Human Subjects Committee of State University of New York (SUNY) at Oswego. Data was excluded from an additional 586 individuals who had identified as living outside the US, not currently matriculated in an undergraduate university or who had failed to complete the survey within, at most, a few missing answers.

The majority of participants came from Le Moyne College in Syracuse, New York (24%) and SUNY Oswego (53%). The other 33% of participants were undergraduates who learned of the survey through Tumblr, a blogging site popular with undergraduates. We administered surveys to SUNY Oswego students during the spring 2015 semester. The classes in which surveys were administered covered a wide range of subjects from both STEM and non-STEM fields. While most SUNY Oswego students took a printed version survey during a class session, several larger classes at SUNY Oswego took an online version, via a link sent by their professors. The online version mirrored the paper version and was created using the website Survey Monkey. Le Moyne College students were prompted to complete the online version, linked in a school-wide email. Le Moyne students and Tumblr readers were incentivized to take the survey with entrance into a drawing for an $80 Apple gift card. A popular blogger on Tumblr initially posted an invitation to take the survey, and other bloggers re-posted or "reblogged" the original post, spreading the invitation to a wider audience. To ensure only the intended participants took the survey, the online version had an added question at the start of the survey: Participants were asked to confirm that they are currently enrolled undergraduates at an American university or college. Those who answered negatively were thanked for their time on the following screen and were not advanced to the next portion of the survey.

Prior to starting the survey, each respondent was informed that the purpose of the study is to examine undergraduate attitudes about academic fields. Additionally, respondents were informed that the study was approved by the Human Subjects Committee of SUNY Oswego and were asked for their participation for the sake of advancing social science.

2.2. Academic Fields

A total of 42 academic fields were included in the surveys. We used three criteria in choosing these fields: overlap with the fields included in the survey used by Leslie et al., the relative size of each field (by Ph.D.'s granted in 2011 National Science Foundation, 2011) and whether an average undergraduate could be expected to meaningfully distinguish the fields from one another. We kept 29 of the 30 fields used by Leslie et. al. rejecting only Comparative Literature on the grounds that the average undergraduate could not be expected to find it meaningfully distinguishable from English Literature, a field which we retained from the group of fields used in Leslie et al.'s survey.

We added an additional 13 fields from a diverse set of disciplines within both STEM and non-STEM fields, with broad variations in female representation. To ensure a relatively objective method for choosing fields and increase the likelihood that undergraduates would be familiar with the chosen fields, all of the 13 added fields were amongst those that produced the largest number of Ph.D.'s in 2011 (National Science Foundation, 2011). We did not include some of those fields on distinguishability grounds. For example, Curriculum and Instruction was excluded due to anticipated difficulties for the average undergraduate to meaningfully distinguish it from Education, which was already included from the Leslie et al.'s group of fields.

2.3. Field-Specific Ability Beliefs Survey Questions

To assess field-specific ability beliefs, participants were asked to rate their agreement with a statement, for each of the 42 fields. The statement was taken directly from Leslie et al.'s original survey of academics (Leslie et al., 2015) and reads, "Being a top scholar in this discipline requires a special aptitude that just can't be taught." Participants rated their agreement on a 10-point scale, which was then converted to a seven-point scale, to match Leslie et al.'s scale.

Participants were randomly assigned to take one of two versions of the survey. In the first version, respondents rated their agreement with the above statement. In the second version of the survey, participants were asked to do the same, but were provided additional information: below the name of each field appeared a percentage bar representing the percent female representation within that field. The same measure used for female representation throughout our analyses was used in the survey: the percent of female Ph.D. recipients for a field in 2011 (National Science Foundation, 2011).

Respondents were told what the percentage bars represent and the source of data. Of all surveys retained for analysis, each of the two surveys was taken by roughly half of participants: 546 participants (40% male, 47% STEM majors) and 518 (37% male, 24% STEM majors) participants took survey version one and two, respectively. Respondents who took the online version of the survey were randomly assigned a survey version. For respondents who took an in-class survey, each class was administered the same survey, but which classes received which version was randomly assigned.

2.4. Education and Demographic Questions

For both versions of the survey, participants were asked a number of questions pertaining to their educational and demographic characteristics. They were asked to report their major(s), future plans after graduation (or to indicate that they did not know, if such was the case), grade point average (GPA), gender, race, class year, and university. Major(s), future plans, GPA, and university were open ended questions, while race (white, black, or African American, American Indian or Alaskan native, Asian American, native Hawaiian, or other Pacific Islander, Latino/Hispanic, middle eastern), gender (female, male or other) and class year (freshman, sophomore, junior, or senior) were multiple choice. Based on the university of the respondent, we also looked up the average American College Test (ACT) score of accepted freshmen to his or her university from collegedata.com and recorded the number for each respondent. This acted as a rough measure of the prestige of his or her university.

To ensure categorical variables with open ended responses in the survey could be usefully parsed and analyzed, responses were coded in a simplified fashion. Major(s) were categorized as STEM, non-STEM or both (when a respondent had double or triple majors that included both STEM and non-STEM). Future plans were coded with attention to the central concerns of this study: the field into which one goes and how far one plans to advance into said field. Future plans after graduation were classified based on the field the respondent intended to pursue, if known, and in what capacity he or she wanted to go into the field "not necessarily academic. A respondent planning to get a Ph.D. in Chemical Engineering, for example, was classified as G-STEM for graduate studies in STEM, while a respondent intending to work in social services was classified as J-SS for job in social services. A full list of questions in the survey and coding schemes are outlined in **Table S1**.

2.5. Data Analyses

Below we discuss our hypothesis-driven analyses, which are numbered to match the hypotheses outlined in the introduction. Additional analyses are described in the results section.

Our study was designed primarily as test of whether undergraduate FABs mirrored those of academics. Toward that end, we replicated several statistical tests used on academic FABs collected by Leslie et al., using, instead, the undergraduate FABs we collected. (1) To test whether undergraduate FABs were, like academic FABs, strongly predictive of female representation, we used Pearson correlations between the average undergraduate FAB for each field and female representation for each field. (2) To test whether undergraduate FABs were, like those of academics, predictive of female representation across all fields, as well as STEM and non-STEM fields separately, we computed Pearson correlations for all fields, for STEM fields and for non-STEM fields. (3) Leslie et al. found that a field's classification as STEM or non-STEM became an unimportant predictor for female representation when academic FABs were added in a stepwise hierarchical regression. To test whether the same was true for undergraduates, we replicated the same analysis. We used a stepwise hierarchical regression with, first, a field's classification as STEM or non-STEM as a predictor for female representation in each field, and second, average undergraduate FAB for each field as an additional predictor. (4) Leslie et al. found that academic FABs were negatively correlated with African American representation for each field. To test whether the same was true with undergraduates, we used a Pearson correlation between average undergraduate FAB score and African American representation.

Our study was also designed to test other several key predictions of the FAB hypothesis. (5) The FAB hypothesis posits that academic FABs influence the FABs of the public. Undergraduates, in constant contact with academics, would thus be expected to have FABs that approach those of academics as they gain more college experience. To test this, we used a two-way ANOVA "with academic field and undergraduate FABs as predictors for female representation" to examine the effect the academic field and the effect that a student's class year (freshman, sophomore, junior or senior) had on undergraduate FAB scores. (6) The FAB hypothesis takes as a working premise that the public has internalized stereotypes that associate men, but not women, with innate brilliance. The FAB hypothesis thus predicts a key difference between the results for survey version one (undergraduates not provided gender ratios) and version two (undergraduates provided gender ratio): That undergraduates informed of the gender ratio within each academic field will, by virtue of these internalized stereotypes, tend to have FABs that are more predictive of female representation relative to undergraduates not informed. To test this, we computed Pearson correlations between average undergraduate FAB and female representation for each field using survey two data and compared these correlations to those computed for survey one.

3. RESULTS

3.1. Hypothesis-Driven Results

We formulated several hypotheses that the FAB hypothesis would predict to be true of undergraduate FABs (see ending section of introduction). We found evidence for hypothesis (1) and found no evidence for hypotheses (2)–(6). In this section, we report the results for those predictions, indicating the number of the corresponding prediction in the text. We also report results for other analyses we used to explore our data. Although this second set of analyses were not hypothesis-driven, insofar as we did not find evidence for most of our hypotheses, we did these additional analyses to generally explore what conclusions our data could suggest.

We did not find evidence to suggest that undergraduate FABs predict female representation. In turn, because academic FABs strongly predict female representation, we found no evidence to suggest that undergraduate FABs strongly resemble those of academics. (1) As the FAB hypothesis would predict, across all 42 fields, undergraduate FAB scores were, like those of academics, negatively correlated with female representation (**Figure 1**). (2) However, when we analyzed STEM and non-STEM fields separately, we did not detect any correlation for both STEM and non-STEM fields (**Table 1**).

A key finding of Leslie et al.'s study on academics was that when academic FABs are taken into account, the STEM/non-STEM divide becomes unimportant for predicting female representation, suggesting that the FAB hypothesis plays an important explanatory role for the gender gap independent

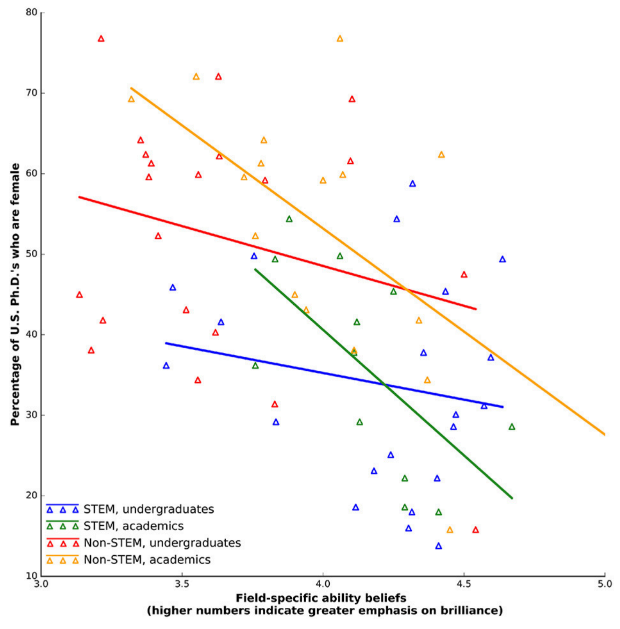


FIGURE 1 | Undergraduate (survey excludes gender ratios) and academic field-specific ability beliefs vs. the percentage of female 2011 U.S. Ph.D.'s.

TABLE 1 | Pearson correlations between field-specific ability beliefs and African American and female representation, stratified by field type (STEM, non-STEM, all fields).

Correlation	Field type	r	p	95% confidence interval
UNDERGRADUATES (SURVEY INCLUDES GENDER RATIOS)				
FABs, percent female Ph.D.s	STEM	−0.31	0.20	−0.67, 0.17
	Non-STEM	0.68	0.002	0.25, 0.82
	All fields	0.28	0.08	−0.03, 0.54
UNDERGRADUATES (SURVEY EXCLUDES GENDER RATIOS)				
FABs, percent female Ph.D.s	STEM	−0.13	0.60	−0.55, 0.35
	Non-STEM	−0.10	0.69	−0.52, 0.36
	All fields	−0.47	0.002	−0.68, −0.19
FAB, percent African American Ph.D.s	All fields	−0.01	0.94	−0.31, 0.29
ACADEMICS				
FABs, percent female Ph.D.s	STEM	−0.64	0.03	−0.88, −0.10
	Non-STEM	−0.65	0.01	−0.86, −0.24
	All fields	−0.63	<0.001	−0.81, −0.34
FAB, percent African American Ph.D.s	All fields	−0.53	0.002	−0.75, −0.21

Results shown for academics (Leslie et al., 2015), as well as for undergraduates with and without gender ratios provided in the survey.

of this divide (Leslie et al., 2015). (3) We replicated the stepwise hierarchical regression used by Leslie et al., using, instead, our undergraduate FABs. For academics, the STEM/non-STEM indicator were a non-significant predictor when academic FABs were added. For undergraduates, on the other hand, the STEM/non-STEM indicator remained significant when FAB scores were added and also mitigated the effect of those scores to the extent that they were a non-significant predictor (**Table 2**).

(4) The FAB hypothesis predicts that the representation of populations who are, like women, stereotyped as lacking innate brilliance will be negatively correlated with FABs (Leslie et al., 2015). African Americans are one such group (Steele and Aronson, 1995). Leslie et al. found that academic FABs were negatively correlated with African American representation. We did not find evidence for the same with undergraduate FABs (**Table 1**).

(5) If academics pass on their FABs to undergraduates, increased exposure to academics should cause undergraduate FABs to converge with academic FABs. In a two-way ANOVA, we found no evidence that class year (freshman, sophomore, junior or senior) [$F_{(1,22413)} = 0.95$, $p = 0.33$] or its interaction with academic field [$F_{(41,22413)}$, $p = 0.87$] had an effect on FAB scores. As a group, we found no evidence that undergraduate's FABs change during college, and, by extension, we found no evidence that they change through prolonged exposure to academics.

Survey version two, for which undergraduates were provided the gender ratio in each academic field, provided no support for our last hypothesis. The FAB hypothesis suggests that (6) informing undergraduates of the gender ratio of each filed should give FABs that rank male-dominated fields as more "brilliance-required"fields (higher FABs) and female-dominated fields as less "brilliance-required" (lower FABs) relative to undergraduates not provided that added information. Our results do not give any evidence for this prediction. Our results, do, however, suggest that this added information dramatically affected undergraduate FAB scores, with markedly different effects for STEM and non-STEM fields. Like survey one, we found no evidence that FAB scores were correlated with female representation in STEM alone for survey two. On the other hand, unlike survey one, FAB scores were positively correlated with female representation in non-STEM alone for survey two (**Figures 2**, **3**; **Table 1**). This is a reversal of the direction of correlation found in survey one. Notably, this reversal did not happen for STEM fields, where the extra information provided in survey two did not change the pattern of responses with respect to survey one.

3.2. Additional Results for Survey One

Our results outlined in the previous subsection were explicitly designed to test the extended FAB hypothesis. The components of the survey and the corresponding results we discuss below were more exploratory than hypothesis-driven, generally designed to explore what conclusions or research directions could be suggested by our data.

To explore which underlying variables can explain patterns in undergraduate FABs, we used exploratory factor analysis (EFA) on undergraduate FAB scores. Two factors were retained, based on parallel analysis, jointly accounting for 65% of the variance (Matsunaga, 2010). The groupings of academic fields that emerged suggested STEM and non-STEM as the latent variables (**Table S2**). Importantly, contrary to what the FAB hypothesis would predict, female representation varied considerably within each group of fields, suggesting that it is whether a field is STEM or non-STEM, rather than the degree of female representation in a field, that is the important underlying influence on undergraduate FABs.

EFA was implemented using a Varimax rotation and, initially, a four factor solution with the factanal function in R. Parallel analysis indicated that two factors should be retained for survey one and three factors for survey two (Matsunaga, 2010). Using an intermediate cut-off loading score of 0.6, no cross loading occurred. EFA results are similar, with some cross loading for smaller cut-offs, for a wide range of commonly used loading score cut-offs (0.4–0.7) (Matsunaga, 2010).

To assess how personal, educational and demographic characteristics might influence undergraduate FABs, respondents were asked for their major(s) (categorized into STEM, non-STEM or — in the case of STEM/non-STEM double majors — both), their post-college plans (sorted into 19 different categories, i.e., "medical school," "graduate school for a STEM field," etc.), GPA, gender, race, class year, and university (**Table S1**). We coded each university's average ACT score of admitted freshmen as a proxy for institutional prestige. To gauge the effects of these characteristics on undergraduate FAB scores, MANOVA was used with FAB scores as the dependent variables and the respondent characteristics as the independent variables. MANOVA was non-significant ($p > 0.1$ for Pillai's Trace) for all characteristics (**Table S4**).

As another way to explore how FABs for STEM fields and non-STEM fields differ, we compared mean FAB scores for STEM and non-STEM fields using Welch's t-test. For academics, we found no evidence that FABs for STEM and non-STEM fields differed

TABLE 2 | Hierarchical regression models predicting female representation from undergraduate data (n = 42 academic fields).

	Model 1			Model 2		
Predictor	$\hat{\beta}$	t	p	$\hat{\beta}$	t	p
STEM/non-STEM categorization of field	−0.56	−4.26	<0.001	−0.42	−2.43	0.02
Undergraduate field-specific ability beliefs				−0.21	−1.20	0.24
Adjusted R^2		0.295			0.303	
F statistic for change in adjusted R^2		18.15			9.90	
P-value for change in adjusted R^2		< 0.001			< 0.001	

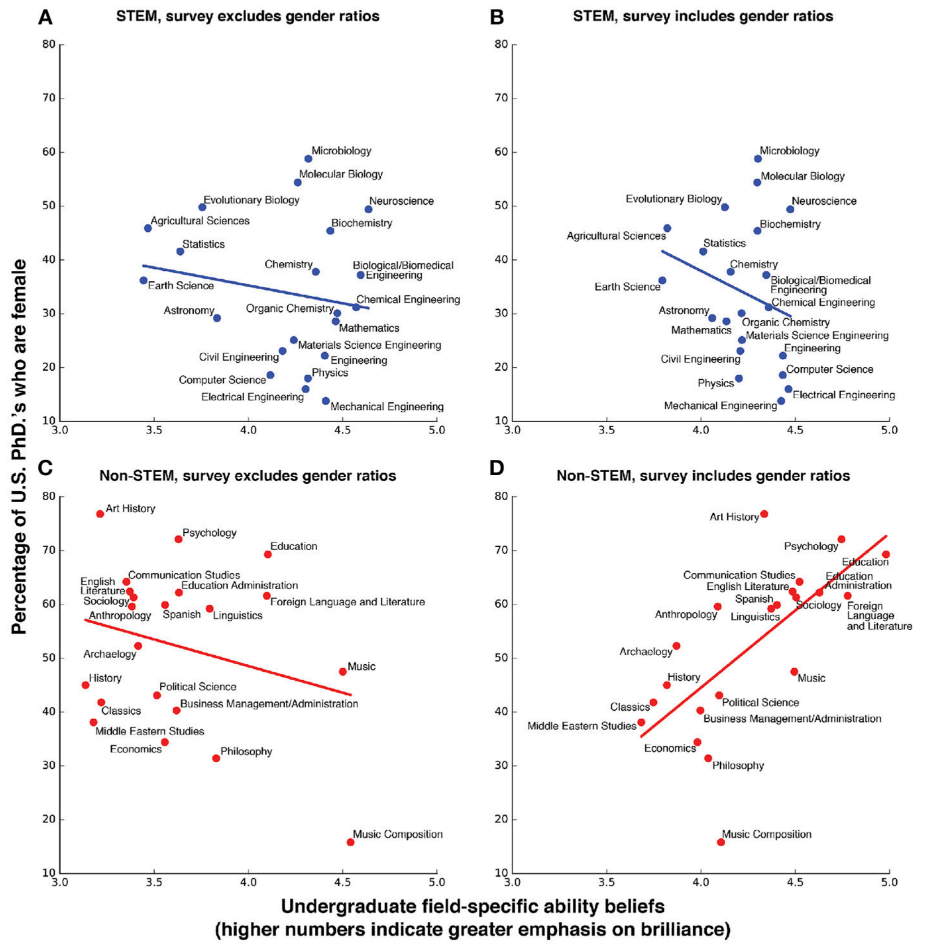


FIGURE 2 | Undergraduate field-specific ability beliefs vs. female representation, with and without field-specific gen- der ratio information provided. Field-specific ability beliefs and the percentage of U.S. female Ph.D.'s in 2011 for STEM fields **(A)** without gender ratios included in the survey and **(B)** with gender ratios and non-STEM fields **(C)** without gender ratios included in the survey and **(D)** with gender ratios.

[$t_{(27)} = -0.87$, $p = 0.390$, 95% confidence interval: –0.37, 0.15]. For undergraduates, on the other hand, we found evidence that FABs for STEM and non-STEM fields were different [$t_{(39)} = -5.61$, $p < 0.001$, 95% confidence interval: –0.85, –0.40], with an average FAB score of 3.61 for STEM and 4.24 for non-STEM fields.

3.3. Additional Results for Survey Two

Additional analyses on the second survey produced results similar to those of the first survey. MANOVA using the personal educational and demographic characteristics of respondents was non-significant (p >0.1 for Pillai's Trace) across all characteristics (**Table S5**). Exploratory factor analysis, with three factors retained, cumulatively accounted for 49% of the variance (Matsunaga, 2010). Like the first survey, the groupings of fields that emerge suggested STEM and non-STEM fields as latent variables (**Table S3**).

3.4. Additional Analyses

Insofar as many of our hypotheses are dependent on our Pearson correlations between percent female representation and undergraduate field-specific ability belief scores, we wanted to test whether those correlations are robust to outliers To test this, we conducted a sensitivity analysis. Looking at our regression plots for undergraduate FABs vs. female representation suggests that no academic field was an outlier for STEM fields while Music Composition was an outlier for non-STEM fields (**Figure 2**). Thus, we re-computed our correlations for non-STEM fields and across all fields, for both versions of the survey, with Music Composition excluded. Our Pearson correlations results were similar with and without the outlier.

Insofar as we used two different survey administration methods (Online and in-class administration), we wanted to test whether, on average, differences in demographic/education variables and/or FAB scores due to the manner in which the survey was presentation (online vs. in-class). To test this, a logistic regression model was fitted separately for each survey version (with gender ratios provided, without gender ratios provided). For each model, the survey presentation type was the outcome and FAB score and demographic/education characteristics were predictors. For both survey versions, we

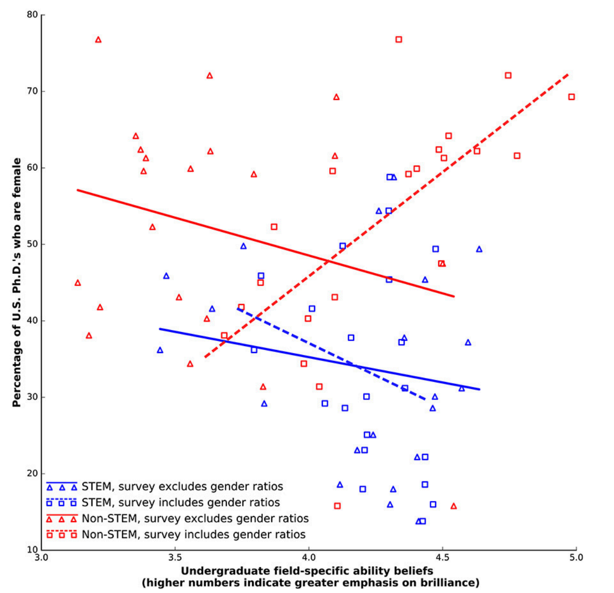


FIGURE 3 | Undergraduate field-specific ability beliefs vs. the percentage of female 2011 US Ph.D.'s in (blue) STEM and (red) Social Sciences/Humanities fields, for both survey that excluded gender ratios (solid trend lines) and included the gender ratios (dotted trend lines).

found no evidence of association between survey presentation type and any predictor ($p > 0.05$ for each predictor).

4. DISCUSSION

Research on female representation in academia has hitherto been largely focused on explaining female underrepresentation in certain STEM fields. The FAB hypothesis raises the possibility that gender representation in STEM can find a unifying explanation, independent of the STEM/non-STEM divide, in widespread beliefs at play across the entire academic spectrum (Ceci and Williams, 2011; Leslie et al., 2015; Bian et al., 2017). Our results, however, do not support a key prediction of the FAB hypothesis—that undergraduate FABs reflect those of academics. Rather, our results collectively suggest that it is the STEM/non-STEM divide that plays the foremost role in shaping undergraduate FABs.

To whatever extent FABs contribute to the gender gap, this study suggests that undergraduate FABs are not indifferent to the divide, but are instead structured by that divide. We did not find evidence that undergraduate FABs are predictive of female representation within STEM and non-STEM fields separately. Although undergraduate FABs were predictive of female representation across all fields, FABs became an unimportant predictor when the STEM/non-STEM divide was taken into account. Exploratory factor analysis reiterated the importance of the divide, showing it "and not female representation" to be the latent variable influencing undergraduate FABs. Average undergraduate FAB scores differed between STEM and non-STEM fields. We also found no evidence that undergraduate FABs change during college, suggesting the STEM/non-STEM divide consistently shapes undergraduate beliefs, even as they gain prolonged exposure to academics.

The striking difference between the results of our two surveys also reinforces the importance of this divide. While the correlation between undergraduate FABs and female representation remained relatively unchanged for STEM fields, this relationship was reversed for non-STEM: when provided the gender ratio in each field, undergraduates rated the fields with more women as requiring more talent. Our results do not point to a definitive interpretation of this reversal. Our study should, ideally, be unaffected by social desirability bias. But because female representation in STEM is currently a well known and controversial topic, including information about the gender ratio in each field may have evoked social desirability bias. Conservatively, we can say our results suggest that undergraduates are comfortable linking females and intellectual talent in non-STEM fields, and that undergraduate beliefs about the relationship between gender and talent in academia are structured by the STEM/non-STEM divide.

The FAB hypothesis builds on Dweck's work on "growth" vs. "fixed" mindset, proposing that entire academic fields may be placed along a spectrum between the "growth" and "fixed" mindset by measuring the FABs of each field. Our results suggest that academic fields can be fruitfully understood as falling along such a spectrum. Indeed, as Leslie et al. found for academics, we found a considerable spread of undergraduate FAB average scores for academic fields. Furthermore, we found that undergraduate FABs were significantly higher for STEM fields than non-STEM fields. In other words, our results suggest that undergraduates on average view STEM fields as "harder"—requiring more of an unteachable, "fixed" spark of brilliance—than non-STEM fields.

Although it is unclear what could be responsible for the difference between academic FABs and undergraduate FABs, our results raise the possibility that the jump from undergraduate to graduate level could be a critical juncture for shaping FABs. We found no evidence that undergraduate FABs change during college, even as undergraduates accumulate more exposure to academics and, by extension, their FABs. Leslie et al., likewise, found no evidence that the FABs of graduate students, postdoctoral researchers and professors differed (Leslie et al., 2015). This does not rule out the possibility, however, that the US undergraduate experience could be considerably different to that of graduate students and beyond, cultivating considerably different FABs.

Our study is not without limitations. Participants opted in, thus raising the spectre of selection bias. Though our sample includes students from 197 geographically diverse US colleges and universities, the majority came from two universities in Central New York. This may have introduced geographical and socioeconomic biases into our sample. Unlike Leslie et al., who surveyed academics exclusively at "high profile research universities" (Leslie et al., 2015), our responses came from broad range of institutions. When we tested for effects of institutional prestige, however, we found it to be a non-significant explanatory variable. We therefore cannot assess to what extent differences in the FABs of undergraduates and academics might be attributable to the effect of institutional rankings. Universities of widely varied rankings, however, could possibly encourage considerably different beliefs about academic talent.

Many pivotal life decisions that mold gender representation in academia happen at the undergraduate level. Overall, our results suggest that the search for strategies to diversify academic fields should take into account how the STEM/non-STEM divide is central to undergraduate perceptions of both gender and talent in academia.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of US Department of Health and Human Services regulations for the protection of human participants in research, as followed by the Human Subjects Committee of State University of New York (SUNY) at Oswego, with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Human Subjects Committee of SUNY Oswego.

AUTHOR CONTRIBUTIONS

KB, DH, and CI: designed research; KB and CI: performed research; KB, SW, AN, and SP: analyzed data; KB, DH, SW, AN, SP, and CI: wrote the paper.

FUNDING

This work was supported by a Student Mini Grant from State University of New York at Oswego and a U.S. Student Fulbright grant.

ACKNOWLEDGMENTS

We thank Charin Smith of Oswego, NY, USA, Melissa Schmitz of Le Moyne College, NY, USA, additional members of the Institute for Quantitative Social Science at Harvard University, the Biostatistics Department at Harvard University, Christian Meisel of the National Institute of Mental Health and the reviewers.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fsoc.2019.00026/full#supplementary-material

REFERENCES

Bennett, M. (1996). Men's and women's self-estimates of intelligence. *J. Soc. Psychol.* 136, 411–412.

Bian, L., Leslie, S.-J., and Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science* 355, 389–391. doi: 10.1126/science.aah6524

Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236

Ceci, S. J. and Williams, W. M. (2007). *Why Aren't More Women in Science?* Washington, DC: American Psychological Association.

Ceci, S. J. and Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3157–3162. doi: 10.1073/pnas.1014871108

Dar-Nimrod, I. and Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science* 314, 435–435. doi: 10.1126/science.1131100

Dweck, C. S. (2006). *Mindset: The New Psychology of Success*. New York, NY: Random House Incorporated.

Kirkcaldy, B., Noack, P., Furnham, A., and Siefen, G. (2007). Parental estimates of their own and their children's intelligence. *Eur. Psychol.* 12, 173–180. doi: 10.1027/1016-9040.12.3.173

Lecklider, A. (2013). *Inventing the Egghead: The Battle Over Brainpower in American Culture*. Pennsylvania, PA: University of Pennsylvania Press.

Leslie, S.-J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 262–265. doi: 10.1126/science.1261375

Matsunaga, M. (2010). How to factor-analyze your data right: do's, don'ts, and how-to's. *Int. J. Psychol. Res.* 3, 98–111. doi: 10.21500/20112084.854

Meyer, M., Cimpian, A., and Leslie, S.-J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Front. Psychol.* 6:235. doi: 10.3389/fpsyg.2015.00235

National Science Foundation (2011). *Survey of Earned Doctorates*. Alexandria, VA: National Science Foundation

Steele, C. M. and Aronson, J. (1995). Stereotype threat and the intellectual test performance of african americans. *J. Person. Soc. Psychol.* 69, 797–811.

Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educ. Stud. Math.* 41, 191–207.

Valian, V. (1999). *Why So Slow?: The Advancement of Women*. Cambridge, MA: MIT press.

Wigfield, A. and Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 24 April 2019
doi: 10.3389/fpsyg.2019.00836

Check for updates

Self-Concept Profiles in Lower Secondary Level – An Explanation for Gender Differences in Science Course Selection?

Steffani Saß* and Nele Kampa

Leibniz Institute for Science and Mathematics Education (IPN), Kiel, Germany

OPEN ACCESS

Edited by:
*Bernhard Ertl,
Universität der Bundeswehr
München, Germany*

Reviewed by:
*Rachael D. Robnett,
University of Nevada, Las Vegas,
United States
Isabelle Schmidt,
University of Trier, Germany
Hanke Korpershoek,
University of Groningen, Netherlands*

***Correspondence:**
*Steffani Saß
sass@ipn.uni-kiel.de*

Specialty section:
*This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Psychology*

Received: *09 January 2019*
Accepted: *29 March 2019*
Published: *24 April 2019*

Citation:
*Saß S and Kampa N (2019)
Self-Concept Profiles in Lower
Secondary Level – An Explanation
for Gender Differences in Science
Course Selection?
Front. Psychol. 10:836.
doi: 10.3389/fpsyg.2019.00836*

One of the most powerful determinants of course selection in upper secondary level is undoubtedly students' self-concept. Students with a high self-concept in a domain are more likely to select a course in that domain. However, according to the dimensional comparison theory, the formation of self-concept includes comparison processes with self-concepts in other domains. Regarding gender, females are less likely to choose physics and are more likely to have lower STEM self-concepts as well as lower aspirations toward STEM careers than males. In Germany, students in Grade 10 choose specific academic tracks to attend during upper secondary school. The academic track choice goes in hand with choosing advanced courses. This choice entails the decision about whether to pursue STEM subjects. We adopted the person-centered approach of latent profile analysis (LPA) to investigate the patterns of students' self-concepts across the five domains, math, biology, reading, English, and physics. Furthermore, we investigated how those patterns influence educational choices regarding science subjects in upper secondary school in Germany. Based on a sample of 1,658 students, we tested whether the distinct profiles of self-concept in different domains in Grade 8 predicted gendered science course selection in Grade 10 as well as career aspirations in science. LPAs yielded four distinct profiles of self-concept that differed in level and shape: high math, high verbal, low overall, and high overall. These profiles were equivalent across gender. Gender differences were manifested in the relative distribution across the four profiles: females were more present in the low overall and high verbal-related self-concept profiles and males in the overall high and high math-related self-concept profiles. The profiles differed regarding abilities, choice of science course in upper secondary level, and science career aspirations.

Keywords: self-concept profiles, dimensional comparison theory, person-centered approach, gender differences, course choice

INTRODUCTION

Despite women increasingly entering higher education and occupations in general, they are still underrepresented in STEM occupations and university attendance (additional material of National Science Foundation, 2003, 2015a,b; Organisation for Economic Co-operation and Development [OECD], 2006, 2016). Field of study choices differ between males and females (e.g., Trusty, 2002;

Nagy et al., 2006; Perez-Felkner et al., 2017) and are already channeled through course choice patterns in upper secondary education. Previous studies have found that, while females already opt out of STEM subjects during school, males more frequently choose advanced science-related or math subjects (Pinxten et al., 2012). Consequently, these first decisions in favor of or against STEM courses in secondary school influence future long-term educational aspirations such as the decision for STEM studies and occupations (Trusty, 2002; Ayalon, 2003; Lörz et al., 2011). One of the most prevalent predictors for opting out of STEM subjects during school is domain-specific self-concept (Wang et al., 2013; Guo et al., 2015). However, most research on course selection focuses on the unique effects of self-concept in a particular domain on course choice in the respective domain. Only few studies have investigated the effects of a combination of different academic self-concepts on course selection (Nagy et al., 2006; Marsh et al., 2009; Guo et al., 2017). These studies found that students compare their perceived abilities in various domains and thus form specific self-concept configurations, which in turn might influence course selection. However, these studies incorporated only math and verbal (and other nonscience) self-concepts (Marsh et al., 2009; Umarji et al., 2018), only science self-concepts (Guo et al., 2017), or only a small selection of science and/or math self-concepts (Nagy et al., 2006). Drawing on the dimensional comparison theory (DCT; Möller and Marsh, 2013), we simultaneously considered self-concepts in various science and nonscience domains and investigated the patterns of self-concept combinations. To the best of our knowledge, we are the first to combine a person-centered approach with a multigroup approach to investigate such a large array of self-concepts relevant for science course selection in order to detect gender differences in these self-concept profiles. Furthermore, we investigated the relation of these profiles to science course selection and career aspirations. Finally, we take a long-term approach by not investigating males and females self-concepts at the time of academic choices toward science but by investigating the impact of early self-concept on later academic choices.

Dimensional Comparisons and Structure of Self-Concept

Students' self-concept is defined as being multifaceted as well as hierarchical and as the self-perception of an individual's ability in a specific domain (Marsh, 2007). According to the DCT (Möller and Marsh, 2013), different dimensional comparisons between domains shape self-concept. Students compare their achievement in one domain with their achievement in another domain. This comparison either leads to self-concepts that are quite distinct from each other (high self-concept in one and low self-concept in the other domain; contrast effect) or to rather similar self-concepts (high self-concept in both domains; assimilation effect). Which effect occurs depends on the perceived subject similarity (Möller and Marsh, 2013; Möller et al., 2015). Rather dissimilar subjects are clearly distinct from each other, for example, subjects containing high levels of math content versus subjects containing high levels of verbal content. This phenomenon is also described as the math-verbal continuum (see Marsh et al., 1988). Research on comparison processes in self-concept formation has confirmed that math and verbal domains, such as German or a foreign language, constitute dissimilar subjects (Haag and Götz, 2012; Helm et al., 2016). Regarding the science and math domains, we assume that particular domains that are perceived as being more similar or less similar differently effect dimensional processes. In this regard, assimilation effects exist for math and physics while reading and physics operate as rather contrasting domains (Jansen et al., 2015). Looking at similarities between science subjects, one can assume that biology is more distinct from math than physics. Möller and Köller (2004) set up a model in which self-concept in physics and chemistry is assigned to an overarching mathematical self-concept factor, and biology self-concept is the only science domain that is assigned to an overarching verbal self-concept factor. Particularly biology and physics self-concepts serve as contrasting domains whereas physics and chemistry appear more similar to each other, thus resulting in assimilating dimensional comparisons (Guo et al., 2017).

Self-concept is reciprocally related to academic achievement; thus, individuals' perception of their abilities is based on prior achievement and vice versa (e.g., Marsh and Craven, 2006). Therefore, achievement serves as an evaluator that shapes self-concept in each domain. School grades or scores of standardized achievement tests usually serve as indicators of achievement. School grades not only reflect a student's academic achievement but also his or her relative position within the classroom setting. Grades are also a direct form of feedback for students regarding their achievement in a specific subject. In contrast, achievement scores in a standardized test are a criterion-referenced measure that is unbiased by group effects and not directly reported back to the student.

Most of the studies mentioned above applied a variable-centered approach (e.g., Nagy et al., 2006; Guo et al., 2017). Because within-person hierarchies of self-concept in various domains influence dimensional comparisons, person-centered approaches are also required. With person-centered approaches, individuals (in our case, students) with similar patterns regarding specific indicators are clustered together and students with disjunctive patterns are placed in different clusters (Lubke and Muthén, 2005). Hence, each student is categorized into a specific group according to his or her specific self-concept pattern. Thus, the approach attempts to carve out homogeneous groups within a heterogeneous population. Compared to the dominant variable-centered approaches, applying a person-centered approach can provide greater insights into students' self-concept combinations. Most of the person-centered studies on motivational patterns used a set of motivational variables rather than self-concept alone (Lazarides et al., 2016; Linnenbrink-Garcia et al., 2018). To date, only a few studies have applied a person-centered approach to investigate self-concept patterns in students using various types of domain-specific self-concept (Marsh et al., 2009; Umarji et al., 2018). Marsh et al. (2009) applied a latent profile analysis (LPA) to a diverse set of academic self-concepts: math, verbal, problem solving, intellectual, artistic, political, technical, and computer self-concept, assessed in the final year of upper secondary school. They were able to distinguish between profiles that showed a combination of quantitative (high, low, and average overall self-concept) and qualitative differences among the single domains (e.g., high math self-concept and low verbal self-concept and vice versa). The authors identified four profiles, which reflected different levels of overall self-concept, that is, one low self-concept pattern, one medium self-concept pattern, and two high self-concept patterns. These profiles also showed opposing math and verbal self-concepts, indicating the verbal-math continuum based on the DCT (Möller and Marsh, 2013). Only one group had an average level of all self-concept indicators with rather similar math and verbal self-concept means. Similarly, a study by Umarji et al. (2018) incorporated the two contrasting domains of math and English (mother tongue) self-concept. The authors identified five groups of students in Grade 7 with three qualitative level differences among the two domains (e.g., low math—medium English, medium math—high English, or high math—medium English) and two groups of students that differed in terms of quantitative overall levels (low in both domains and high in both domains). Although the self-concept indicator as well as the number of profiles varied in both studies, there is evidence of at least two profiles representing contrasting domains (high math-related and low verbal-related vs. high verbal-related and low math-related). Moreover, both studies also identified groups that differed only in their quantitative level but did not show differences between math and verbal self-concept. The described studies on self-concept suggest that, within the framework of person-centered approaches, one can expect to find at least two profiles showing shape differences and two profiles showing level differences (overall high and overall low self-concept).

Domain-Specific Self-Concept and Academic Choices Toward Science

Motivational variables such as self-concept are important predictors of course selection (Marsh and Yeung, 1997) and aspirations toward a career in science (DeWitt et al., 2013). Several studies have revealed that domain-specific self-concept predicts academic choices over and above prior academic achievement (e.g., Marsh and Yeung, 1997; Köller et al., 2000; Watt, 2006; Nagy et al., 2008; Wang and Degol, 2013). Students with a high self-concept in a STEM domain are more likely to remain on a STEM path even after it is no longer compulsory (Halpern et al., 2007). Students with a high math-specific self-concept are more likely to select math courses in high school (Nagy et al., 2006; Watt, 2006). The effects in math ranged between OR = 2.15 to OR = 2.95 after controlling for several cognitive variables in logistic regressions (Köller et al., 2000; Nagy et al., 2008). Comparable analyses were performed to predict the choice of STEM occupations and they showed that math-specific self-concept in the 12th grade was also one of the stronger predictors of this choice (OR = 1.46). Another study revealed that self-concept in science predicts aspirations for a science career (Nagengast and Marsh, 2012).

The presented studies incorporated only same-domain self-concept in order to predict academic choices in the respective domain, we not only considered self-concept and achievement in the chosen domain but also incorporated several domain-specific science and nonscience self-concepts to account for various comparison processes. Research on the DCT model has already proved the applicability of intraindividual dimensional comparisons in explaining academic choices while using variable-centered approaches (Nagy et al., 2006; Parker et al., 2012; Guo et al., 2017) as well as person-centered approaches (Wang et al., 2013; Umarji et al., 2018). Regarding the dual role of math and language as contrasting domains, results have revealed that students with a higher math self-concept relative to verbal self-concept are more likely to choose a math intensive major (Nagy et al., 2008; Parker et al., 2012; Guo et al., 2015; Umarji et al., 2018). This is also the case in dimensional comparisons between math and biology self-concepts and course selection in the respective subject. Students with a high self-concept in math were more likely to choose an advanced math course and less likely to opt for an advanced biology course (Nagy et al., 2006). Regarding science-specific comparisons, Guo et al. (2017) showed that self-concept in physics negatively predicted coursework aspirations in biology and vice versa. Taken together, science course selection is based not only on self-concept and achievement in a particular science domain. Rather, students seem to compare their achievement and self-concept with other domains such as reading as a rather dissimilar domain or math as a similar domain and thus evaluate their individual relative strengths. However, up until now, studies have not included a wide spectrum of science and nonscience self-concepts. Thus, a person-centered approach incorporating these various self-concepts could be especially fruitful to investigate whether students' science course selection and career aspirations are associated with students' intraindividual patterns of self-concept in various domains.

Gender Differences in Self-Concept and Educational Choices

Although gender differences in certain abilities such as math and reading seem to have become increasingly negligible (Hyde, 2005; Else-Quest et al., 2010; Lindberg et al., 2010), gender differences in self-concept are still visible. A number of studies have revealed that, from early school years onward, the self-concepts of females and males differ depending on the academic domain (Herbert and Stipek, 2005; Spinath et al., 2014). Females, for instance, often have a lower positive self-concept in science and math than males do (Marsh and Yeung, 1998; Köller et al., 2000; DeWitt et al., 2013), whereas males show relatively lower self-concepts in verbal abilities (Schilling et al., 2006). This gender gap in the perception of one's own abilities even exists after prior achievement has been controlled for (Wilkins, 2004; Sikora and Pokropek, 2012; Jansen et al., 2014). Thus, females seem to underestimate their abilities in math, which in turn might lower their math self-concept.

Taking a more differentiated look at science, gender differences vary across scientific fields (Britner, 2008). In physics, males tend to have a higher self-concept than females do (Schilling et al., 2006; Jansen et al., 2014). The results on self-concept in biology are mixed. While Schilling et al. (2006) found a higher self-concept in biology for males, other studies reported a higher self-concept in biology for females (Nagy et al., 2006;

Jansen et al., 2014). Because gender differences do occur in terms of self-concept, different mechanisms behind males' and females' academic choices for and against STEM subjects could be derived from these self-concept differences. These different perceptions of one's own abilities in a scientific field could lead to males choosing math- and science-related courses more often (Smyth and Hannan, 2006). Differences in self-concept are a key predictor in explaining gender differences in educational choices (Nagy et al., 2006; Watt, 2006). For instance, females with high abilities and a high self-concept in biology are more likely to choose advanced biology courses (Nagy et al., 2006). Whereas, for males, the same effect of math achievement and self-concept on math course selection is more pronounced than for females (Guo et al., 2015).

Taken together, self-concept not only differs across domains but also across gender. This effect has repeatedly been found in variable-centered studies. So far, studies have not tested for gender-specific profiles by combining a multigroup (gender) and a person-centered approach (homogeneous groups within one gender category). In such a framework, gender differences can mainly occur in two different ways. First, males and females can have different profile patterns, which lead to different choices regarding science. Second, both genders can show the same profile patterns, but the distribution across the profiles may differ across gender. Some studies have revealed that males were more likely to fall into a profile with high math self-concept, whereas females were more likely to be in a high verbal self-concept profile (Marsh et al., 2009; Umarji et al., 2018). However, the studies did not systematically test for gender invariance in profiles and did not apply these methods to investigate gendered paths to STEM, for example, academic choices toward science. The nature of profiles can differ across gender in terms of number of profiles, structure of profiles, as well as the relationship between profiles and predictors, or outcomes. Thus, taking a multigroup perspective within a person-centered approach into account not only helps to systematically assess gender differences in the nature of profiles but also to fully disentangle gendered pathways into science choices and career aspirations. We took this innovative approach and investigated several science and nonscience self-concepts that may lead to staying in or dropping out of science courses. In addition, we took a new approach by focusing not on self-concept at the time of the academic choice but on the long-term effects of self-concepts for these academic choices.

The German Secondary School System

Despite each federal state in Germany having its own education system, the systems also share common features. First, science education takes place in separate courses for biology, chemistry, and physics, and typically starts in Grade 7. Prior to Grade 5, elementary schools provide a combined course that includes social science, history, geography, science, and technology. In the federal state of Hamburg, where our study took place, in Grades 5 and 6, academic secondary schools offer an integrated science/technology course. In Grade 7, the structure changes to separate courses in biology, chemistry, and physics. Second, when entering upper secondary education (10th grade) students have to choose advanced courses that they would like to focus on and basic courses that they wish or need to continue. These choices take place in track systems or course systems, depending on the federal state. Hamburg has a track system. At the end of Grade 10, students have to choose between five tracks, namely, science/technology, language, social sciences, arts, and sports.

The Present Study

The aim of the present study was to identify males' and females' self-concept profiles using LPA. Further, we incorporated several relevant predictors to account for the influence of achievement and outcome variables to fully discern the mechanisms behind gendered differences in self-concept profiles and in academic choices toward science. Unlike previous studies, we incorporated a wider range of science and nonscience self-concepts, namely, math, verbal, biology, and physics; we also focused on the impact of early self-concept on later academic choices.

First, we examined students' self-concept profiles across the domains of math, biology, physics, German, and English, separately for each gender. Applying a person-centered approach, we expected to find different profiles that show qualitative and quantitative self-concept differences. We expected to find profiles that demonstrate the verbal-math continuum based on the DCT (Möller and Marsh, 2013). Therefore, we assumed that at least four profiles would emerge. We expected to find two qualitatively different profiles; one profile showing high self-concept values in the math-related domains of math and physics, and low self-concept values in the verbal-related domains of English and reading; and a second profile showing low self-concept values in the math-related domains and high self-concept values in the verbal-related domains. Because results concerning the placement of biology on the math-verbal continuum are mixed, we did not formulate a specific hypothesis regarding the assignment of biology. Studies investigating patterns of motivational variables have demonstrated that profiles also differ in terms of quantity (high and low overall self-concept, in some instances, a medium overall self-concept), in which the values of math-related and verbal-related self-concepts are rather similar (Marsh et al., 2009; Umarji et al., 2018). Hence, we also expected to find at least two profiles, one with overall low self-concept values and one with overall high self-concept values.

Second, applying the novel approach for testing gender invariance in LPA, we systematically assessed whether the profiles were similar across gender (Morin, 2016). We did not specify hypotheses regarding differences between males and females due to the lack of studies using the person-centered approach to investigate the profiles of domain-specific self-concepts across gender in a multigroup framework. Thus, we conducted an explorative test regarding differences in the number of profiles and the qualitative differences of profiles across gender. However, research regarding gender differences in self-concept has shown that males have higher self-concept values in math-related domains and females have higher self-concept values in verbal-related domains. Hence, we expected that proportionally more females would be in profiles characterized by a high verbal self-concept whereas males would more frequently belong to profiles characterized by a high math self-concept.

Third, achievement in the respective domain influences the formation of students' self-concept and students rely on achievement in different domains to shape their self-concept. Therefore, we hypothesized that high achievement in math domains and low achievement in verbal domains would be positively associated with high math-related self-concept (HMRSC) profiles whereas a low math and high verbal achievement constellation would be positively associated with profiles characterized by high verbal-related self-concept (HVRSC) profiles. Due to our novel approach, which combined the person-centered and multigroup perspective, we were not able to formulate hypotheses regarding gender differences within the relations between achievement and self-concept.

Finally, we expected that group membership would predict science course selection and science aspirations. More precisely, we hypothesized that students in high math- and science-related self-concept profiles would be more likely to choose a science course and hold higher aspirations toward science occupations than students in high verbal self-concept profiles. Because several studies have revealed that males are more likely to choose a science course than females in upper secondary school, we hypothesized that, in profiles with quite identical math and verbal self-concept values, females would be more likely to choose science courses.

MATERIALS AND METHODS

Procedure and Sample

We performed secondary analyses on a subsample of the longitudinal "Competencies and Attitudes of Students" study [KESS] (Bos and Gröhlich, 2010) conducted in the German federal state of Hamburg. The sample of the KESS study comprised various school types. We investigated the subsample of 1,658 students attending 61 academic schools (*Gymnasium*). Only this school type leads to upper secondary education and comprises 3 years more than the other school types. We used data from the measurement points in Grade 8 and Grade 12, measured in 2007 and 2011, respectively. Domain-specific self-concept, science aspirations, and ability in various domains (achievement test scores and grades) were measured at the end of the eighth grade. In the 12th grade, students stated which track they selected in Grade 10.

Measures

Domain-Specific Self-Concept

Students' domain-specific self-reported perception of their abilities was assessed with separate scales for each domain. Self-concept in biology, physics, and math was each measured using three scales adopted from the Academic Self-Description Questionnaire II scale (ASDQ II; Marsh, 1990). Each scale comprised three items, for example, "I have always been good in [biology/physics/math.]" (biology: $\alpha = 0.87$, physics: $\alpha = 0.90$, and math: $\alpha = 0.93$). Self-concept in reading (German) was assessed using seven items adopted from the Progress in International Reading Literacy Study (PIRLS; Kelly, 2003), for example, "I have difficulties understanding texts." ($\alpha = 0.75$). Self-concept in English was assessed using five items adopted from Jerusalem (1984). A typical item was: "Nobody can be good in every subject. I am just not talented in English" ($\alpha = 0.91$). All items were rated on four-point Likert-type scales.

Science Track Selection

Students indicated their course choices for Grades 11 and 12. We used the choice of the first course as an indicator of the choice of a science or nonscience track (values 1 and 0). The subject choices biology, chemistry, and physics indicated science track selection. All remaining first course choices were an indicator of nonscience track selection.

Science Aspirations

On a four-point Likert-type scale, students answered an item asking whether they were interested in choosing an occupation in a scientific field after completing school. The item was measured in Grade 8.

Standardized Achievement Test

Students' achievement in science, math, and reading was measured using standardized achievement tests covering the curriculum in the respective domain. Applying an anchor-item design and IRT scaling across the three measurement points ensured the comparability of the achievement test scores (Davier and Davier, 2007). A common longitudinal metric was built ($M = 100$, $SD = 30$). All reliabilities reported in the manuals of the respective studies were satisfactory (Bos et al., 2010).

School Grades and Gender

Students indicated their grade in math, German, English, biology, and physics from the latest school report card. School grades in lower secondary level range between one and six, with one being the highest score. We recoded the grades for the sake of clarity, thus, higher scores reflect higher achievement. Finally, students reported their gender (0 = male, 1 = female).

Statistical Analyses

All analyses were conducted in M*plus* 7.4 (Muthén and Muthén, 1998–2015) using maximum likelihood with robust standard errors (MLR). Our initial analyses included correlation analyses and tests for mean differences between males and females. The data set contained data missing at random in the indicator variables for the LPA as well as in the covariates. Therefore, we handled the missing data with the imputation function in M*plus* and produced 50 imputed data sets (Graham et al., 2007; Enders, 2010; Hickendorff et al., 2018). We integrated these 50 data sets with the function TYPE = IMPUTATION. We also took the hierarchical data structure into account by adjusting standard errors using the TYPE = COMPLEX option in M*plus*, specifying schools as clusters. LPA was applied using *z*-standardized domain-specific self-concept indicator variables.

In the first step, we examined the number of profiles for males and females independently in consecutive LPAs. In each model, indicator means were freely estimated while the variance of the indicators was held equal across profiles[1]. Models were estimated based on 5,000 random sets of start values and 100 iterations per start. The decision for an optimal profile solution was based on a variety of statistical fit indexes. Thus, we evaluated the Akaïke information criterion (AIC), the consistent AIC (cAIC), the Bayesian information criterion (BIC), and the sample-adjusted BIC (SABIC). Simulation studies recommend the use of the cAIC, BIC, and SABIC (e.g., Nylund et al., 2007; Tofighi and Enders, 2008). Lower values indicate a better fit. As the fit indices in large sample sizes frequently improve with the number of profiles, we also inspected elbow plots to detect the point at which the decrease in the fits became negligible (Morin et al., 2016a). We also considered the theoretical meaning of the profile solutions (Muthén, 2003; Marsh et al., 2009) as profiles should be meaningful and interpretable as well. Hence, we inspected the self-concept composition of additional profiles. We preferred models with fewer profiles when an additional profile was quite similar to the already existing profiles. Finally, the entropy and the average posterior probabilities, varying from 0 to 1, gives information about how well students can be classified to their most likely profile. The higher the value, the smaller the classification error. Nagin (2005) suggests an average posterior probability value of at least 0.70 as the cutoff for an acceptable classification error. In the following steps, we used a multigroup approach using the KNOWNCLASS function to investigate whether gender differences occurred in (a) the nature of the profiles, (b) the prediction, and (c) the outcome (Morin, 2016; Morin et al., 2016b).

In the second step, we examined the degree of similarity across gender in a sequence of models after identifying the final number of profiles for males and females. First, a baseline model with freely estimated profile-specific indicator variables across gender was estimated (configural model). Then, the models were successively constrained with regard to profile indicators across gender (i.e., means and variance). Thus, we tested whether the means within the profiles were similar across gender (structural similarity) and we tested whether the variance within profiles were similar across gender (dispersion similarity). Finally, we examined whether the relative profile sizes were equal across gender (distributional similarity). Models with equality constraints across gender were compared to a less restrictive model using cAIC, BIC, and SABIC. Lower values on at least two information criteria indicated a favorable model (Morin, 2016).

In the third step, we added predictors to the final model. A multinomial logistic regression was used to predict profile membership. We further tested for differences across gender by comparing a model constraining logistic regression coefficients to be equal between males and females (predictive similarity) with a model where coefficients were freely estimated between males and females. We then compared the model fit. In a last step, we added the outcome variables, science course selection, and science aspirations to the final model. Again, we first estimated a model with freely estimated within-profile outcomes across gender and compared the model to a model with equality constraints of the outcome variables across gender (explanatory similarity). To ensure that including covariates and outcomes did not change the profile solution, all models were estimated applying the manual three-step method (Asparouhov and Muthén, 2014; Morin and Litalien, 2017). Moreover, if the model with freely estimated coefficients between males and females was superior, we further checked pairwise differences of coefficients across profiles and gender using the model constraint function.

[1]We also tested models with freely estimated variance. However, these models did not converge, possibly due to an overparameterization (Chen et al., 2001). We therefore used fixed-variance models in all subsequent analyses.

RESULTS

Before proceeding to the hypothesis testing, we report the descriptive findings of self-concept and the achievement means of males and females (see **Table 1**). Correlations are reported in the Appendix (see **Appendix Tables A1**, **A2**). The *t*-test and effect sizes show that gender differences in favor of males occurred for self-concept in physics and math. The effect sizes for both are quite substantial.

Females showed higher English self-concept than males but the effect size shows that the difference had only little practical relevance. Many gender differences occurred regarding achievement. Males had significantly higher scores in standardized achievement tests in math and science whereas females showed higher scores in reading and English. However, the effect sizes indicated a medium effect for reading whereas the

TABLE 1 | Means of males' and females' self-concepts as well as achievement in different domains.

	Males (n = 714)	**Females (n = 944)**	**Cohen's *d***	**Mean Test between-gender difference**
Self-concept	**Mean (*SD*)**	**Mean (*SD*)**		***T*-Test**
Biology	3.32 (0.74)	3.31 (0.65)	0.01	0.29
Physics	3.19 (0.79)	2.72 (0.80)	0.59	11.97***
German	3.30 (0.50)	3.31 (0.49)	0.001	0.29
Math	3.33 (0.75)	2.85 (0.84)	0.59	11.91***
English	3.23 (0.75)	3.35 (0.68)	0.17	3.49***
Achievement				
Test scores				
Science	155.83 (31.23)	147.39 (28.24)	0.29	5.76***
Reading	150.94 (18.94)	159.79 (16.78)	0.50	10.05***
Math	165.03 (23.26)	159.22 (21.43)	0.26	5.27***
English	144.86 (20.72)	151.11 (18.26)	0.30	6.51***
Grades				
Biology	4.21 (0.84)	4.34 (0.84)	0.16	3.31***
Physics	4.23 (0.90)	4.26 (0.90)	0.03	0.69
German	4.05 (0.82)	4.44 (0.76)	0.49	9.85***
Math	4.13 (0.99)	4.16 (0.99)	0.04	0.77
English	3.94 (0.92)	4.31 (0.89)	0.41	8.26***

****$p < 0.001$; **$p < 0.01$; *$p < 0.05$; all achievement variables were measured in Grade 8.*

remaining domains showed rather small effect sizes. With regard to school grades, math and physics did not significantly differ across gender. Females received significantly better grades in the verbal domains.

Males' and Females' Self-Concept Profiles

We ran consecutive LPAs from one to six profiles. **Table 2** displays the fit indices, average probabilities of profile membership, and classification accuracy (entropy) for the subsamples males and females. The normality and local independence assumptions of LPA were met.

In both subsamples, the fit indices continuously decreased as the number of profiles increased up to the fifth profile. The cAIC, BIC, and SABIC increased, thus, the sixth profile was statistically unsatisfactory. However, we also inspected a graphical representation of the fit indices. The indices attenuated at the fourth profiles (see **Appendix Figures A1**, **A2**). Adding a fifth profile did not result in a further meaningful profile. In fact, this fifth profile resembled an existing profile (see **Appendix Figures A3**, **A4**). In both samples, the 4-profile solution showed sufficient classification accuracy (entropy). The average probabilities of profile membership in the profiles were high and ranged between 0.76 and 0.85 in the females' subsample and between 0.76 and 0.87 in the males' subsample. Thus, the models can distinguish quite well between the profiles. Taking these fit values into account, as well as the theoretical foundations of the profiles, and the inspection of the emerging profile in the 5-profile solution, the 4-profile solution was retained for subsequent analysis in both samples.

Analyses of Profile Similarities Between Males and Females

The LPA of each subsample revealed the same number of profiles for males and females. Thus, in a second step, we investigated the profile similarity between males and females using a multigroup approach. **Table 3** shows the fit indices for the models. First, a 4-profile model with freely estimated means across subsamples was estimated (configural model). The configural model was compared with a more restrictive model in which the means were held equal across gender (structural model).

The structural model resulted in lower cAIC and BIC values, indicating that the four profiles showed similar within-profile means across gender. Next, a model with means and within-profile variance held equal across gender was estimated (dispersion model). This model resulted in lower fit indices than the prior model and thus further supports the structural profile similarities across gender. Finally, the distributional model was estimated by constraining the profile sizes of the profiles to be equal across gender. All fit indices were higher than in the dispersion model, indicating that the relative size of the latent profiles was not similar across gender, which means that the profile membership differed as a function of gender.

The final multigroup 4-profile model, with equal within-profile means and within-profile variability across gender but freely estimated profile probabilities was retained for further analysis. **Figure 1** exhibits the profiles and **Table 4** the gender distribution across the four profiles.

The first profile reflects high math-related self-concepts (math and physics) and low verbal-related self-concepts (German and English). We labeled this profile *high mathrelated self-concept* (HMRSC). As we hypothesized, this group included 19% of the males but only 7% of the females. Even though we had refrained from formulating hypotheses regarding biology self-concept, we found that this domain-specific self-concept did not show directionality on the math-verbal continuum. The second profile (11% males and 18% females) comprised low levels of self-concept across all domains with rather identical physics, math, and reading self-concept. Biology self-concept, though still low, was slightly higher than the remaining self-concepts whereby English self-concept was slightly lower than all other self-concepts. We labeled this profile *low overall self-concept* (LOSC). The third profile showed differences between math-related and

TABLE 2 | Results from the latent profiles analyses.

Model	LL	# of par.	AIC	cAIC	BIC	SABIC	ϕ Profile prob.	entropy
Females								
1 profile	−6697.39	10	13414.78	13473.28	13463.28	13431.52	n.a.	n.a.
2 profile	−6520.61	16	13073.21	13166.81	13150.81	13100.00	>0.88	0.66
3 profile	−6441.81	22	12927.62	13056.32	13034.32	12964.45	>0.81	0.69
4 profile	−6403.64	28	12863.29	13027.09	12999.09	12910.17	>0.76	0.70
5 profile	−6360.91	34	12789.82	12988.72	12954.72	12846.74	>0.79	0.76
6 profile	−6352.72	40	12785.44	13019.44	12979.44	12852.40	>0.80	0.79
Males								
1 profile	−5065.61	10	10151.22	10206.93	10196.93	10165.18	n.a.	n.a.
2 profile	−4911.66	16	9855.32	9944.46	9928.46	9877.65	>0.90	0.69
3 profile	−4863.50	22	9771.00	9893.56	9871.56	9801.70	>0.81	0.69
4 profile	−4837.89	28	9731.78	9887.77	9859.77	9770.86	>0.76	0.72
5 profile	−4817.99	34	9703.98	9893.39	9859.39	9751.43	>0.69	0.68
6 profile	−4802.44	40	9684.89	9944.46	9867.72	9740.71	>0.71	0.70

LL, log likelihood; # of par., number of free parameters; AIC, Akaïke information criteria; CAIC, constant AIC; BIC: Bayesian information criterion; SABIC, sample-size adjusted BIC; n.a., not applicable.

TABLE 3 | Test of profile similarity, predictive similarity, and explanatory similarity.

	LL	# of par.	AIC	cAIC	BIC	SABIC
Profile similarity						
Configural: All means free	−12264.12	57	24642.23	25007.79	24950.80	24769.71
Structural: All means equal	−12318.68	37	24711.37	24948.66	24911.66	24794.20
Dispersion: All equal	−12328.13	32	24720.26	24925.49	24893.49	24791.83
Distributional	−12373.87	29	24805.74	24991.73	24962.73	24870.60
Predictive similarity: test score						
Freely estimated	−2897.35	31	5856.70	6055.52	6024.52	5926.04
Equality across gender	−2904.96	19	5847.91	5969.77	5950.77	5890.41
Predictive similarity: grades						
Freely estimated	−2706.43	37	5486.85	5724.15	5687.14	5569.60
Equality across gender	−2718.08	22	5480.16	5621.26	5599.26	5529.37
Explanatory similarity						
Freely estimated	−5989.74	25	12029.48	12189.810	12164.81	12085.39
Equality across gender	−6055.97	17	12145.95	12254.975	12237.98	12183.97

LL, log likelihood; # of par., number of free parameters; AIC, Akaike Information Criteria; CAIC, Constant AIC; BIC, Bayesian Information Criterion; SABIC, Sample-Size Adjusted BIC; n.a., not applicable.

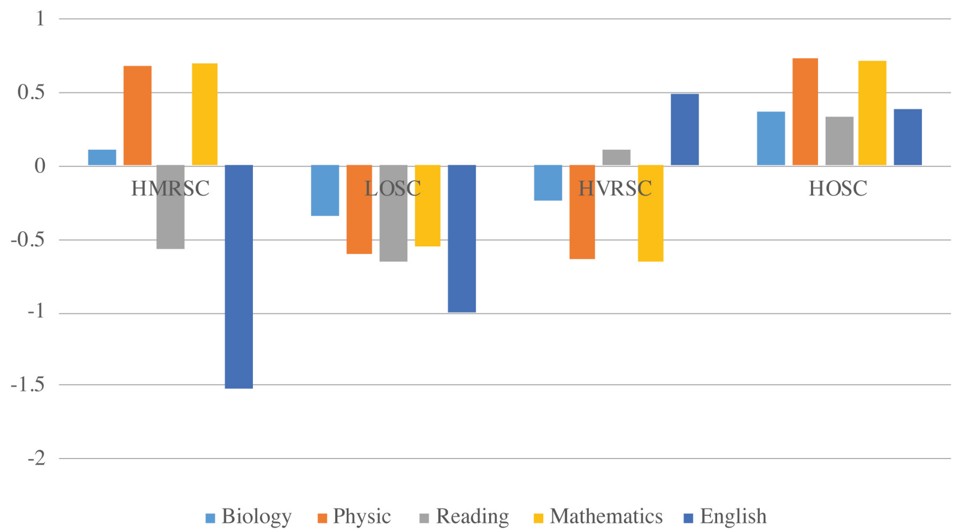


FIGURE 1 | Final profile solution. HMRSC, high math-related self-concept; LOSC, low overall self-concept; HVRSC, high verbal-related self-concept; HOSC, high overall self-concept profile.

verbal-related self-concepts and was somewhat the opposite of the HMRSC profile. Students in this profile had high verbal self-concepts (German and English) and low math-related self-concepts (physics and math). Self-concept in biology did not show directionality toward math-related self-concept on the verbal-math continuum. Thus, we labeled the profile *high verbal-related self-concept* HVRSC. As we expected, this profile was prevalent for almost half of the females (48%) but only for 28% of the males. Finally, the last profile constituted most of the students and contained 43% of the males' subsample and 28% of the females' subsample. The profile was characterized by high math and physics self-concept, which was slightly larger than reading, biology, and English self-concept. This configuration corresponds to a *high overall self-concept* profile (HOSC). As we had expected, these four profiles represent level (LOSC and HOSC) and shape differences (HMRSC and HVRSC). Therefore, again as we expected, we found profiles that differed both quantitatively and qualitatively.

We also conducted a chi-square analysis and adjusted standardized residuals to investigate gender differences in profile membership (see **Appendix Table A3**). As hypothesized, gender differences occurred in all profiles following a stereotypical pattern. Females were overrepresented in the HVRSC and the LOSC profile whereas males were overrepresented in the HMRSC profile and the HOSC.

Predictive Similarity

We added covariates to the final multigroup model to investigate the prediction of profile membership via multinomial regression. A multigroup approach was used to investigate whether

TABLE 4 | Gender distribution across profiles.

	High math-related self-concept	Overall low self-concept	High verbal-related self-concept	Overall high self-concept
Males	19%	11%	28%	43%
Females	7%	18%	48%	28%

covariates differentially predicted profile membership for males and females. We analyzed two sets of covariates: achievement in a standardized test and school grades. We again compared two models, a model freely estimating the predictors across gender and a model constraining the paths to be equal across gender. The constrained model resulted in lower fit values, indicating predictive similarity across gender. **Table 5** presents the results of the multinomial logistics regression. We report odds ratios (OR). ORs reflect the change in the likelihood of being in a profile versus a comparison profile for each unit of increase in the predictor. An OR of 1 indicates that the likelihood to be placed in a profile versus a comparison profile is equal, and an OR above 1 indicates a higher likelihood to be placed in the comparison profile.

Students with high achievement test scores in math and in science were more likely to belong to the HMRSC profile rather than to the LOSC profile or the HVRSC profile. Students with lower achievement test scores in math and science were more likely to belong to the HVRSC profile than to the HOSC profile. Students with high English achievement test scores were more likely to belong to the HVRSC profile than to the HOSC profile. Students with low English achievement test scores were more likely to be in the HMRSC profile than in any of the three others or were more likely to be in the LOSC profile than the HVRSC profile.

The effects of school grades on group membership revealed similar results. Again, students with higher grades in math and in physics were more likely to be members of the HMRSC profile in comparison to the LOSC profile and the HVRSC profile. Lower grades in math and in physics led to higher membership chances in the HVRSC profile and the LOSC profile compared to the HOSC profile. Students with better grades in German and in English were more likely to be a member of HVRSC profile than the HOSC profile. Students with lower grades in English were more likely to belong to the HMRSC in comparison to all other profiles. However, students with lower grades in English were also more likely to be in the LOSC profile than in the HOSC profile or the HVRSC profile. Regarding grades in biology, students with lower grades had a higher likelihood of membership in the HVRSC and in the LOSC profiles compared to the HOSC profile.

In a last step, we added outcome variables to the final model. Again, we compared a model with freely estimated mean-level outcomes with a model in which the mean-level outcomes were constrained to be equal across gender. We tested the models for each of the outcome variables. **Table 3** shows that the fit indices were lower for the freely estimated model, thereby indicating that course selection varies as a function of gender.

Pairwise mean differences were tested between the profiles and across gender (**Table 6**). The likelihood of choosing a science course at the beginning of upper secondary school was highest in the HMRSC profile and in the HOSC profile as compared to the LOSC and HVRSC self-concept profile. This was true for both males and females. However, in the HMRSC profile and in the HOSC profile, the probability of choosing a science course was lower for females than for males. We found the same pattern for science aspirations, indicating that students in the HMRSC and HOSC profile were more likely to choose a science career in the future. However, with regard to gender differences, only the HOSC profile revealed significant differences, indicating that males' aspirations for a science career were higher than females'.

DISCUSSION

The focus of our study was to explore gendered differences within self-concept profiles and their relation to pursuing a scientific path. In order to gain deeper insights into choices made regarding science at upper secondary school, we examined

TABLE 5 | Relations of achievement measures to self-concept profile membership (Multinomial logistic regression, separate analyses for test scores and school grades).

	HMRSC vs. LOSC	HMRSC vs. HVRSC	HMRSC vs. HOSC	LOSC vs. HOSC	LOSC vs. HVRSC	HVRSC vs. HOSC
	OR (SE_b)	OR (SE_b)	OR (SE_b)	OR (SE)	OR (SE_b)	OR (SE_b)
Achievement test scores						
Math	1.07 (0.02)***	1.07 (0.01)***	1.01 (0.01)	0.95 (0.01)***	1.00 (0.01)	0.95 (0.01)***
Reading	0.99 (0.01)	0.99 (0.01)	0.99 (0.01)	1.00 (0.01)	1.00 (0.01)	1.00 (0.01)
Science	1.03 (0.01)**	1.03 (0.01)***	1.00 (0.01)	0.98 (0.01)***	1.03 (0.01)*	0.97 (0.01)***
English	0.95 (0.02)**	0.89 (0.01)***	0.93 (0.01)***	0.96 (0.01)***	0.97 (0.01)**	1.04 (0.01)***
School grades						
Math	10.97 (0.41)***	10.79 (0.24)**	1.64 (0.24)*	0.26 (0.25)***	1.68 (0.23)*	0.15 (0.20)***
German	0.56 (0.30)***	0.51 (0.22)*	1.21 (0.21)	2.31 (0.23)***	0.99 (0.21)	2.35 (0.18)***
Biology	1.67 (0.31)	1.32 (0.22)	0.90 (0.23)	0.58 (0.22)*	0.85 (0.20)	0.69 (0.19)*
Physics	4.17 (0.33)***	5.28 (0.24)**	1.16 (0.20)	0.30 (0.22)***	1.36 (0.18)	0.22 (0.18)***
English	0.23 (0.47)***	0.06 (0.30)*	0.11 (0.29)***	0.28 (0.28)***	0.15 (0.22)*	1.94 (0.24)**

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; all achievement variables were measured in Grade 8. OR, odds ratio; SE_b, standard error of beta; HMRSC, high math-related self-concept; LOSC, low overall self-concept; HVRSC, high verbal-related self-concept; HOSC, high overall self-concept profile.

TABLE 6 | Associations between profile membership course selection and career aspirations in science.

Outcome	HMRSC (1)	LOSC (2)	HVRSC (3)	HOSC (4)
	OR	OR	OR	OR
Science course male	1.81 (0.11)[2,3]	1.20 (0.08)[1,4]	1.17 (0.03)[1,4]	1.52 (0.06)[2,3]
Science course female	1.32 (0.06)[2,3,a]	1.13 (0.03)[1,4]	1.13 (0.02)[1,4]	1.29 (0.03)[2,3,a]
	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)
Science aspirations male	2.94 (0.11)[2,3]	2.03 (0.17)[1,4]	2.00 (0.06)[1,4]	2.88 (0.08)[2,3]
Science aspirations female	2.91 (0.11)[2,3,4]	2.13 (0.11)[1,4]	1.89 (0.06)[1,4]	2.613 (0.06)[2,3,1,a]

[a]significant differences between gender, $p < 0.05$. Indices indicate the other profile that has a significantly different value compared to the respective profile, for example: 1.81[2,3] of HMRSC profile (1) is significantly higher than 1.20 of LOSC profile (2) and higher than 1.17 of HVRSC profile (3). HMRSC, high math-related self-concept; LOSC, low overall self-concept; HVRSC, high verbal-related self-concept; HOSC, high overall self-concept profile.

the joint effects of prior domain-specific self-concept (i.e., math, reading, physics, biology, and English) on course selection in secondary education. By adopting a person-centered approach, we explored how dimensional comparisons between these science and nonscience school subjects lead to science course selection instead of only taking a single domain into account. In contrast to previous studies, we incorporated these self-concepts to focus on science course selection and science aspirations, to systematically check for gender invariances within the person-centered approach, and to evaluate whether gender-specific differences in choosing a science path occur. We first conducted an LPA to analyze self-concept profiles separately for males and females and we then adopted a multigroup LPA to further investigate gender differences in profiles, profile membership, predictors, and outcomes.

Self-Concept Profiles of Males and Females

As expected, the LPA revealed four distinct profiles that differed qualitatively (in shape) and quantitatively (in level). Thus, already in Grade 8, students develop a differentiated pattern of science and nonscience self-concepts. Three profiles reflected the verbal-math continuum. The first profile is characterized by high self-concept in math and physics and low self-concept in the verbal domains (i.e., reading and English). The opposing profile reflected high verbal self-concept and low math self-concept. One profile reflected a high overall self-concept across all domains, but with slightly higher math-related self-concepts compared to verbal-related self-concepts. The remaining profile reflected overall low levels of self-concept across all domains. In all profiles, math and physics self-concepts had comparable means within each profile. Biology, a self-concept component we explored without formulating hypotheses, could not be related to any end of the verbal-math continuum. This self-concept showed different positioning in the profiles. It rather belonged to the math-related self-concepts in the HVRSC profile, to the high verbal-related self-concepts in the HOSC profile, whereas, in the HMRSC and the LOSC profile, no clear allocation to the continuum could be made. Therefore, our results did not support the assumption that biology self-concept belongs to either one of the ends of the verbal-math continuum. Regarding the placement of biology self-concept within the verbal-math continuum, we might have revealed a reason for the inconsistent results from previous variable-centered studies. Whether or not biology can be placed on the verbal end of the verbal-math continuum might depend on the students' self-concept profile. The nature of our profiles was similar to profiles identified in earlier studies; however, the number of profiles differed as the number of domain-specific self-concepts used in former studies differed as well (Marsh et al., 2009; Umarji et al., 2018). Thus, our results confirm the importance of an individual's self-evaluation of strengths and weaknesses, which has already been established in the context of the DCT (Möller and Marsh, 2013).

The shape of the profiles was similar across gender, indicating that both genders had the same structure of self-concept profiles. We did not find distinctive self-concept patterns for males and females. Even though we did not find gender distinctive patterns, as expected, we did find differences in the distribution across profiles as a function of gender. The HVRSC profile and the LOSC profile were more frequent among females than males. In contrast, males were overrepresented in the profiles with higher math-related self-concepts (i.e., the HOSC profile and the HMRSC profile). About 62% of the males either had a high overall self-concept or a HMRSC, while this was the case for less than 35% of the females. The distributional differences suggest that a specific group of males' and females' self-concept should be strengthened according to their profile as early as Grade 8. As this was only 1 year after the beginning of science teaching in Germany, our results show that self-concepts should be fostered from the very beginning of teaching these subjects. In order to keep females on the science path, specific interventions should foster the math and physics self-concepts of females in the HVRSC profile and in the LOSC profile.

Predictors

Stereotypical gender differences emerged in the achievement indicators. Although recent studies underline that gender differences in achievement almost disappear between males and females, contrary to the self-concepts, where we found more gender differences for physics and science, we found substantial differences between achievement indicators mainly for the verbal domains (i.e., German and English). Interestingly, we did not find any gender differences for grades in math and physics whereas males showed higher test scores in math and science than females. Moreover, the dimensional comparisons of the students' achievement seemed to influence membership in the self-concept profiles. Just as has been shown for math, English, and German, we again showed that students rank their achievement in comparison to other domains in order to get information about their relative abilities in a specific domain (Marsh, 1990; Möller and Marsh, 2013) and this comparison forms their self-concept pattern. Students with high math and low verbal abilities rather belong to the HMRSC profile than to the HVRSC profile and vice versa. Interestingly, students with low abilities in English fall into the HMRSC profile rather than into the HOSC. Therefore, English (as a foreign language) seems to be a more contrasting domain in finding one's own self-concept profile than the more general German (mother tongue) comparison. Hence, the perception of one's abilities in the foreign languages is a quite important (negative) predictor of females staying on the STEM path. Moreover, as expected, the grade in math is the most powerful predictor of membership in the HMRSC profile for both males and females. Even though males and females did not differ regarding their math grades, females were less likely to be in profiles with HMRSCs. One explanation might be the attributional gender bias: males' success in math is attributed to ability whereas females' success is attributed to effort (e.g., Räty et al., 2002). Perhaps, females might not use their actual math grade to examine their abilities. Thus, other mechanisms seem to influence gender-specific distribution across profiles. Exploring these mechanisms should be the focus of future studies.

Outcome

Students may be more likely to persevere in a subject in which they feel more competent than in another subject (Jonsson, 1999; Uerz et al., 2004). For instance, Marsh and Yeung (1997) were able to show that domain-specific self-concept predicts the choice of that particular subject more than self-concept in other subjects. Moreover, students tend to choose a science course more frequently when their abilities in math are clearly higher than their abilities in reading (Uerz et al., 2004). Our results support these assumptions. As we expected, students belonging to profiles with math-related self-concepts more frequently chose a science course and had higher aspirations toward a science career than students in the LOSC and the HVRSC profile. Because both profiles with math-related self-concepts drives science course selection, dimensional comparisons might lead and might not lead to science course selection at the time. Not only students with higher self-concept in mathematics compared to the verbal domain choose science courses but also students with high self-concept in both domains. This effect was even stronger for males in two ways. First, males tended to be in the two profiles favoring science course selection and science aspirations more often. Second, compared to females, males in the HMRSC or the HOSC profile more frequently chose a science course than females. In particular, females who perceived themselves as being good in both math and verbal subjects were less likely to choose science course than males. Having high abilities or self-concept in several domains leads to a broader range of possibilities in educational choices. Students might choose academic courses and later careers according to stereotypical interests or fields to which their environment directs them. For example, stereotypes regarding science as a male domain (Cheryan et al., 2011) and gender-role behaviors in students' environment (Eccles, 1993) might influence educational decisions. These mechanism for this specific group needs to be addressed in future studies in order to tailor future interventions to keep females on the scientific path. Even though our study did not lead to specific interventions, we were able to show that there cannot be one strategy to keep females on this path. Furthermore, we showed that – just as for males – the group of females dropping out of science careers is not a homogenous group, just as the group of females that pursues a scientific career is not homogenous. Understanding the mechanisms behind the specific groups is an important task for future interventions. Our conclusion that females within specific profiles should be supported in strengthening their math- and physics-related self-concepts should ideally also result in females' increased interest in a scientific career as well as in the choice of science courses during upper secondary education. Our study showed that already in Grade 8 self-concept predicts career aspirations. Although we were able to reveal this mechanism for STEM paths, future studies could investigate whether these mechanisms also appear for other paths such as choosing humanities or languages for males and for females.

Limitation and Future Directions

Some limitations, which future studies could attempt to overcome, are worth mentioning. First, our study focused on general science course selection. The subsamples of students who chose advanced physics and chemistry courses were too small because these two school subjects are not very popular in Germany. Therefore, analyses of these fine-grained levels were not feasible. Future studies might look at how self-concept patterns (e.g., covering the differential self-concepts of our study) differentially influence the uptake of physics, biology, and chemistry courses. It would be interesting to investigate the impact of self-concepts that include high math contents (such as physics, which we were able to place on the verbal-math continuum, in comparison to biology) on science course selection.

Some methodological limitations are to mention. Similar to previous studies, we assessed domain self-concept with separate domain-specific Likert scales (i.e., for biology, physics, math, English, and German). Additionally, we could not rely on paralleled items across all scales. This approach may have the drawback that it was not able to measure exactly the same construct in all of the domains. This is a methodological problem that should be addressed in the future, especially in the light of DCT studies. Further, the entropy values in the female and male subsamples were about 0.70 suggesting that precision of classification might not be highly accurate. However, the average posterior probabilities ranged between 0.76 and 0.88. The cross probabilities values, which ranged between 0.01 and 0.16, indicated that students had also a 1–16% chance of belonging to another profile. Thus, the chosen model can distinguish quite well between the profiles.

Future studies could also focus on the predictors of self-concept patterns. We showed that individuals' achievement is a source of self-concept formation but peers, teachers, and family also play an important role in perceptions of self-concept. Previous studies have already demonstrated that school and family influence career choices (Mau and Bikos, 2000; Fouad and Santana, 2017). The shaping of self-concept can provide valuable information about why females are less likely to choose STEM. Moreover, future studies could incorporate a wide array of motivational variables that might explain gender differences in course selection, such as task values, interest, and self-efficacy (Guo et al., 2015; Korhonen et al., 2016; Lin et al., 2018; Linnenbrink-Garcia et al., 2018) and these individual predictors could be combined with school, family, and peer predictors. Incorporating these additional variables and utilizing the DCT as well as the person-centered and the multigroup approach could reveal whether the same mechanisms apply to all motivational variables for males and females alike.

Lastly, we focused on the impact of self-concept on educational decisions 2 years before students decide upon their courses and basically right after the start of science teaching in Germany. Previous studies have shown that early self-concept already impacts later educational decisions (Nagy et al., 2006; Umarji et al., 2018). For instance, the effects of math and English self-concepts on math intensive college majors were rather similar for the students in the 7th and 10th grades (Umarji et al., 2018). Moreover, there is evidence that motivational profiles are rather stabile from Grades 9 to 10 (e.g., Lazarides et al., 2019). Further, our study did not take self-concept development throughout secondary level into account. On the basis of our results, studies incorporating science and nonscience self-concepts into growth mixture models and using several measurement points in lower secondary education would be suitable to identify subgroups of students that differ in their growth rates of achievement or in their self-concept patterns. One study has already been able to show that students who increased their math self-concept throughout secondary school were more likely to choose a math intensive course than students whose math self-concept declined over the course of secondary school (Musu-Gillette et al., 2015). Those results could be expanded by using a wider array of self-concept variable patterns and by focusing on science course selection, which could then be linked to several outcomes that we used for our study. Thus, future research could thereby provide deeper insights into the substantial changes that might occur in the development of self-concept in different domains and that could influence later science course selection.

CONCLUSION

Our study contributes to the body of research on self-concept by incorporating self-concept in several domains and the DCT to understand gendered paths into STEM. To the best of our knowledge, this is the first study that separately tested the self-concept profiles of females and males in upper secondary school in a person-centered and multigroup framework in order to disentangle gendered pathways to academic choices regarding science. Our study further provides evidence for the robustness of former findings regarding multidimensional self-concept profiles using LPA.

ETHICS STATEMENT

Participation in the achievement tests was mandatory for all students with the consent of the Ministry of Schools and Vocational Training of the Free and Hanseatic City of Hamburg, whereas participation in the questionnaires was voluntary. This study was carried out in accordance with the ethical guidelines for research with human participants as proposed by the American Psychological Association (APA). The study materials and procedures were approved by the Ministry of Schools and Vocational Training of the Free and Hanseatic City of Hamburg.

AUTHOR CONTRIBUTIONS

SS wrote the first and consecutive drafts of the manuscript. Both the authors listed have made a substantial, direct and intellectual contribution to the work, including the approval for publication.

ACKNOWLEDGMENTS

This manuscript uses data from the longitudinal studies LAU and/or KESS. Both data sets were generated by the Free and Hanseatic City of Hamburg through the Ministry of Schools and Vocational Training between 1995 and 2012 and were provided to the MILES (Methodological Issues in Longitudinal Educational Studies) scientific consortium for a limited period with the aim of conducting in-depth examinations of scientific questions. MILES is coordinated by the Leibniz Institute for Science and Mathematics Education (IPN).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00836/full#supplementary-material

REFERENCES

Asparouhov, T., and Muthén, B. (2014). Auxiliary variables in mixture modeling: three-step approaches using *Mplus*. *Struct. Equ. Modeling* 21, 329–341. doi: 10.1080/10705511.2014.915181

Ayalon, H. (2003). Women and men go to university: mathematical background and gender differences in choice of field in higher education. *Sex Roles* 48, 277–290. doi: 10.1023/A:1022829522556

Bos, W., and Gröhlich, C. (eds) (2010). *KESS 8. Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8 [Competencies and Attitudes of Students at the End of Grade 8]*. Münster: Waxmann.

Bos, W., Gröhlich, C., Dudas, D. F., Guill, K., and Scharenberg, K. (2010). *KESS 8. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.

Britner, S. L. (2008). Motivation in high school science students: a comparison of gender differences in life, physical, and earth science classes. *J. Res. Sci. Teach.* 45, 955–970. doi: 10.1002/tea.20249

Chen, F., Bollen, K. A., Paxton, P., Curran, P. J., and Kirby, J. B. (2001). Improper solutions in structural equation models: causes, consequences, and strategies. *Sociol. Methods Res.* 29, 468–508. doi: 10.1177/0049124101029004003

Cheryan, S., Siy, J., Vichayapai, M., Drury, B., and Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Soc. Psychol. Pers. Sci.* 2, 656–664. doi: 10.1177/1948550611405218

Davier, M. V., and Davier, A. A. V. (2007). A unified approach to IRT scale linking and scale transformations. *Methodology* 3, 115–124. doi: 10.1027/1614-2241.3.3.115

DeWitt, J., Osborne, J., Archer, L., Dillon, J., Willis, B., and Wong, B. (2013). Young children's aspirations in science: the unequivocal, the uncertain and the unthinkable. *Int. J. Sci. Educ.* 35, 1037–1063. doi: 10.1080/09500693.2011.608197

Eccles, J. S. (1993). "School and family effects on the ontogeny of children's interests, self-perceptions, and activity choice," in *Nebraska Symposium on Motivation, 1992: Developmental perspectives on motivation*, ed. J. Jacobs (Lincoln, NB: University of Nebraska Press), 145–208.

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York, NY: Guilford Publications.

Fouad, N. A., and Santana, M. C. (2017). SCCT and underrepresented populations in STEM fields: moving the needle. *J. Career Assess.* 25, 24–39. doi: 10.1177/1069072716658324

Graham, J. W., Olchowski, A. E., and Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev. Sci.* 8, 206–213. doi: 10.1007/s11121-007-0070-9

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., and Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: internal comparison processes and expectancy-by-value interactions. *Learn. Instr.* 49, 81–91. doi: 10.1016/j.learninstruc.2016.12.007

Guo, J., Parker, P. D., Marsh, H. W., and Morin, A. J. S. (2015). Achievement, motivation, and educational choices: a longitudinal study of expectancy and value using a multiplicative perspective. *Dev. Psychol.* 51, 1163–1176. doi: 10.1037/a0039440

Haag, L., and Götz, T. (2012). Mathe ist schwierig und deutsch aktuell: vergleichende studie zur charakterisierung von schulfächern aus schülersicht [math is difficult and german up to date: a comparative study on the characterization of subject domains from students' perspective]. *Psychologie in Erziehung und Unterricht* 59, 32–46. doi: 10.2378/peu2012.art03d

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., and Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychol. Sci. Public Inter.* 8, 1–51. doi: 10.1111/j.1529-1006.2007.00032.x

Helm, F., Müller-Kalthoff, H., Nagy, N., and Möller, J. (2016). Dimensional comparison theory: perceived subject similarity impacts on students' self-concepts. *AERA Open* 2, 1–9. doi: 10.1177/2332858416650624

Herbert, J., and Stipek, D. (2005). The emergence of gender difference in children's perceptions of their academic competence. *J. Appl. Dev. Psychol.* 26, 276–295. doi: 10.1016/j.appdev.2005.02.007

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., and Trezise, K. (2018). Informative tools for characterizing individual differences in learning: latent class, latent profile, and latent transition analysis. *Learn. Individ. Diff.* 66, 4–15. doi: 10.1016/j.lindif.2017.11.001

Hyde, J. S. (2005). The gender similarities hypothesis. *Am. Psychol.* 60, 581–592. doi: 10.1037/0003-066X.60.6.581

Jansen, M., Schroeders, U., and Lüdtke, O. (2014). Academic self-concept in science: multidimensionality, relations to achievement measures, and gender differences. *Learn. Individ. Differ.* 30, 11–21. doi: 10.1016/j.lindif.2013.12.003

Jansen, M., Schroeders, U., Lüdtke, O., and Marsh, H. W. (2015). Contrast and assimilation effects of dimensional comparisons in five subjects: an extension of the I/E model. *J. Educ. Psychol.* 107, 1086–1101. doi: 10.1037/edu0000021

Jerusalem, M. (1984). *Selbstbezogene Kognitionen in Schulischen Bezugsgruppen [Self-referential Cognition in School Reference Groups]*. Berlin: Freie Universität.

Jonsson, J. O. (1999). Explaining sex differences in educational choice an empirical assessment of a rational choice model. *Eur. Sociol. Rev.* 15, 391–404. doi: 10.1093/oxfordjournals.esr.a018272

Kelly, D. L. (2003). "Developing the PIRLS background questionnaires," in *PIRLS 2001 Technical Report*, eds M. O. Martin, I. V. S. Mullis, and A. M. Kennedy (Chestnut Hill, MA: Boston College), 29–40.

Köller, O., Daniels, Z., Schnabel, K. U., and Baumert, J. (2000). Kurswahlen von mädchen und jungen im fach mathematik: zur rolle von fachspezifischem selbstkonzept und interesse [math course selection by boys and girls: the role of domain-specific self-concept and interest]. *Zeitschrift für Pädagogische Psychologie* 14, 26–37. doi: 10.1024//1010-0652.14.1.26

Korhonen, J., Tapola, A., Linnanmäki, K., and Aunio, P. (2016). Gendered pathways to educational aspirations: the role of academic self-concept, school burnout, achievement and interest in mathematics and reading. *Learn. Instr.* 46, 21–33. doi: 10.1016/j.learninstruc.2016.08.006

Lazarides, R., Dietrich, J., and Taskinen, P. (2019). Stability and change in students' motivational profiles in mathematics: the role of perceived teaching. *Teach. Teach. Educ.* 79, 164–175. doi: 10.1016/j.tate.2018.12.016

Lazarides, R., Viljaranta, J., Aunola-Aro, K., Pesu, L., and Nurmi, J.-E. (2016). The role of parental expectations and students' motivational profiles for educational aspirations. *Learn. Individ. Differ.* 51, 29–36. doi: 10.1016/j.lindif.2016.08.024

Lin, L., Lee, T., and Snyder, L. A. (2018). Math self-efficacy and STEM intentions: a person-centered approach. *Front. Psychol.* 9:2033. doi: 10.3389/fpsyg.2018.02033

Lindberg, S. M., Hyde, J. S., Petersen, J. L., and Linn, M. C. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psychol. Bull.* 136, 1123–1135. doi: 10.1037/a0021276

Linnenbrink-Garcia, L., Wormington, S. V., Snyder, K. E., Riggsbee, J., Perez, T., Ben-Eliyahu, A., et al. (2018). Multiple pathways to success: an examination of integrative motivational profiles among upper elementary and college students. *J. Educ. Psychol.* 110, 1026–1048. doi: 10.1037/edu0000245

Lörz, M., Schindler, S., and Walter, J. (2011). Gender Inequalities in higher education. Extent, development and mechanisms of gender differences in enrolment and field of study choice. *Irish Educ. Stud.* 30, 179–198. doi: 10.1080/03323315.2011.569139

Lubke, G. H., and Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychol. Methods* 10, 21–39. doi: 10.1037/1082-989X.10.1.21

Marsh, H. W. (1990). A multidimensional, hierarchical model of self-concept: theoretical and empirical justification. *Educ. Psychol. Rev.* 2, 77–172. doi: 10.1007/BF01322177

Marsh, H. W. (2007). *Self-Concept Theory, Measurement and Research Into Practice: The Role of self-Concept in Educational Psychology*. Leicester: British Psychological Society.

Marsh, H. W., Byrne, B. M., and Shavelson, R. J. (1988). A multifaceted academic self-concept: its hierarchical structure and its relation to academic achievement. *J. Educ. Psychol.* 80, 366–380. doi: 10.1037/0022-0663.80.3.366

Marsh, H. W., and Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: beyond seductive pleasure and unidimensional perspectives. *Perspect. Psychol. Sci.* 1, 133–163. doi: 10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., Lüdtke, O., Trautwein, U., and Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: synergy of person- and variable-centered approaches to theoretical models of self-concept. *Struct. Equ. Modeling* 16, 191–225. doi: 10.1080/10705510902751010

Marsh, H. W., and Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: gender differences in the development of math and English constructs. *Am. Educ. Res. J.* 35, 705–738. doi: 10.3102/00028312035004705

Marsh, H. W., and Yeung, S. A. (1997). Coursework selection: relations to academic self-concept and achievement. *Am. Educ. Res. J.* 34, 691–720. doi: 10.3102/00028312034004691

Mau, W. C., and Bikos, L. H. (2000). Educational and vocational aspirations of minority and female students: a longitudinal study. *J. Couns. Dev.* 78, 186–194. doi: 10.1002/j.1556-6676.2000.tb02577.x

Möller, J., Helm, F., Müller-Kalthoff, H., Nagy, N., and Marsh, H. W. (2015). Dimensional comparisons and their consequences for self-concept, motivation, and emotion. *Int. Encycl. Soc. Behav. Sci.* 26, 430–436. doi: 10.1016/B978-0-08-097086-8.26092-3

Möller, J., and Köller, O. (2004). Die genese akademischer selbstkonzepte: effekte dimensionaler und sozialer vergleiche [on the development of academic self-concepts: the impact of social and dimensional comparisons]. *Psychologische Rundschau* 55, 19–27. doi: 10.1026/0033-3042.55.1.19

Möller, J., and Marsh, H. W. (2013). Dimensional comparison theory. *Psychol. Rev.* 120, 544–560. doi: 10.1037/a0032459

Morin, A. J. S. (2016). "Person-centered research strategies in commitment research," in *The Handbook of Employee Commitment*, ed. J. P. Meyer (Cheltenham: Edward Elgar), 490–508. doi: 10.4337/9781784711740.00050

Morin, A. J. S., Boudrias, J.-S., Marsh, H. W., Madore, I., and Desrumaux, P. (2016a). Further reflections on disentangling shape and level effects in person-centered analyses: an illustration exploring the dimensionality of psychological health. *Struct. Equ. Modeling* 23, 438–454. doi: 10.1080/10705511.2015.1116077

Morin, A. J. S., and Litalien, D. (2017). *Webnote: Longitudinal Tests of Profile Similarity and Latent Transition Analyses*. Montreal, QC: Substantive Methodological Synergy Research.

Morin, A. J. S., Meyer, J. P., Creusier, J., and Biétry, F. (2016b). Multiple-group analysis of similarity in latent profile solutions. *Organ. Res. Methods* 19, 231–254. doi: 10.1177/1094428115621148

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., and Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educ. Res. Eval.* 21, 343–370. doi: 10.1080/13803611.2015.1057161

Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling: comment on bauer and curran (2003). *Psychol. Methods* 8, 369–377. doi: 10.1037/1082-989X.8.3.369

Muthén, L. K., and Muthén, B. O. (1998–2015). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

Nagengast, B., and Marsh, H. W. (2012). Big fish in little ponds aspire more: mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *J. Educ. Psychol.* 104, 1033–1053. doi: 10.1037/a0027697

Nagin, D. S. (2005). *Group-based Modeling of Development*. Harvard: Harvard University Press. doi: 10.4159/9780674041318

Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., and Eccles, J. (2008). "Gendered high school course selection as a precursor of gendered careers: the mediating role of self-concept and intrinsic value," in *Gender and Occupational Outcomes: Longitudinal Assessment of Individual, Social, and Cultural Influences*, eds H. Watt and J. Eccles (Washington: American Psychological Association), 115–143.

Nagy, G., Trautwein, U., Baumert, J., Köller, O., and Garrett, J. (2006). Gender and course selection in upper secondary education: effects of academic self-concept and intrinsic value. *Educ. Res. Eval.* 12, 323–345. doi: 10.1080/13803610600765687

National Science Foundation (2003). *August 14. The science and engineering workforce: Realizing America's potential (NSB 03-69)*. Washington, DC: National Science Foundation.

National Science Foundation (2015a). *Revisiting the STEM Workforce: A Companion to Science and Engineering Indicators 2014. (NSB*-2015-10). Arlington, VA: National Science Foundation.

National Science Foundation (2015b). *Women, minorities, and persons with disabilities in science and engineering: 2015 (Special Report NSF 15-311)*. Arlington, VA: National Science Foundation, National Center for Science and Engineering Statistics.

Nylund, K. L., Asparouhov, T., and Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Struct. Equ. Modeling* 14, 535–569. doi: 10.1080/10705510701575396

Organisation for Economic Co-operation and Development [OECD] (2006). *Women in Scientific Careers: Unleashing the Potential*. Paris: Organisation for Economic Co-operation, and Development.

Organisation for Economic Co-operation and Development (2016). *Data on Education by Sex and Country. In Education at a Glance*. Paris: Organisation for Economic Co-operation, and Development.

Parker, P., Schoon, I., Tsai, Y., Nagy, G., Trautwein, U., and Eccles, J. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multicontext study. *Dev. Psychol.* 48, 1629–1642. doi: 10.1037/a0029167

Perez-Felkner, L., Nix, S., and Thomas, K. (2017). Gendered pathways: how mathematics ability beliefs shape secondary and postsecondary course and degree field choices. *Front. Psychol.* 8:386. doi: 10.3389/fpsyg.2017.00386

Pinxten, M., de Fraine, B., van den Noortgate, W., van Damme, J., and Anumendem, D. (2012). Educational choice in secondary school in Flanders: the relative impact of occupational interests on option choice. *Educ. Res. Eval.* 18, 541–569. doi: 10.1080/13803611.2012.702991

Räty, H., Vänskä, J., Kasanen, K., and Kärkkäinen, R. (2002). Parents' explanations of their child's performance in mathematics and reading: a replication and extension of Yee and Eccles. *Sex Roles* 46, 121–128. doi: 10.1023/A:1016573627828

Schilling, R. S., Sparfeldt, J., and Rost, D. (2006). Facetten schulischen selbstkonzepts: welchen unterschied macht das geschlecht? [gender differences in subject-specific academic self-concepts]. *Zeitschrift für Pädagogische Psychologie* 20, 9–18. doi: 10.1024/1010-0652.20.12.9

Sikora, J., and Pokropek, A. (2012). Gender segregation of adolescent science career plans in 50 countries. *Sci. Educ.* 96, 234–264. doi: 10.3102/0002831211435229

Smyth, E., and Hannan, C. (2006). School effects and subject choice: the uptake of scientific subjects in Ireland. *Sch. Effect. Sch. Improv.* 12, 303–327. doi: 10.1080/09243450600616168

Spinath, B., Eckert, C., and Steinmayr, R. (2014). Gender differences in school success: what are the roles of students' intelligence, personality and motivation? *Educ. Res.* 56, 230–243. doi: 10.1016/j.paid.2013.07.131

Tofighi, D., and Enders, C. K. (2008). "Identifying the correct number of classes in growth mixture models," in *Advances in Latent Variable Mixture Models*, eds G. R. Hancock and K. M. Samuelsen (Charlotte, NC: Information Age Publishing), 317–341.

Trusty, J. (2002). Effects of high school course-taking and other variables on choice of science and mathematics college majors. *J. Counsel. Dev.* 80, 464–474. doi: 10.1002/j.1556-6678.2002.tb00213.x

Uerz, D., Dekkers, H., and Béguin, A. A. (2004). Mathematics and language skills and the choice of science subjects in secondary education. *Educ. Res. Eval.* 10, 163–182. doi: 10.1076/edre.10.2.163.27908

Umarji, O., Mcpartlan, P., and Eccles, J. (2018). Patterns of math and english self-concepts as motivation for college major selection. *Contemp. Educ. Psychol.* 53, 146–158. doi: 10.1016/j.cedpsych.2018.03.004

Wang, M. T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001

Wang, M.-T., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice: individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937

Watt, H. M. G. (2006). The role of motivation in gendered educational and occupational trajectories related to maths. *Educ. Res. Eval.* 12, 305–322. doi: 10.1080/13803610600765562

Wilkins, J. L. M. (2004). Mathematics and science self-concept: an international investigation. *J. Exp. Educ.* 72, 331–346. doi: 10.3200/JEXE.72.4.331-346

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 30 April 2019
doi: 10.3389/fpsyg.2019.00897

The Interest Profiles and Interest Congruence of Male and Female Students in STEM and Non-STEM Fields

Bernhard Ertl* and Florian G. Hartmann

Department for Education, Universität der Bundeswehr München, Neubiberg, Germany

The goal of the following study is to investigate whether first-year students in STEM fields that have a low proportion of females (STEM-L) show vocational interests that fit their vocational aspirations. To place our investigation into a broader context, we compared students in STEM-L with students of STEM subjects with a medium proportion of women (STEM-M) as well as with other subjects with a medium or a high proportion of females. We analyzed their vocational interests, vocational aspirations and their interest congruence. In both the comparison regarding interest profiles and the comparison of vocational aspirations, we focused on the things-orientation and people-orientation, all while taking respective gender differences into account. Following the suggestion from previous studies, in a further step we differentiated between subjects within STEM-L. Using data from the German National Educational Panel Study (NEPS), we analyzed the interest congruence of 5,530 male and 7,406 female students in STEM majors (with a low or medium proportion of women) and non-STEM majors (with a medium or high proportion of women). Students from different subjects showed different magnitudes regarding their things- and people-orientation. STEM-L students had a high things-orientation and a low people-orientation regarding both their interests and aspired occupations. Students of STEM-L and STEM-M showed a lower interest congruence than students from other subjects. With the exception of education, gender differences regarding the people- and things-orientation also existed within most of the subjects. Gender differences partly remain when distinguishing between the different subjects within STEM-L. And so, the result that not all STEM-L subjects are "created equal" is discussed in the context of their theoretical and methodological aspects.

Keywords: STEM students, non-STEM students, vocational interests, RIASEC, choice of major, congruence, O*NET, large scale study

OPEN ACCESS

Edited by:
Michael S. Dempsey, Boston University, United States

Reviewed by:
Evangelia Karagiannopoulou, University of Ioannina, Greece
JenJang Sheu, National Chung Hsing University, Taiwan

****Correspondence:***
Bernhard Ertl
bernhard.ertl@unibw.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *26 December 2018*
Accepted: *04 April 2019*
Published: *30 April 2019*

Citation:
Ertl B and Hartmann FG (2019) The Interest Profiles and Interest Congruence of Male and Female Students in STEM and Non-STEM Fields. Front. Psychol. 10:897. doi: 10.3389/fpsyg.2019.00897

INTRODUCTION: CONTEXTS OF FEMALES IN STEM

Science and technology are drivers of societal benefit. This is one of the reasons why the EU Commission established the goal of increasing the number of STEM (**s**cience, **t**echnology, **e**ngineering, and **m**athematics) graduates (Hingel et al., 2008, p. 10). However, the proportion of students enrolled in STEM subjects hasn't noticeably changed within the past decade (Eurostat, 2018), and the proportion of females in STEM remains low (Eurostat, 2018). This phenomenon requires a deeper look into the STEM subjects and the students (both female and male) who take them to potentially identify any peculiarities that may help achieve measures to encourage more students and especially more females to pursue a career in STEM.

When first observing the notion of STEM, it's important to acknowledge that this term is somewhat blurry with respect to its subjects. Although the EU Commission (Hingel et al., 2008) only focuses on (physical) science, technology (including engineering), and mathematics, newer publications (European Commission, 2015) also include life sciences such as medicine (also Eccles and Wang, 2016). Publications from the United States (e.g., Su and Rounds, 2015) furthermore include social sciences. These gray areas provide challenges when analyzing gendered pathways into STEM. Engineering subjects have low proportions of females, often far less than 30% (see Su and Rounds, 2015; Destatis [Statistisches Bundesamt], 2018b). In contrast, social sciences like education have very high proportions of females, ranging in some cases above 70% (ibid.). So, any analysis that focuses on females in STEM needs to acknowledge that there is variance even within the STEM fields themselves.

Recent research has often clustered STEM subjects according to different criteria to deal with this heterogeneity and achieve more differentiated insights into the characteristics of STEM students (e.g., Ertl et al., 2014, 2017; Eccles and Wang, 2016; Watt et al., 2017). For example, Eccles and Wang (2016) grouped health, biological, and medical sciences, contrasting them with mathematics, physics, engineering, and computer sciences. They found that differences in subject choice for one of both groups resulted primarily from gender differences in occupational and lifestyle values. Ertl et al. (2014) distinguished STEM subjects with respect to their proportions of females, finding differences with respect to motivation, academic self-concept, and the impact of stereotypes. Watt et al. (2017) focused the subject groups of mathematics, physics, chemistry, and biology, revealing differences in gendered processes of influence by prior mathematical performance, motivation, and mothers' perceptions. Recent research in sum has found differences in the student characteristics between the different STEM subjects or subject groupings.

Su and Rounds (2015) followed an approach of classifying STEM subjects very narrowly according to an interest profile that distinguishes the dimensions of *things* and *people*. Based on this profile, they projected a proportion of females for different STEM subjects and compared this with their actual proportion. This projection fit quite well in several cases, even though it showed a noticeable deviation in the fields of engineering and computer sciences in which it overestimated the proportion of women. This was also the case in applied mathematics and medical services where it underestimated the proportion of females (Su and Rounds, 2015, p. 15). Thus, they were able to show on a macro/subject level that the interest profile of a subject corresponds to the proportion of females within it.

This paper examines the vocational interests of individuals on a micro level and aims at revealing characteristics of interest profiles and their fit to aspired occupations. It will classify students' vocational interests according to the RIASEC model of Holland (1997) and compare these profiles with the RIASEC profiles of the respective students' vocational aspiration. The paper will apply a vector-based measure of congruence according to Eder (1998) that considers the Euclidean distance of the two interest vectors as congruence measure (see Prediger, 1982; Tracey and Sodano, 2013). This approach will reveal how far the individual has an interest congruence with his or her vocational aspiration, providing a more individual perspective than the comparison of subject-level interest profiles and proportions of females.

This investigation acknowledges the reported differences of students in STEM subjects. These may be a result of the proportion of females within a subject and/or its broader field. The paper will first classify subjects at the finest possible level according to their proportion of females: low (with a proportion of females less than 30%), medium (with a proportion of females between 30 and 70%), and high (with a proportion of females higher than 70%). It will then distinguish the broader subject area. The focus of this paper is on the STEM subjects of physical sciences, technology, engineering, and mathematics. However, for a point of comparison, to compare the results with previous research, and to interpret the results better within a broader context, the paper will also include the subject groups of medicine, economics, education, and languages.

VOCATIONAL INTERESTS AND CAREER CHOICES

Dealing with the issue of why so few females go into STEM means starting with models of why individuals decide on specific career paths. The following will briefly introduce three theoretical models that have different emphases on explaining these career paths and at the same time are well compatible with each other. We will first describe Holland's (1997) theory of occupational choice that deals with an individual's interests as the basis for selecting an appropriate (congruent) occupation. This model emphasizes the aspect of vocational interests as a crucial factor for career decisions. The Gottfredson (1981, 2005) model of circumscription and compromise highlights the developmental process behind the individual's occupational choice. Her model introduces the issue of sex-type regarding different occupations that may shape an individual's choice based on an evaluation of how far an occupation may be considered "typical" for males or females. Here, the model describes the exclusion of what are considered "inappropriate" occupations at an early developmental stage. As a third approach, we will cite the social cognitive career theory (SCCT; Lent et al., 1994; Lent, 2013; Lent and Brown, 2013) that considers the perspectives of person-environment fit approaches (e.g., Holland, 1997) as well as developmental career theories (e.g., Gottfredson, 1981, 2005) which additionally includes cognitive variables being less stable and therefore more malleable than personality dispositions (Lent, 2013; c.f. Hartmann, 2018).

Types of Vocational Interests

Holland's (1997) theory of occupational choice focuses on vocational interests and distinguishes six ideal types: Realistic (R), Investigative (I), Artistic (A), Social (S), Enterprising (E), and Conventional (C). These six types are not only used to describe an individual's personality including vocational interests. They also characterize potential work environments. According to Holland (1997), people seek work environments that fit their vocational interests. This means that people who resemble the realistic (R) personality type are interested in mechanical or technical activities. They prefer working with tools or machines. Therefore, they are supposed to choose a realistic occupation that includes these tasks and objects. For example, given realistic interests, it would be a consistent occupational choice to become a surveyor or radiologist. The investigative (I) type are people interested in mathematical and scientific activities. They prefer occupations like aerospace engineer or general internist. Artistic (A) individuals are interested in creative and artistic activities. Occupations such as architectural drafter or geneticist are well suited to their type. Social (S) people are interested in activities that emphasize social interaction and interpersonal relations. They enjoy teaching or helping other people. Following their interests, they could become music therapists or midwives. Enterprising (E) persons are interested in leading and convincing other people. They prefer occupations like clinical research coordinators or natural sciences managers. People resembling the conventional (C) type follow their preference for ordering and repetitive tasks and choose occupations like actuary or electronic drafter. All these occupations can be assigned to the broadest context of STEM fields (O*NET OnLine, 2018), making clear that STEM occupations and their corresponding training paths can be very different in terms of the required vocational interests (see also Su and Rounds, 2015). According to Holland (1997) students who have interests in STEM fields should choose STEM work environments as their courses of study, and later on as their occupations. The inverse is also true: STEM work environments should choose students or graduates who display STEM interests. Empirical evidence indicates that although investigative interests are crucial for going into STEM, very different interest profiles can cause people to choose STEM subjects or occupations; these are not all created equal, especially when it comes to the question of whether the work environments are typically male or female. In other words, STEM fields have gender differences both within specific subject-related interests as well as differences in the actual percentage of women. Both gender-related differences can be explained by the things-orientation (R) and the people-orientation (S) of the STEM fields (Prediger, 1982), with the things-orientation attracting more men and the people-orientation attracting more women. Consequently, female students tend to be interested in and choose STEM fields that are people-oriented and avoid STEM fields that are things-orientated (Su et al., 2009; Su and Rounds, 2015).

When it comes to the question of how vocational interests emerge, Holland (1997) describes a rather general model in which an individual's career-related development is based on the interplay of genes and environmental influences.

Development of Vocational Aspirations

Gottfredson (1981, 2005) looked deeper into the development of vocational aspirations, providing a specific explanation for the development of gender-specific differences regarding the choice of occupations. In her model of circumscription and compromise, an individual's career choice is described as a process that is mainly based on two personal developmental processes: the individual's cognitive growth, which is the development of cognitive skills, and the self-creation, which is the individual's development of a self-concept. The development of the self-concept involves the gradual development of cognitive skills, as well as the successive exclusion of occupations that are are no longer compatible with the current self-concept. At an early developmental stage (orientation to sex roles; 6–8 years of age) children classify people as well as work environments as "male" or "female." They become aware of their own gender and exclude occupations that are too male or too female and therefore outside of their tolerable sex-type boundary. Individuals then exclude occupations whose prestige is too low (tolerable-level boundary) and whose requirements are too high (tolerable-effort boundary). Gottfredson (2005) claims that the tolerable-level boundary is more important for men's career choices because typically male occupations show a higher variance with regard to their prestige level than typical female occupations. All three boundaries together define the individual's social space on the cognitive map of occupations within which occupations are eventually chosen based on personal interests in the sense of Holland (1997). If an individual is not able to find a work environment that is compatible with his or her specific self-concept, the restrictions are solved successively, with the tolerable sex-type boundary being perpetuated most strongly. Along with the mere congruence of occupational interests and occupational characteristics of the Holland (1997) model, Gottfredson (1981, 2005) understands gender as a category of occupational choice. This is a category that may restrict a woman's social space, perhaps making the choice of a STEM field unlikely even if her interests do in fact match it. The empirical evidence for Gottfredson (1981, 2005) theory is equivocal with studies often lacking appropriate methods when investigating the processes of circumscription and compromise, especially when trying to assess individuals' social space (Gottfredson, 2005; Junk and Armstrong, 2010).

With the objective of bringing more women into typically male occupations, the focus becomes which STEM occupations often are more "malleable," and have attributes that influence occupational choice.

Socio-Cognitive Career Theory

Here, the SCCT (Lent et al., 1994; Lent, 2013; Lent and Brown, 2013) provides a framework that is compatible with both person-environment fit approaches and developmental career theories. The SCCT is based on Bandura's (1986) social cognitive theory. It highlights three cognitive variables that are less stable than personality variables and that are related to individuals' interest development and career choices. These variables are self-efficacy beliefs, outcome expectations, and personal goals. Self-efficacy beliefs are personal assessments of one's own potential for being able to perform certain actions and cause related outcomes. Outcome expectations are evaluations of the results that possibly come with certain actions; they can be influenced by self-efficacy beliefs. Personal goals refer to the actions a person wants to carry out or to the outcome a person wants to produce (Lent, 2013). The SCCT includes four models that focus on different aspects of vocational behavior. According to the *choice model*, the development of vocational interests depends on self-efficacy expectations and outcome expectations, which are themselves the result of an individual's learning experiences. The learning experiences in turn depend on personal variables like gender or ethnicity and background influences like gender role socialization or the presence of different career role models. For example, if a person is encouraged to perform the realistic activities suggested by his or her parents and experiences positive outcomes, he or she may develop a high self-efficacy and positive outcome expectations toward realistic fields. He or she may develop strong realistic interests as a result. Vocational interests in this model have an impact on choice goals and actions (e.g., studying a STEM subject) with outcomes (e.g., exam results) that are the basis for new learning experiences. Along with the more distal background influences, there are environmental variables that can directly impact choice goals (and choice actions) or even moderate their relation to vocational interests (e.g., if a young woman is lacking emotional support for a career into STEM, the impact of her realistic orientation on her choice goals may become weaker). The choice model of the SCCT (Lent et al., 1994; Lent, 2013; Lent and Brown, 2013) could be confirmed by several empirical studies (Lent et al., 1994; Rottinghaus et al., 2003; Sheu et al., 2010). With regard to gender-specific differences, males show a higher level of confidence toward realistic and investigative issues (mechanical, outdoor/physical, mathematics, science) than females (Betz and Wolfe, 2005).

Consequences for Investing Females' Careers in STEM

Although the models described provide frameworks for career choice processes in general, more research is needed on the impacts of the factors on females' choices for or against STEM careers. This research should start from the macro perspective of Su and Rounds (2015), with "men preferring working with things and women preferring working with people" (Su et al., 2009, p. 880). This also includes a closer look at the interests of the individuals within the different STEM fields. Furthermore, much of the previous research was built on convenience samples comprised of students from a specific university or region. This makes it essential to systematically analyze samples of male and female students that either go into STEM or something else. It's also necessary to consider that not all STEM fields are created equal regarding vocational interests; this of course includes both the people-orientation as well as the things-orientation (Su and Rounds, 2015).

As it is, men are more interested in things (R) and women are more interested in people (S) (Su et al., 2009; Su and Rounds, 2015; Morris, 2016). Su and Rounds (2015) conclude – in line with Gottfredson (1981, 2005), Holland (1997), and the SCCT (Lent et al., 1994; Lent, 2013; Lent and Brown, 2013) – that interests are strong predictors of vocational choices. Women tend to choose social work environments and avoid STEM fields, especially those that only require realistic interests. Su and Rounds (2015) suspect two parallel processes lying behind the choice of social work environments and the avoidance of STEM fields by females. Both processes take into account quantitative skills mostly required to be successful in STEM. On the one hand, they assume a *constraining* process within which women may develop strong people-orientated interests and skills, but weak things-orientated interests and weak quantitative skills. Because these women have lower quantitative skills than others, they do not go into STEM fields or do not have the option to do so (c.f. Gottfredson, 1981, 2005). On the other hand, they assume a *broadening* process within which females who develop strong realistic interests and quantitative skills also have a better chance than males to develop social interests and people-orientated skills, reducing the chance to go into STEM fields (Wang et al., 2013; Woodcock et al., 2013).

MEASURING THE CONGRUENCE BETWEEN A PERSON'S INTERESTS AND THE VOCATIONAL ASPIRATION

The theory of Holland (1997) postulates that a person seeks the best fit between his or her interest profile and the profile of his or her (aspired) environment. Testing this hypothesis means creating profiles of both persons and environments and calculating the fit between both (this is also called *congruence*). There are several methods available to do this.

Representing Interest Profiles and Profiles of Work Environments

If inventoried interests (Super, 1957) in the sense of Holland (1997) are available, the complete interest profile usually consists of six scores, each indicating the similarity of a person to one of the six RIASEC types. Based on the six scores two different methods are commonly used to represent an interest profile. The first method uses the three dimensions the person is most similar to and creates a three-letter code (Holland, 1997). For example, if a person resembles the investigative (I) type the most followed by the realistic (R) type and the social (S) type, the person is assigned the code IRS (see **Figure 1A**). The second commonly used method relies on Holland's (1997) assumption that the RIASEC types can be mapped onto a regular hexagon, describing each interest value as a single vector directed toward the respective corner in the hexagon (see small arrows in **Figure 1B**). The interest profile can then be described by a vector resulting from the sum of the six single vectors (see thick arrow in **Figure 1B**). This vector has a length that describes the differentiation of the profile, and a main direction toward one of the RIASEC dimensions (Eder, 1998; c.f. Prediger, 1982).[1] **Figure 1** illustrates the differences between interest profiles represented by a three-letter code (**Figure 1A**) and by an interest vector (**Figure 1B**). Representing interest profiles by the three-letter code (**Figure 1A**)

[1]The method by Eder (1998) is similar to calculating the coordinates in the two-dimensional space using Prediger's (1982) data/ideas and things/people dimensions. The two methods reveal identical results.

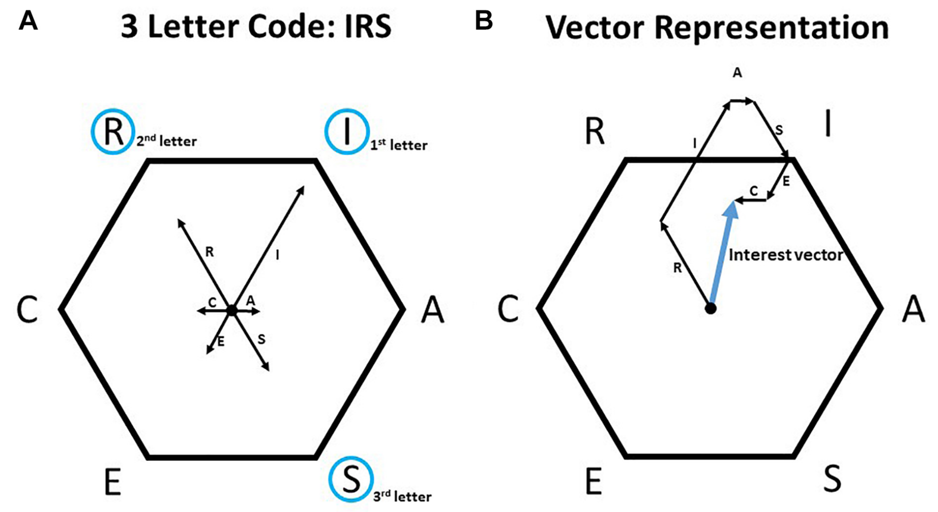


FIGURE 1 | Representation of the interest profiles using the three-letter code (**A**; Ertl and Hartmann, 2019, reprinted with permission) and using an interest vector **(B)**. The values in the single RIASEC dimensions are represented by the thin arrows, while the interest profile for the three-letter code is visualized by the blue circles and the numbers indicating the position of each letter **(A)**. This is represented by the thick blue arrow for the vector representation **(B)**.

allows a ranking of the three main dimensions (e.g., IRS). However, it does not address the magnitude of the differences between the dimensions, nor does it take into account the information given by the three remaining RIASEC dimensions. In contrast, the vector representation (**Figure 1B**) includes all the dimensions for representing an interest profile. Along with three-letter codes (e.g., IRS), an interest profile can be represented by the dominant type only (e.g., I) or by a two-letter code indicating the dominant and the second dominant type (e.g., IR). Depending on the data available, the methods of one-, two- and three-letter codes as well as the method of interest vectors can also be applied to work environments. The characterization of a work environment is mainly based on three different methods: expert ratings, assessing respondents' personality, and assessing the content and demands of work environments. In the case of expert ratings, occupational analysts characterize work environments and derive an environmental profile (e.g., three-letter codes or numerical profiles) using occupational information and Holland's (1997) RIASEC types (e.g., see Rounds et al., 2013). If the respondents' personality is assessed, Holland's (1997) environmental assessment technique (EAT) can be applied, i.e., the information about respondents having the same occupation is used to derive a one-, two-, or three-letter code that reflects the distribution of the personal dominant RIASEC types in that occupation. In these cases, the work environment is characterized by the personality of its inhabitants. A third option for creating an environmental profile is to ask inhabitants to fill out tests like the Position Classification Inventory (PCI; Gottfredson and Holland, 1991) that contains items assessing the content and demands of an occupation. The information from single questionnaires can be aggregated again to a one-, two-, or three-letter code. Given numerical profiles, an interest vector can also be calculated using the three methods described. An online source containing extensive information about work environments is available from O*NET OnLine (2018).

Analyzing Congruencies Between a Person and the Environment

A wide selection of different congruence indices are available for analyzing the congruence between a person's interest profile and the profile of his or her environment (e.g., first-letter agreement based on the hexagon by Holland, 1963; two-letter agreement index by Healy and Mourton, 1983; Z-S index by Zener and Schnuelle, 1976; M-Index by Iachan, 1984; ranked comparison congruence scale by Robbins et al., 1978). These algorithms based on Holland's codes differ for example in the consideration of the hexagonal model; in the number of letters considered; and in the weighting of differences and letter positions (c.f. Tracey and Sodano, 2013; Hartmann, 2018). Studies comparing different congruence indices reveal that their similarities range from $r = 0.05$ to $r = 0.98$ (Camp and Chartrand, 1992; Brown and Gore, 1994; Young et al., 1998) causing different results concerning the relation of congruence with outcome variables like occupational satisfaction (e.g., Assouline and Meir, 1987; Tranberg et al., 1993; Young et al., 1998; Tsabari et al., 2005). More recent studies apply the profile correlation (Allen and Robbins, 2010; Tracey et al., 2012; Wille et al., 2014; Xu and Tracey, 2016), the angle difference of two vectors in the hexagon (angular displacement; Tracey and Robbins, 2006; Hartmann et al., 2015), or the Euclidian distance of two vectors in the hexagon (Tracey and Robbins, 2005, 2006; Tracey et al., 2005, 2012, 2014; Neumann et al., 2009; Wille et al., 2014) to measure congruence between two profiles using full profile information. Studies using more than one of those methods reveal a rather moderate similarity between them (e.g., Tracey and Robbins, 2006; Tracey et al., 2012; Wille et al., 2014; Hartmann, 2018). All three methods are related in expected

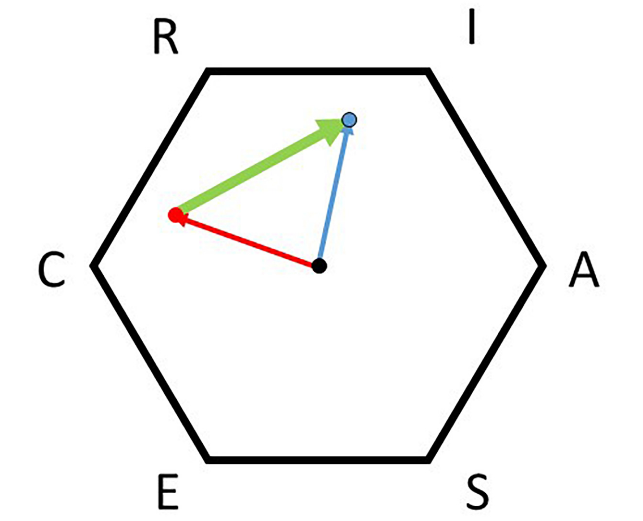


FIGURE 2 | Representation of the congruence between a person and his or her environment. Blue represents the person's interests, red the environment's, and green the measure of congruence.

ways to their outcomes (Tracey and Robbins, 2005; Tracey et al., 2005, 2012; Durr and Tracey, 2009; Neumann et al., 2009; Allen and Robbins, 2010). According to Tracey and Sodano (2013) the Euclidean distance is preferable because of its "ease of calculation, the unneeded assumption of independence of scales, and the easy extrapolation to more than two dimensions..." (p. 115). An application of the Euclidean distance is illustrated in **Figure 2**. The blue arrow represents the vector of the person's interests with the three main dimensions of IRS, while the red arrow represents the environment's profile with the main dimensions of CRS. The congruence between both profiles is estimated by the difference between both vectors (see the thick green arrow). With a higher congruence of person and environment, the difference and, consequently, the green arrow gets shorter until its value is zero, indicating maximum congruence. With a lower congruence, the green arrow grows until its length reaches twice the diameter of the hexagon for maximum divergence (see also Eder, 1998).

Context for Investigating Interest Congruence for STEM Professions

STEM fields are of great importance to the future development of society. However, especially the number of females who choose to take STEM subjects or STEM professions is limited. One possible explanation for the low proportion of females is that the interest profiles of women do not match the profiles of STEM fields.

The Holland (1997) model describes a well-received background for assessing the fit between the individual and her or his vocational environment. Yet, while the model is well received in research and career counseling, very few is known about applying this model for explaining STEM careers. A Scopus inquiry for "vocational interest STEM" in March 2019 just resulted in 36 hits of which just five deal more specific with this issue. Most of these five discuss the aspects of differences in the things and people orientation (e.g., Su et al., 2009; Yang and Barth, 2015) while one of them (Babarović et al., 2018) maps specific interests for STEM into the RIASEC hexagon. Thus, it is essential to analyze the interest profiles of females and males in STEM and also to compare these with other occupational fields to get reliable knowledge about the characteristic interest profiles of females in STEM.

This is even more true when looking at the aspect of interest congruence. Among other topics, previous research dealing with interest congruence focused on the measurement of interest congruence (e.g., Camp and Chartrand, 1992; Brown and Gore, 1994), its connection to outcome variables (e.g., Tsabari et al., 2005; Nye et al., 2012) or the congruencies between the individual's interests and the interests of her or his socialization group (e.g., Luttenberger et al., 2014; Etzel et al., 2018; Hartmann, 2018; Ertl and Hartmann, 2019). So far, research has barely dealt with the more specific topic of interest congruence within STEM fields: the Scopus research just revealed one relevant hit that investigates interest congruence as one variable among others for predicting IT job satisfaction (Carpenter et al., 2018). Thus, also in the area of interest congruence more research is needed that investigates how far the interest congruence of females and males varies within STEM and distinguishes from other fields.

METHODS, MODES OF INQUIRY, AND RESEARCH QUESTIONS

Research Questions

This leads to the following research questions and hypotheses:

1. *To what extent do female and male students of selected STEM/non-STEM fields distinguish themselves with respect to their RIASEC interest profiles?*
2. *To what extent do female and male students of selected STEM/non-STEM fields distinguish themselves with respect to the congruence between their interests and their vocational aspirations?*

Research Rationale and Hypotheses

Previous research has shown that it is crucial to differentiate within STEM fields when characterizing students' interest profiles. Therefore, using Holland's (1997) RIASEC model, the research questions analyze to what extent female and male students from selected STEM and non-STEM subjects with different proportions of women differ with respect to their interest profiles, their vocational aspirations and the congruence between their interests and aspirations. The following will apply a vector-based analysis while considering the aspects of conceptualizing interests, profiles, and congruence (e.g., Eder, 1998).

The first research question is comprised of a descriptive part that characterizes the interests and vocational aspirations of different student populations. Its results will provide insights into vectors as well as into the predominant three-letter codes representing the respondents' interests and vocational aspirations. In a second step it aims to analyze the extent of gender differences regarding interests and vocational aspirations. According to Su and Rounds (2015) one STEM field can be very heterogeneous in terms of its things-orientation and people-orientation and, in consequence, regarding the actual proportion of females. Here, we focus on the STEM field with a low proportion of females (STEM-L), comparing it with other fields of study. In addition, we look deeper into the gender differences within STEM-L by differentiating between subjects. In the context of the first research question we aim to test three hypotheses:

1. Students in STEM-L show higher realistic interests and lower social interests than students in other fields.
2. Students' vocational aspirations in STEM-L show a stronger realistic orientation and weaker social orientation than students' vocational aspirations in other fields.
3. Within STEM-L, female students show lower realistic and higher social interests than male students.

We generally assume a higher homogeneity within STEM-L, which means that interest differences decrease.

Research question 2 analyzes the match between the interests of the individuals and the interest profile of the vocational aspiration, which is the congruence between the individual and environment. Generally, we would hypothesize that people seek occupations that suit their interests (Gottfredson, 1981, 2005; Lent et al., 1994; Holland, 1997; Lent, 2013; Su and Rounds, 2015). Although there may be a higher chance for female students in STEM-L to have competing social interests that reduce their congruence, we assume that these females are able to seek a job that fits their interest profile. In the context of research question 2, we aim to investigate two alternative assumptions:

a. According to Su and Rounds (2015) we would assume that there are no differences regarding congruence between different fields of study.
b. Students have to overcome obstacles in subjects where they are under-represented. These may include stereotypes for females in STEM (see Ertl et al., 2017) or the low prestige of jobs for males in education and languages (see e.g., Gottfredson, 1981). According to SCCT (Lent et al., 1994) a higher congruence in interests could mitigate these obstacles to students choosing the respective field.

MATERIALS AND METHODS

Data sources used for the analysis were the cohort of first year students (SC5:10.0.0) of the German National Educational Panel Study (Blossfeld et al., 2011; see also acknowledgments) that started in the winter term of 2010/2011 (FDZ-LIfBi, 2018b). All students from this cohort gave informed consent to participate in the panel.

The dataset (SC5:10.0.0) contains 10 different waves of surveys (FDZ-LIfBi, 2018b) at different points in time. All analyses for this study come from wave 1 that was surveyed right after students' university entrance.

Sample and Sampling Procedures

The sampling applied students' *study subject* as filter having STEM subjects as main focus. These were represented by a three-digit classification of the German Federal Statistical Office (Destatis [Statistisches Bundesamt], 2018a). Analyzing the proportion of females in each study subject, the results showed that the NEPS dataset had an oversampling, with 60% of all cases being female. Looking at the respective German data, the German Federal Statistical Office reports in its statistics on German university entrants in the winter term 2010/2011 (Destatis [Statistisches Bundesamt], 2018b) that only 50% were females. Therefore, each student was assigned a variable with the proportion of females within the first study subject based on the data provided by the German Federal Statistical Office (Destatis [Statistisches Bundesamt], 2018b). Students were classified with respect to their first study subject according to the number of females within it (see e.g., Ertl et al., 2014; low-proportion: less than 30% females; high proportion: more than 70% females; moderate proportion: percentages in between). The subjects were furthermore clustered according to their area of study (e.g., STEM, medicine, economics, educational sciences, and languages). This paper focuses on female students in STEM and distinguishes them according to subjects with a low proportion of females (less than 30%; STEM-L) and a moderate proportion of females (between 30 and 70%; STEM-M). The sample also includes medicine with a moderate proportion of females because this field is often discussed within the context of STEM careers (European Commission, 2015; Su and Rounds, 2015; Ertl et al., 2017).

The control sample for comparison included economics with a moderate proportion of females, educational sciences with a high proportion of females, and languages with a high proportion of females (mainly German, English, and the Romance languages). These sub-samples were selected because they had comparably high numbers of students in the respective category and no admission restrictions. With 12,936 out of a total sample of 17,910 students, the data set analyzed comprises more than 70% of the total data set.

Variables and Analysis Procedures

The variables analyzed include:

- students' RIASEC values,
- and their vocational aspirations as ISCO-08 codes.

Regarding the RIASEC values, we applied the NEPS-generated IILS-II scales values for each RIASEC dimension (see FDZ-LIfBi, 2018a, pp. 699–704). The IILS-II (interest inventory lifespan; Wohlkinger et al., 2011) is comprised of three items per dimension (two of them stemming from the AIST, Bergmann and Eder, 2005). Each had a range from one to five. The internal consistency of these scales was best for the social dimension and worst for the enterprising dimension (Cronbach's a for Realistic: $\alpha = 0.704$; Investigative: $\alpha = 0.625$; Artistic: $\alpha = 0.629$; Social: $\alpha = 0.749$; Enterprising: $\alpha = 0.523$; Conventional: $\alpha = 0.561$). Please find the means and standard deviations for each dimension in **Supplementary Table 1**. Testing the hexagonal structure of vocational interests, we used Tracey's (1997) program RANDALL to conduct the randomization test of hypothesized order relations (Hubert and Arabie, 1987). For our sample, the correspondence index (CI) was 0.81, p = 0.017, indicating a good fit of the data with the postulated hexagon. Students' *three-letter codes* were generated from these values.

Students' *vocational aspirations* were provided as ISCO-08 codes. The ISCO-08 codes were matched with the occupational information provided by O*NET (2018) to obtain information about the RIASEC classification of these aspirations. This procedure had three steps: First, as there are more O*NET-SOC codes than ISCO codes, RIASEC classifications of the O*NET interest table were aggregated regarding the first six digits of the O*NET-SOC code. This step focused on the O*NET main categories that can be matched with ISCO codes. Second, the O*NET-SOC codes were translated into ISCO codes based on the ISCO-08 to the 2010 SOC crosswalk table (Bureau of Labor Statistics, 2019). RIASEC values were aggregated in cases when the crosswalk provided more O*NET codes as equivalent for one ISCO code. Finally, the resulting RIASEC classifications codes were assigned to the NEPS ISCO classifications of students' vocational aspirations. In total, 9,860 out of 10,196 vocational aspirations (96.7%) were able to be RIASEC-classified via this procedure. Three hundred and eighteen vocational aspirations were NEPS coded as ISCO code 2100, which is a container category usually not foreseen as code in the ISCO-08 classification; the coding reflects a relatively vague aspiration in the context of doing something with technology. As this category is quite broad (mathematics and sciences), no RIASEC code could be assigned to those students' aspirations, although they are in fact possibly part of the STEM fields. A total of 18 students (0.2%) without a RIASEC code for their vocational aspiration remained. Please find the means and standard deviations for the RIASEC dimensions of the aspirations in **Supplementary Table 4**.

All analyses were done with SPSS 25.0.

RESULTS

Research Question 1: Students' Interest Profiles

Students' Interest Profiles

Table 1 gives an overview of the students' interest profiles. The analysis shows that most females and males in all, but the STEM subjects had vector directions toward the same RIASEC dimensions in their interests. With STEM-M, 31.1% of the females had a predominant social dimension, while 19.8% of the males had a predominantly enterprising dimension. Differences also occurred for the STEM-L subjects in which 21.2% of the females had a predominant investigative dimension, while 34.5% of the males had a realistic dimension. With over 60%, educational sciences and females in the languages showed the most distinct vector directions, while overall the STEM subjects showed the lowest. This shows that students' interests are more diverse for the STEM subjects than for other subject groups.

We can recognize large variability when looking at the most frequent three-letter codes (**Table 1**). In almost all fields (except medicine) the occurrences of the most frequent code were below 10%, and even in medicine the most frequent code occurred just around 15% of the time. Particularly in the STEM fields we could see a high variability, with even the first most frequent codes contributing often less than 5%[2]; these codes were also comprised of opposite dimensions, e.g., R and S or I and E.

Looking at the things and people orientation, we can see significant differences in the realistic interests in all but the educational sciences (see **Figure 3**). Students' realistic interests were noticeably higher in STEM-L than in all other fields. For females, the R values in STEM-M were higher than in the remaining other fields. For medicine, the means were between STEM-M and the remaining fields, although the confidence intervals were partially overlapping. Considering 3 as the scale mean, we can furthermore observe that students' realistic interests in all but the STEM fields were below this.

This is different for students' social interests (see **Figure 4**). Here, we can see that all but the interests of the males in STEM-L were above the middle of the scale of 3. For the social interests, females showed significantly higher values for all but the educational sciences. Students showed the highest social interests in medicine and the educational sciences, and the lowest in STEM-L and economics, which showed significant but only marginally higher values than STEM-L. Social interests in STEM-M were significantly higher than in economics, and in the languages they were even higher than in STEM-M.

In light of the hypothesis that females have higher social interests (go more into people-related careers) and males have more realistic interests (go more into thing-related careers), we can verify this hypothesis for all but the educational sciences, which showed significant differences for neither of the dimensions. The results furthermore showed that STEM-L is the most things-oriented, and medicine and educational sciences are the most people-oriented. Economics is neither, and its values are among the lowest for both categories.

Interests Profiles of Students' Vocational Aspirations

When it comes to students' vocational aspirations, **Table 2** provides insights into the three-letter codes and the main direction of the interest vectors. Considering the vector directions of students' vocational aspiration, females and males had similar, and most frequent predominant letters in all areas but economics. Students' aspirations in STEM-L had a focus on realistic activities. Students in STEM-M, educational sciences, and the languages had a focus on social activities. In medicine their focus was on investigative activities, while in economics the focus was on enterprising activities for the female students and on conventional ones for the males. These kinds results are also reflected in the distributions of the three-letter codes of the different aspirations.

Focusing on the homogeneity of students' vocational aspirations (**Table 2**), we can see extensive differences between

[2] An equal distribution would show 0.83% for each code.

TABLE 1 | Most- and second-frequent three-letter codes and vector directions derived from students' interest values for the different subject areas including the proportion and absolute numbers (N) of students that showed the respective letter/codes.

	Female						Male					
	Three-letter code			Vector direction			Three-letter code			Vector direction		
	Code	Prop.	*N*	Dir.	Prop.	*N*	Code	Prop.	*N*	Dir.	Prop.	*N*
STEM-L	IRS	1.3	10	I	21.2	778	RIE	3.6	109	R	34.5	3041
	RIS	1.0	8				IRE	2.1	64			
STEM-M	ISE	2.7	49	S	31.1	1812	ISE	2.1	20	E	19.8	947
	SIE	2.3	42				IES	1.4	13			
MED-M	SIE	15.2	72	A	40.0	473	SEI	14.9	31	A	37.5	208
	SIA	4.4	21				SIE	9.6	20			
ECO-M	ECS	4.3	52	E	46.2	1196	ECS	4.4	34	E	53.9	775
	ESC	3.8	46				ESC	4.4	34			
EDU-H	SEA	8.6	89	S	68.1	1034	SEA	8.5	12	S	60.3	141
	SAE	7.0	73				SAE	6.3	9			
Lang-H	SEA	6.0	126	S	61.2	2100	SEA	4.2	17	S	49.0	406
	ASE	4.5	94				SEC	3.0	12			

STEM-L, STEM subjects with a low proportion of females; STEM-M, STEM subjects with a medium proportion of women; MED-M, medicine (with a moderate proportion of females); ECO-M, economics (with a moderate proportion of females); EDU-H, education (with a high proportion of females); Lang-H, languages (with a high proportion of females). Code, frequent three-letter codes composed of the RIASEC dimensions: R, realistic; I, investigative; A, artistic; S, social; E, enterprising; C, conventional; Prop., proportion of three-letter codes or vector directions. Dir., frequent vector directions according to the RIASEC dimensions.

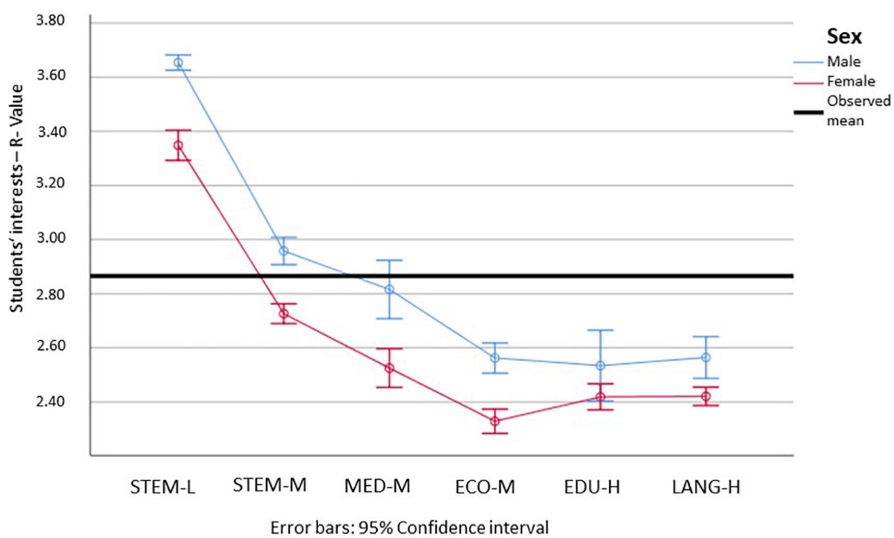


FIGURE 3 | Means and 95% confidence intervals for the students' realistic interests, for male (blue) and female (red) students in the different subject groups (min = 1; max = 5). STEM-L, STEM subjects with a low proportion of females; STEM-M, STEM subjects with a medium proportion of women; MED-M, medicine (with a moderate proportion of females); ECO-M, economics (with a moderate proportion of females); EDU-H, education (with a high proportion of females); Lang-H, languages (with a high proportion of females). Please find the means, standard errors, and confidence intervals for each group in **Supplementary Table 2**.

the frequencies of students' three-letter codes in STEM-L and STEM-M. While most students in STEM-M head into social professions with a high proportion of teaching activities, students in STEM-L show much more diverse aspirations.

Looking now at our hypothesis that males go more into things and females more into people, we can analyze the interest profiles of students' vocational aspirations. In terms of the realistic or things dimension (see **Figure 5**), we can only observe one significant difference between females and males in STEM-L, which is noticeably small when comparing it with the differences between the STEM-L group and the other groups in the study. Of particular note is also that the medical aspirations have a very high amount of realistic interests, which is in contrast to the values regarding students' individual realistic and social interests. When looking at the social interest profile of students' aspirations (see **Figure 6**), we see several significant but small differences between males and females in STEM and economics. This indicates that females tend toward more social aspirations

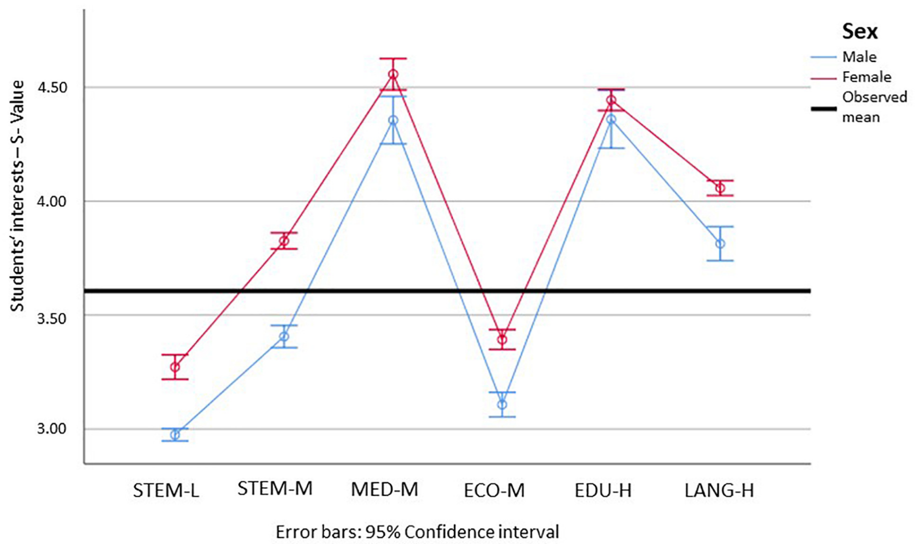


FIGURE 4 | Means and 95% confidence intervals for the students' social interests, for male (blue) and female (red) students in the different subject groups (min = 1; max = 5). STEM-L, STEM subjects with a low proportion of females; STEM-M, STEM subjects with a medium proportion of women; MED-M, medicine (with a moderate proportion of females); ECO-M, economics (with a moderate proportion of females); EDU-H, education (with a high proportion of females); Lang-H, languages (with a high proportion of females). Please find the means, standard errors, and confidence intervals for each group in **Supplementary Table 3**.

TABLE 2 | Most- and second-frequent three-letter codes and vector direction derived from students' vocational aspirations for the different subject areas including the proportion and absolute numbers (*N*) of students that showed the respective letter/codes.

	Female						Male					
	Three-letter code			Vector direction			Three-letter code			Vector direction		
	Code	Prop.	*N*	Dir.	Prop.	*N*	Code	Prop.	*N*	Dir.	Prop.	*N*
STEM-L	RIC	9.4	73	R	52.8	472	IRA	10.5	320	R	65.5	1926
	IRA	6.7	52				IRC	9.4	288			
STEM-M	SAE	49.3	894	S	72.3	1533	SAE	42.3	401	S	60.1	735
	SAI/C	8.9	161				IRC	10.1	96			
MED-M	ISR	87.8	416	I	97.8	451	ISR	90.9	189	I	99.0	200
ECO-M	ECS	15.4	185	E	34.8	728	ECS	14.0	109	C	45.1	519
	SAE	6.3	76				CEI	11.0	86			
EDU-H	SIA	29.2	303	S	75.7	845	SIA	24.6	35	S	76.3	118
	SAE	13.3	138				SAE	15.5	22			
Lang-H	SAE	64.0	1346	S	90.5	1954	SAE	74.1	301	S	91.3	379
	SAI/C	13.2	277				AEC	3.9	16			

STEM-L, STEM subjects with a low proportion of females; STEM-M, STEM subjects with a medium proportion of women; MED-M, medicine (with a moderate proportion of females); ECO-M, economics (with a moderate proportion of females); EDU-H, education (with a high proportion of females); Lang-H, languages (with a high proportion of females). Code, frequent three-letter codes composed of the RIASEC dimensions: R, realistic; I, investigative; A, artistic; S, social; E, enterprising; C, conventional; Prop., proportion of three-letter codes or vector directions. Dir., frequent vector directions according to the RIASEC dimensions.

within each field of study. Comparing both dimensions, we can see that all, but the STEM-L and medical students' aspirations have realistic values below the middle of the scale of 4. With the social dimension, almost all aspirations besides STEM-L and economics show high to very high characteristics of this dimension. This allowed us to partially verify the hypothesis that the fields of study are linked to differences in the things dimension and people dimension. However, for this particular student population, we obtained stronger evidence for the things dimension.

Interest Profiles of Students in STEM-L

Focusing now specifically on STEM-L, the following reports the data from subjects with more than 100 students in the sample. Regarding the realistic dimension, we can see that females score notably lower in all subjects except general and electrical engineering, and that the confidence intervals are generally larger for females than for males (which can be attributed to the reduced sample size). We can furthermore observe that the significant difference between males and females seen in **Figure 3** disappears

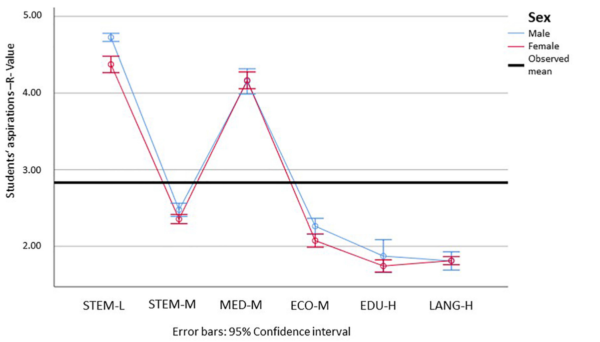


FIGURE 5 | Means and 95% confidence intervals for interest profiles of students' aspirations, realistic dimension, for male (blue) and female (red) students in the different subject groups (min = 1; max = 7). STEM-L, STEM subjects with a low proportion of females; STEM-M, STEM subjects with a medium proportion of women; MED-M, medicine (with a moderate proportion of females); ECO-M, economics (with a moderate proportion of females); EDU-H, education (with a high proportion of females); Lang-H, languages (with a high proportion of females). Please find the means, standard errors, and confidence intervals for each group in **Supplementary Table 5**.

on the subject level for all subjects except mechanical engineering and traffic engineering (see **Figure 7**).

Regarding the social dimension, we can generally observe higher values of females in this dimension, as well as generally higher confidence intervals of females (see **Figure 8**). Significant differences could be found for about half of the subjects including computer science, physics, and mechanical, electrical, and civil engineering.

Looking now at our hypothesis, we can still see significant differences persisting, although only for a few subjects. Of note here are the significant differences in the social dimension and partially in the realistic dimension.

Research Question 2: Congruence Between Students' Interests and Their Vocational Aspirations

Keeping in mind the significant gender differences regarding students' individual interests and the comparably small differences in the field of occupational aspirations, it's interesting to delve deeper into the congruencies of students' interests and their aspirations. With the congruence vectors, it is important to note that a lower value means a higher congruence. The vector values are around 0.8 (see **Table 3**). Considering the maximum congruence of 0 and a theoretical minimum of 4, these vectors are within the highest quartiles for females as well for males, which indicates that students chose a subject that is in line with their individual interests. Descriptively, STEM subjects show lower congruencies, while educational sciences and languages show higher ones. The highest congruence is found with males in medicine, and the lowest with males in STEM-M.

Looking now at significant differences (**Figure 9**), gender differences are seen for STEM-M and the languages, with females showing a higher congruence in both areas. Regarding the differences between the subjects, females show the highest congruencies in educational sciences and the languages. These are slightly but significantly lower in medicine, lower again in STEM-M and economics, and the lowest in STEM-L. Males in contrast show the highest congruencies in medicine and educational sciences, lower congruencies in the languages, lower ones in economics, and the lowest in STEM.

Looking more specifically at the STEM-L subjects, we were unable to identify significant differences for any of the

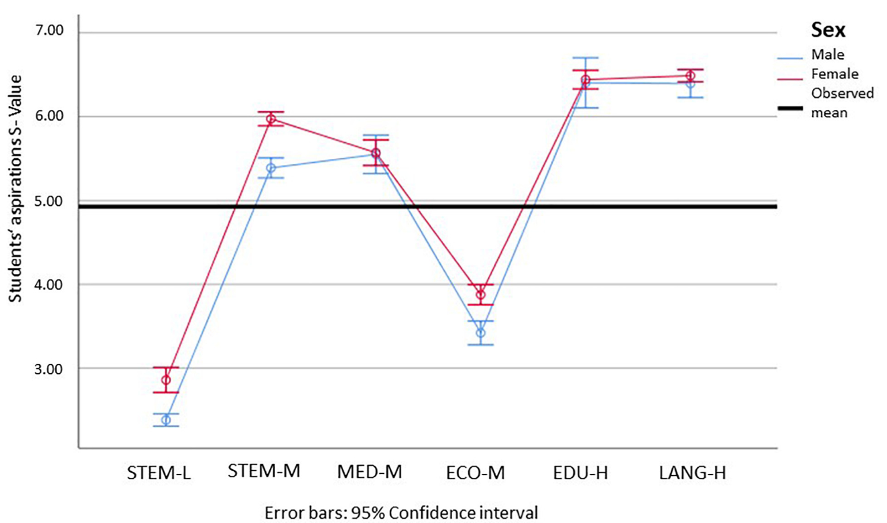


FIGURE 6 | Means and 95% confidence intervals for interest profiles of students' aspirations, social dimension, for male (blue) and female (red) students in the different subject groups (min = 1; max = 7). STEM-L, STEM subjects with a low proportion of females; STEM-M, STEM subjects with a medium proportion of women; MED-M, medicine (with a moderate proportion of females); ECO-M, economics (with a moderate proportion of females); EDU-H, education (with a high proportion of females); Lang-H, languages (with a high proportion of females). Please find the means, standard errors, and confidence intervals for each group in **Supplementary Table 6**.

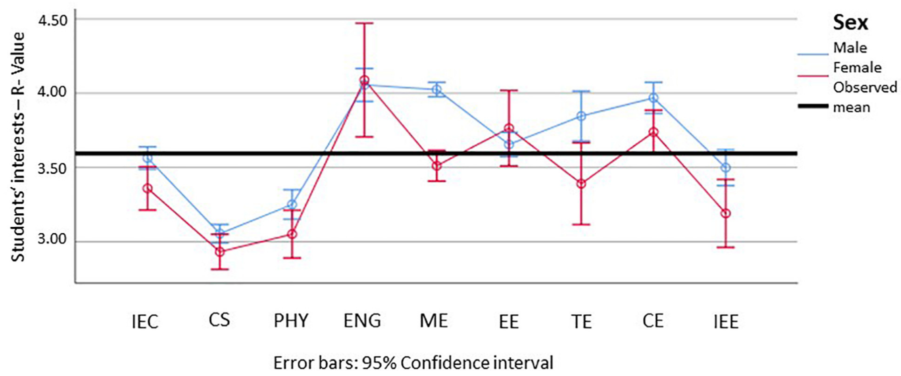


FIGURE 7 | Means and 95% confidence intervals for the students' realistic interests, for male (blue) and female (red) students in the different subject groups (min = 1; max = 5). IEE, Industrial Engineering, focus on economics; CS, Computer Science; PHY, Physics, Astronomy; ENG, Engineering, general; ME, Mechanical Engineering; EE, Electrical Engineering; TE, Traffic Engineering; CE, Civil Engineering; IEE, Industrial Engineering, focus on engineering. Please find the means, standard errors, and confidence intervals for each group in **Supplementary Table 7**.

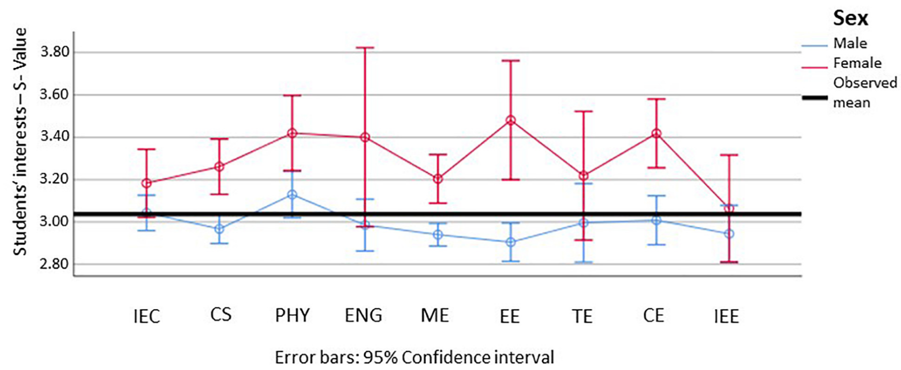


FIGURE 8 | Means and 95% confidence intervals for the students' social interests, for male (blue) and female (red) students in the different subject groups (min = 1; max = 5). IEE, Industrial Engineering, focus on economics; CS, Computer Science; PHY, Physics, Astronomy; ENG, Engineering, general; ME, Mechanical Engineering; EE, Electrical Engineering; TE, Traffic Engineering; CE, Civil Engineering; IEE, Industrial Engineering, focus on engineering. Please find the means, standard errors, and confidence intervals for each group in **Supplementary Table 8**.

TABLE 3 | Mean length of students' congruence vectors.

	Female			Male		
	M	***SD***	***N***	***M***	***SD***	***N***
STEM-L	0.963	0.341	470	0.941	0.320	1923
STEM-M	0.872	0.295	1533	0.973	0.304	735
MED-M	0.715	0.240	450	0.648	0.250	200
ECO-M	0.846	0.327	727	0.842	0.329	517
EDU-H	0.657	0.331	842	0.678	0.302	118
Lang-H	0.666	0.275	1952	0.777	0.296	379

STEM-L, STEM subjects with a low proportion of females; STEM-M, STEM subjects with a medium proportion of women; MED-M, medicine (with a moderate proportion of females); ECO-M, economics (with a moderate proportion of females); EDU-H, education (with a high proportion of females); Lang-H, languages (with a high proportion of females).

subjects (see **Figure 10**). Descriptively, we can observe the high confidence intervals of females and a noticeably lower congruence, although this was not significant in electrical engineering and traffic engineering.

Looking at our alternative assumptions, we see that both are unable to satisfactorily explain the results. The assumption that students in fields in which they are under-represented also show a higher congruence regarding their interests and vocational aspirations could only be confirmed for male students in the educational sciences and the languages, and only when compared to males in STEM. Moreover, when interpreting the confidence intervals, it could be assumed that a higher sample size of females may in fact disclose that they show a significant lower congruence in STEM-L than males.

SUMMARY AND DISCUSSION

Summarizing our results, we can see on an *individual interest* level that females show significantly higher social interests, and males higher realistic ones across all study subjects except educational science. This expands the research of Su et al. (2009) by showing that the phenomenon of females having more social interests and males having more realistic ones holds across the

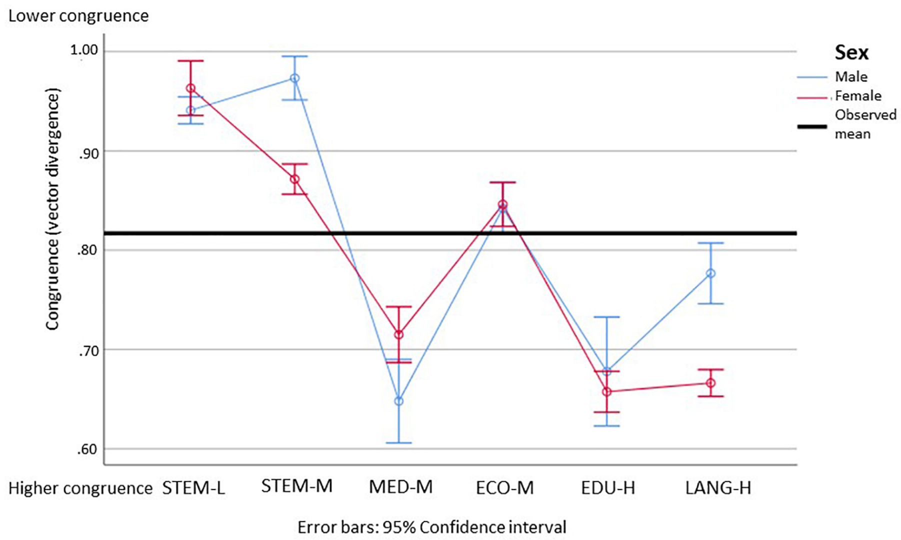


FIGURE 9 | Means and 95% confidence intervals of students' congruence between their interests and vocational aspirations for female (red) and male (blue) students based on an analysis with interest vectors (lower means indicate higher congruence). STEM-L, STEM subjects with a low proportion of females; STEM-M, STEM subjects with a medium proportion of women; MED-M, medicine (with a moderate proportion of females); ECO-M, economics (with a moderate proportion of females); EDU-H, education (with a high proportion of females); Lang-H, languages (with a high proportion of females). Please find the means, standard errors, and confidence intervals for each group in **Supplementary Table 9**.

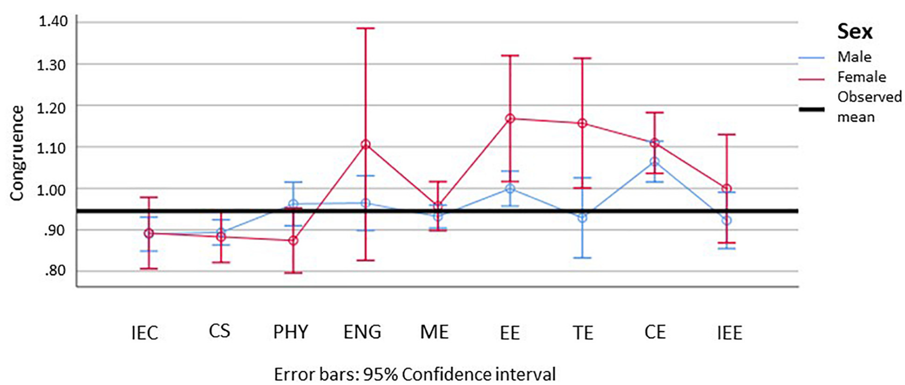


FIGURE 10 | Means and 95% confidence intervals of students' congruence between their interests and vocational aspirations for female (red) and male (blue) students based on an analysis with interest vectors (lower means indicate higher congruence). IEE, Industrial Engineering, focus on economics; CS, Computer Science; PHY, Physics, Astronomy; ENG, Engineering, general; ME, Mechanical Engineering; EE, Electrical Engineering; TE, Traffic Engineering; CE, Civil Engineering; IEE, Industrial Engineering, focus on engineering. Please find the means, standard errors, and confidence intervals for each group in **Supplementary Table 10**.

disciplines. Of note, students in STEM-L showed a much higher level of realistic interests than all other fields, and a lower level of social interests than most other fields.

We obtained a quite different observation regarding students' occupational aspirations. Here, we only found one significant difference between females and males in the realistic dimension in the area of STEM-L. This was also the case for the social dimension where we obtained less significant differences for the subjects of STEM and economics only. These differences were in line with the approach of Su and Rounds (2015) for predicting the proportion of females within a subject area via the difference in the things-people dimension. Interestingly, the significant differences held, consistent with Su and Rounds (2015), for the STEM-L area, which indicates that there are further factors affecting the proportion of females in this area. Overall, we observed the highest levels of realistic for occupational aspirations in STEM-L and medicine, and the lowest levels of social for occupational aspirations in STEM-L and economics.

Looking now at the congruencies between the individuals' interests and the respective profile of their occupational aspiration, we could only see significant differences for STEM-M and the languages. In both, females showed a higher congruence than males. Remarkably, students in STEM and especially males in STEM-M showed the lowest levels of congruence, while students in education, males in medicine, and females in the languages showed the highest. Recalling our hypothesis (b) according to SCCT (Lent et al., 1994; Lent, 2013; Lent and Brown, 2013) that students show higher levels of congruence in fields where they are under-represented, we were unable to verify this, especially when it comes to females in STEM-L.

Discussion

The paper started with the societal challenge to increase the number of STEM graduates (see Hingel et al., 2008). Looking, however, into students' interest profiles and especially into the congruence between their interests and their vocational aspirations, we can see that this congruence is noticeably lower for STEM and especially for STEM-L than for any other of the fields analyzed – for females as well as for males. Keeping in mind that the analyses included 70% of a stratified student sample of German first year students, we have to assert that students' interest profiles fit less for STEM environments than for other occupational environments.

The High Focus on Things as a Challenge for STEM Career Choices

One reason for this discrepancy may result from the strong focus on realistic interests, the *things* dimension, that STEM, especially STEM-L occupations require. We realized that students in STEM fields with a low proportion of females (STEM-L) showed noticeably higher realistic interests and lower social interests than male and female students in other subjects, including STEM fields with a moderate proportion of females (STEM-M). In other words, STEM fields in which students have high realistic interests and low social interests have a low proportion of females. Considering the strong gender differences regarding the things-orientation and people-orientation (Su et al., 2009; Morris, 2016), and in line with previous studies (Woodcock et al., 2013; Su and Rounds, 2015), we can state that men and women follow their gender-specific interests when choosing a study subject. In the current study, this was especially true for realistic interests because the subjects' proportion of females was connected to the magnitude of the corresponding subject-specific mean of realistic interests.

This was less evident with respect to social interests, the *people* dimension, which were around or above the middle of the scale for all fields investigated. Besides STEM, also students of economics show a medium average in terms of social interests, which does not fit the moderate proportion of females. This would appear to be a combination of low realistic and low social interests that lead men and women to the decision to study economics. In addition, they on average show high enterprising and high conventional interests. Please find the means, standard errors, and confidence intervals for the enterprising and conventional dimension for each group in **Supplementary Tables 11** and **12**. Students in medicine with a moderate proportion of females on the other hand show social interests that are as high as or even higher than the mean social interests in study subjects with a high proportion of females (education, languages). This deviation is in line with the study of Su and Rounds (2015, p. 15) who expected a higher percentage of women in medical science given the observed gender differences in interests related to this field.

Similar differences between the study subjects can be observed regarding the things-orientation of students' vocational aspirations. Here, students in STEM-L choose occupations that have a stronger realistic orientation than the occupations chosen by students in other fields of study. In addition, students in STEM-M aspire toward occupations with a higher realistic orientation than students in study subjects with a high proportion of women (education, languages). The only clear exceptions here are medical students aspiring toward medical occupations containing a relatively strong realistic orientation. This discrepancy is also in line with the study by Su and Rounds (2015, p. 15) who, based on the gender differences in interests, expected less than the actual proportion of women in the field of medical services. Su and Rounds (2015) conclude that there must be factors in addition to a high people-orientation (and a low things-orientation) within work environments that attract women (e.g., working conditions; e.g., Babarović et al., 2018). The magnitude of the mean social orientation of aspired occupations predominantly reflects the proportion of women in the different study subjects. Students in STEM-L aspire toward occupations with a low people-orientation, while students of STEM-M and medicine aspire toward occupations with a moderate people-orientation that is lower than the mean people-orientation in education and the languages. Again, the field of economics is characterized by a low people-orientation and a low things-orientation.

Effects of and Reasons for a Low Congruence Between Personal Interests and the Occupational Interest Profile

Compared to students in other subjects, students in STEM show smaller interest congruence with respect to their vocational aspirations. Apparently, students in STEM choose aspirations that are less compatible with their interest profiles. According to Holland (1997) and empirical evidence (e.g., Tsabari et al., 2005; Nye et al., 2012) lower congruence is connected to lower vocational satisfaction and performance. Thus, if students in STEM take the occupations they aspire it could be expected that they are less satisfied and perform lower than students of other fields.

In accordance to the RIASEC model people seek occupations that fit their interests and vocational environments seek people that fit their requirements. Since this tendency seems to be relatively weak in STEM, the question arises as to why a STEM subject or profession is chosen, even though one's own interests do not fit (perfectly) with the requirements (or why a vocational environment chooses such an individual). As Gottfredson's (1981, 2005) theory and the SCCT (Lent et al., 1994; Lent, 2013; Lent and Brown, 2013) point out there are multiple reasons for aspiring an occupation. Apart from fitting interests the sex-type or prestige of an occupation or outcome expectations like favorable prospects in the job market could be alternative reasons for an occupational choice. Vice versa, if there is a lack of candidates a vocational environment may also choose aspirants showing only moderate or weak fit.

Following the suggestions of Su and Rounds (2015), we aimed to investigate STEM fields on a finer level by distinguishing between different subjects within STEM-L. We assumed that within the different STEM-L subjects, the gender differences regarding realistic and social interests should vanish or at least decrease. This assumption could be confirmed for some but not all of the subjects, showing again that not all STEM fields are equal, even on a finer level.

Limitations

The analyses rely on the huge sample of a large-scale panel study that applied professional sampling procedures. Here, the results are different from convenience samples applied in other studies of occupational interests. This strength, however, does in fact have some limitations, e.g., that only short scales of occupational interests could be processed. Although more detailed instruments would be desirable, the vector-analytical approach is very robust when it comes to analyzing congruencies in the context of these short scales. This robustness comes from the construction of the interest vectors as well as from the calculation of the Euclidean distance as a congruence measure. Both methods are far less sensitive to the side effects of short scales (e.g., equal values of dimensions) than ranking-based algorithms.

A further and more general problem when studying the characteristics of STEM fields on a fine level and in the context of other fields of study is the small number of women in work environments with a high things-orientation. The large confidence intervals in **Figures 8** and **10** reflect this challenge. Considering that the NEPS SC5 has around 18,000 participants, it would be desirable to conduct an effective oversampling of under-represented student populations.

For the Future

We could generally observe that students' values for the realistic interests were comparably low, even for the STEM-L field, especially when compared to the social dimension. This may result from the construction of interest inventories for a wide population. Looking at the initial samples of interest inventories (e.g., Bergmann and Eder, 2005; or the meta-study of Su et al., 2009), it's obvious that these were either mainly administered with school children or with a broad range of professions that may also include students or professions requiring a university degree. These inventories may, especially for the realistic dimension, comprise a high amount of skilled manual work activities that often don't perfectly match occupations that students aspire to after finishing their university degrees. They hardly can account for different working profiles within occupations (this is by the way the pitfall of all kinds of interest inventories) even if O*NET (2018) provides a very fine-grained structure of occupations. In any case, this will require further development in the interest inventories, e.g., by distinguishing the R dimension with respect to physical or manual technical work and more white-collar professions. Su et al. (2018) followed this approach by developing an eight-dimensional interest model.

CONCLUSION

One of the most astonishing outcomes of this study is its low congruence of all students in STEM-L and the even lower congruence of males in STEM-M. This indicates a worse fit between individual interests and the vocational aspirations of students in these areas compared to students in other areas. At first glance, this finding appears similar to the results of Su and Rounds (2015) who had discrepancies in predicting the proportion of females in the areas with low proportions of females. However, it, has to be acknowledged that the approach of this study was different from Su and Rounds (2015). While they analyzed the things-people discrepancy on a macro level and compared this with the proportion of females within a subject area, our study compared the individual interests of students within different fields of study with the interest profile of the individual's own occupational aspiration. Whatever the outcome, the results look similar: in the area of STEM-L either the proportion of females only vaguely meets the prediction or, in our case, in the area of STEM-L there are still remaining significant differences in both the realistic and social dimensions. This indicates that, especially for the area of STEM-L, other variables mediate the impact of interests. These may be, e.g., prestige (Gottfredson, 1981, 2005), stereotypes (Konrad et al., 2000; Ertl et al., 2014, 2017), aptitudes and motivational beliefs (Eccles and Wang, 2016), contextual variables like career-related network contacts (Lent, 2013; Lent and Brown, 2013), socialization factors (e.g., Bleeker and Jacobs, 2004), or working conditions (Ferriman et al., 2009; Moss-Racusin et al., 2012). These should be analyzed in a further study. For STEM-M on the other hand, we could observe that females show a noticeably higher congruence than males, although they are in fact the second-lowest for females in the subjects investigated. We also could observe a higher proportion of females aspiring to teaching or counseling professions in this field than males. Here, further research should investigate how far this difference results from females aspiring toward more social job profiles within an occupational area (see e.g., Gottfredson, 1981, 2005; Su and Rounds, 2015).

Implications

The analyses in this paper show how necessary it is to distinguish different STEM subjects – at least with respect to the corresponding proportion of females. Just using the term "STEM" may blur obvious differences of subjects like biological science and engineering (see e.g., Ertl et al., 2014; Su and Rounds, 2015). Determining the proportion of females in a subject is one approach to distinguish this. It might also be worthwhile to develop a more fine-grained but researchable clustering of STEM in an effort to differentiate the investigated effects more effectively.

The theories above offer useful information in terms of bringing more women into STEM. According to Gottfredson (1981, 2005), possible steps to promote the realistic interests of girls/women should be taken at an early developmental stage before the things-orientated work environments are excluded due to a lack of consistency with the own self-concept. Consistent with the SCCT (Lent et al., 1994; Lent, 2013; Lent and Brown, 2013), interventions may also work at a later developmental stage provided that they are able to cause shifts in interests by influencing learning experiences, self-efficacy beliefs, and outcome expectations. We agree with Su and Rounds (2015) when they propose an emphasis on the social aspects of STEM fields. This approach may be more promising than the attempt to promote the development of a differentiated realistic interest profile and could generally accentuate the importance of social values when it comes to societal development.

ETHICS STATEMENT

The analyses of this manuscript are secondary analyses of data published previously (Blossfeld et al., 2011). Data sources used for the analyses were the cohort of first year students (doi: 10.5157/NEPS:SC5:10.0.0) of the German National Educational Panel Study (Blossfeld et al., 2011). All students from this cohort gave informed consent to participate in the panel by providing their phone number for being contacted for telephone interviews after being informed about the purposes of the study. Specific information about the recruitment process can be found in the field report of the study (Steinwede and Aust, 2012). All data analyses were performed via a remote terminal (RemoteNEPS) at the LIfBi in Bamberg, Germany that provided a controlled privacy environment for data access. Furthermore, an ethics approval for the analyses was obtained by the local ethics committee.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This manuscript uses data from the National Educational Panel Study (NEPS): Starting Cohort First-Year Students, *doi: 10.5157/NEPS:SC5:10.0.0*. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

This manuscript includes information from the O*NET 22.3 Database by the United States Department of Labor, Employment and Training Administration (USDOL/ETA). Used under the CC BY 4.0 license, O*NET® is a trademark of USDOL/ETA. The authors have modified all or some of this information. The USDOL/ETA has not approved, endorsed, or tested these modifications.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.00897/full#supplementary-material

REFERENCES

Allen, J., and Robbins, S. (2010). Effects of interest-major congruence, motivation, and academic performance on timely degree attainment. *J. Couns. Psychol.* 57, 23–35. doi: 10.1037/a0017267

Assouline, M., and Meir, E. I. (1987). Meta-analysis of the relationship between congruence and well-being measures. *J. Vocat. Behav.* 31, 319–332. doi: 10.1037/bul0000120

Babarović, T., Dević, I., and Burušić, J. (2018). Fitting the STEM interests of middle school children into the RIASEC structural space. *Int. J. Educ. Vocat. Guid.* 19, 111–128. doi: 10.1007/s10775-018-9371-8

Bandura, A. (1986). *Social Foundations of Thought and Action. A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bergmann, C., and Eder, F. (2005). *AIST-R: Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R)*, 3rd Edn. Göttingen: Beltz test Gesellschaft.

Betz, N. E., and Wolfe, J. B. (2005). Measuring confidence for basic domains of vocational activity in high school students. *J. Career Assess.* 13, 251–270.

Bleeker, M. M., and Jacobs, J. E. (2004). Achievement in Math and Science: Do Mothers' Beliefs Matter 12 Years Later? *J. Educ. Psychol.* 96, 97–109. doi: 10.1037/0022-0663.96.1.97

Blossfeld, H.-P., Roßbach, H.-G., and von Maurice, J. (2011). *Education as a lifelong process: the German National Educational Panel Study (NEPS).* Zeitschrift für Erziehungswissenschaft - Sonderheft (Berlin: Springer)

Brown, S. D., and Gore, P. A. J. (1994). An evaluation of interest congruence indices: distribution characteristics and measurement properties. *J. Vocat. Behav.* 45, 310–327.

Bureau of Labor Statistics (2019). *Crosswalks between the 2010 SOC and Systems Used by other Federal and International Statistical Agencies*. Washington, DC: Bureau of Labor Statistics.

Camp, C. C., and Chartrand, J. M. (1992). A comparison and evaluation of interest congruence indices. *J. Vocat. Behav.* 41, 162–182.

Carpenter, D., Young, D. K., McLeod, A., and Maasberg, M. (2018). IT career counseling: are occupational congruence and the job characteristics model effective at predicting IT job satisfaction? *J. Inform. Syst. Educ.* 29, 225–238.

Destatis [Statistisches Bundesamt] (2018a). *Studierende an Hochschulen - Fächersystematik -*. Available at: https://www.destatis.de/DE/Methoden/Klassifikationen/BildungKultur/StudentenPruefungsstatistik.pdf?__blob=publicationFile (accessed October 30, 2018).

Destatis [Statistisches Bundesamt] (2018b). *Studienanfänger: Deutschland, Semester, Nationalität, Geschlecht, Studienfach* [*First-Year Students: Germany, Semester, Nationality, Gender, Study Subject] [xlx]*. Available at: https://www-genesis.destatis.de/genesis/online/data;jsessionid=78FDF13B4054B9F106F1C72BFFA22385.tomcat_GO_2_1?operation=abruftabelleBearbeiten&levelindex=1&levelid=1531475934506&auswahloperation=abruftabelleAuspraegungAuswaehlen&auswahlverzeichnis=ordnungsstruktur&auswahlziel=werteabruf&selectionname=21311-0012&auswahltext=%23Z-01.01.2011&werteabruf=Werteabruf (accessed July 13, 2018).

Durr, M. R., and Tracey, T. J. G. (2009). Relation of person–environment fit to career certainty. *J. Vocat. Behav.* 75, 129–138.

Eccles, J. S., and Wang, M. T. (2016). What motivates females and males to pursue careers in mathematics and science? *Int. J. Behav. Dev.* 40, 100–106.

Eder, F. (1998). "Differenziertheit der Interessen als Prädiktor der Interessenentwicklung [Differentiation of interests as a predictor of interest development]," in *Pädagogisch-Psychologische Interessenforschung in Studium und Beruf* [*Pedagogical-Psychological Interest Research in Study and Occupation*], eds J. Abel and C. Tarnai (Münster: Waxmann), 63–77.

Ertl, B., and Hartmann, F. G. (2019). Interest congruency of female STEM students regarding their vocational aspirations and their parents' occupations. *Paper Presented at the 2019 Annual Meeting on AERA*, Toronto, CA.

Ertl, B., Luttenberger, S., and Paechter, M. (2014). Stereotype als Einflussfaktoren auf die Motivation und die Einschätzung der eigenen Fähigkeiten bei Studentinnen in MINT-Fächern. *Gruppendyn. Organisationsberatung* 45, 419–440. doi: 10.1007/s11612-014-0261-3

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The Impact of Gender Stereotypes on the Self-Concept of Female Students in STEM Subjects with an Under-Representation of Females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Etzel, J. M., Lüdtke, O., Wagner, J., and Nagy, G. (2018). Similarity of vocational interest profiles within families: a person-centered approach for examining associations between circumplex profiles. *J. Pers.* doi: 10.1111/jopy.12418 [Epub ahead of print].

European Commission (2015). *She Figures 2015*. Available at: https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-final.pdf doi: 10.2777/744106

Eurostat (2018). *Mathematics, Science and Technology Enrolments and Graduates*. Available at: https://ec.europa.eu/eurostat/web/products-datasets/-/educ_thflds (accessed March 15, 2019).

FDZ-LIfBi (2018a). *Codebook. NEPS Starting Cohort 5—First-Year Students. From Higher Education to the Labor Market. Scientific Use File Version 10.0.0*. Available at: https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/10-0-0/SC5_10-0-0_Codebook_en.pdf (accessed October 23, 2018).

FDZ-LIfBi (2018b). *Data Manual. NEPS Starting Cohort 5—First-Year Students. From Higher Education to the Labor Market. Scientific Use File Version 10.0.0*. Available at: https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC5/10-0-0/SC5_10-0-0_Datamanual.pdf (accessed October 23, 2018).

Ferriman, K., Lubinski, D., and Benbow, C. P. (2009). Work preferences, life values, and personal views of top math/science graduate students and the profoundly gifted: developmental changes and sex differences during emerging adulthood and parenthood. *J. Pers. Soc. Psychol.* 97, 517–532. doi: 10.1037/a0016030

Gottfredson, G. D., and Holland, J. L. (1991). *PCI. Position Classification Inventory*. Odessa, FL: Psychological Assessment Resources.

Gottfredson, L. S. (1981). Circumscription and compromise: a developmental theory of occupational aspirations. *J. Couns. Psychol.* 28, 545–579. doi: 10.1037/0022-0167.28.6.545

Gottfredson, L. S. (2005). "Applying Gottfredson's theory of circumscription and compromise in career guidance and counseling," in *Career Development and Counseling. Putting Theory and Research to Work*, eds S. D. Brown and R. W. Lent (Hoboken, N.J: John Wiley & Sons), 71–100.

Hartmann, F. G. (2018). *Analysen zur Übereinstimmung beruflicher Interessen in der Familie. Eine Studie zur Bestimmung der dyadischen Ähnlichkeit bei mehrdimensionalen Konstrukten im familialen Kontext* [*Analyzes of the Similarity of Vocational Interests in the Family. A Study to Determine the Dyadic Similarity of Multidimensional Constructs in the Familial Context*]. Münster: Waxmann.

Hartmann, F. G., Tarnai, C., and von Maurice, J. (2015). "Berufliche Interessen in der Kernfamilie [Vocational interests in the nuclear family]," in *Berufliche Interessen. Beiträge zur Theorie von J. L. Holland* [*Vocational Interests. Contributions on the Theory of J. L. Holland*], eds C. Tarnai and F. G. Hartmann (Münster: Waxmann), 87–114.

Healy, C. C., and Mourton, D. L. (1983). Derivatives of the Self-Directed Search: potential clinical and evaluative uses. *J. Vocat. Behav.* 23, 318–328.

Hingel, A., Saltelli, A., and Mercy, J.-L. (2008). *Progress Towards the Lisbon Objectives in Education and Training. Indicators and benchmarks 2008*. Brussels: Commission of the European Communities.

Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement: II. A four-year prediction study. *Psychol. Rep.* 12, 547–594.

Holland, J. L. (1997). *Making Vocational Choices. A Theory of Vocational Personalities and Work Environments*. Lutz, FL: Psychological Assessment Resources.

Hubert, L., and Arabie, P. (1987). Evaluating order hypotheses within proximity matrices. *Psychol. Bull.* 102, 172–178.

Iachan, R. (1984). A measure of agreement for use with the Holland classification system. *J. Vocat. Behav.* 24, 133–141.

Junk, K. E., and Armstrong, P. I. (2010). Stability of career aspirations: a longitudinal test of Gottfredson's theory. *J. Career Dev.* 37, 579–598.

Konrad, A. M., Ritchie, J. E., Lieb, P., and Corrigall, E. (2000). Sex differences and similarities in job attribute preferences: a meta-analysis. *Psychol. Bull.* 126, 593–641. doi: 10.1037/0033-2909.126.4.593

Lent, R. W. (2013). "Social cognitive career theory," in *Career Development and Counseling. Putting Theory and Research to Work*, eds S. D. Brown and R. W. Lent (Hoboken, N.J: John Wiley & Sons), 115–146.

Lent, R. W., and Brown, S. D. (2013). Social cognitive model of career self-management: toward a unifying view of adaptive career behavior across the life span. *J. Couns. Psychol.* 60, 557–568. doi: 10.1037/a0033446

Lent, R. W., Brown, S. D., and Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* 45, 79–122.

Luttenberger, S., Aptarashvili, I., Ertl, B., Ederer, E., and Paechter, M. (2014). Niedrige Übereinstimmung zwischen Interessen und Berufswunsch: Ein bislang vernachlässigtes Risiko in der Berufsorientierung Jugendlicher. *Gruppendyn. Organisationsberatung* 45, 359–377. doi: 10.1007/s11612-014-0252-4

Morris, M. L. (2016). Vocational interests in the United States: sex, age, ethnicity, and year effects. *J. Couns. Psychol.* 63, 604–615. doi: 10.1037/cou0000164

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16474–16479. doi: 10.1073/pnas.1211286109

Neumann, G., Olitsky, N., and Robbins, S. B. (2009). Job congruence, academic achievement, and earnings. *Labour Econ.* 16, 503–509.

Nye, C. D., Su, R., Rounds, J., and Drasgow, F. (2012). Vocational interests and performance. A quantitative summary of over 60 years of research. *Perspect. Psychol. Sci.* 7, 384–403. doi: 10.1177/1745691612449021

O*NET (2018). *O*NET® 22.3 Database [xls]*. Available at: https://www.onetcenter.org/database.html?p=2 (accessed March 5, 2019).

O*NET OnLine (2018). *Browse STEM Occupations*. Available at: https://www.onetonline.org/find/stem?t=0 (accessed March 26, 2019).

Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *J. Vocat. Behav.* 21, 259–287.

Robbins, P. I., Thomas, L. E., Harvey, D. W., and Kandefer, C. (1978). Career change and congruence of personality type: an examination of DOT-derived work environment designations. *J. Vocat. Behav.* 13, 15–25.

Rottinghaus, P. J., Larson, L. M., and Borgen, F. H. (2003). The relation of self-efficacy and interests: a meta-analysis of 60 samples. *J. Vocat. Behav.* 62, 221–236.

Rounds, J., Su, R., Lewis, P., and Rivkin, D. (2013). *Occupational Interest Profiles for New and Emerging Occupations in the O*NET System: Summary*. Available at: https://www.onetcenter.org/dl_files/OIP_NewEmerging.pdf (accessed June, 2013).

Sheu, H.-B., Lent, R. W., Brown, S. D., Miller, M. J., Hennessy, K. D., and Duffy, R. D. (2010). Testing the choice model of social cognitive career theory across Holland themes: a meta-analytic path analysis. *J. Vocat. Behav.* 76, 252–264.

Steinwede, J., and Aust, F. (2012). *Methodenbericht. NEPS Startkohorte 5 - CATI-Haupterhebung Herbst 2010 B52*. Bonn: infas Institut für angewandte Sozialwissenschaften.

Su, R., and Rounds, J. (2015). All STEM fields are not created equal: people and things interests explain gender disparities across STEM fields. *Front. Psychol.* 6:189. doi: 10.3389/fpsyg.2015.00189

Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychological Bulletin* 135, 859–884. doi: 10.1037/a0017364

Su, R., Tay, L., Liao, H.-Y., Zhang, Q., and Rounds, J. (2018). Toward a dimensional model of vocational interests. *J. Appl. Psychol.* doi: 10.1037/apl0000373 [Epub ahead of print].

Super, D. E. (1957). *The Psychology of Careers: An Introduction to Vocational Development*. New York, NY: Harper & Row.

Tracey, T. J. G. (1997). RANDALL: a microsoft FORTRAN program for a randomization test of hypothesized order relations. *Educ. Psychol. Meas.* 57, 164–168.

Tracey, T. J. G., Allen, J., and Robbins, S. B. (2012). Moderation of the relation between person-environment congruence and academic success: environmental constraint, personal flexibility and method. *J. Vocat. Behav.* 80, 38–49.

Tracey, T. J. G., and Robbins, S. B. (2005). Stability of interests across ethnicity and gender: a longitudinal examination of grades 8 through 12. *J. Vocat. Behav.* 67, 335–364.

Tracey, T. J. G., and Robbins, S. B. (2006). The interest-major congruence and college success relation: a longitudinal study. *J. Vocat. Behav.* 69, 64–89.

Tracey, T. J. G., Robbins, S. B., and Hofsess, C. D. (2005). Stability and change in interests: a longitudinal study of adolescents from grades 8 through 12. *J. Vocat. Behav.* 66, 1–25.

Tracey, T. J. G., and Sodano, S. M. (2013). "Structure of interests and competence perceptions," in *Handbook of Vocational Psychology: Theory, Research, and Practice*, eds W. B. Walsh, M. L. Savickas, and P. J. Hartung (New York, NY: Taylor & Francis), 104–120.

Tracey, T. J. G., Wille, B., Durr, M. R. II, and De Fruyt, F. (2014). An enhanced examination of Holland's consistency and differentiation hypotheses. *J. Vocat. Behav.* 84, 237–247.

Tranberg, M., Slane, S., and Ekeberg, S. E. (1993). The relation between interest congruence and satisfaction: a meta-analysis. *J. Vocat. Behav.* 42, 253–264.

Tsabari, O., Tziner, A., and Meir, E. I. (2005). Updated meta-analysis on the relationship between congruence and satisfaction. *J. Career Assess.* 13, 216–232.

Wang, M.-T., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice: individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937

Watt, H. M. G., Hyde, J. S., Petersen, J., Morris, Z. A., Rozek, C. S., and Harackiewicz, J. M. (2017). Mathematics – a critical filter for STEM-related career choices? A longitudinal examination among Australian and U.S. adolescents. *Sex Roles* 77, 254–271. doi: 10.1007/s11199-016-0711-1

Wille, B., Tracey, T. J. G., Feys, M., and De Fruyt, F. (2014). A longitudinal and multi-method examination of interest-occupation congruence within and across time. *J. Vocat. Behav.* 84, 59–73.

Wohlkinger, F., Ditton, H., von Maurice, J., Haugwitz, M., and Blossfeld, H.-P. (2011). 10 Motivational concepts and personality aspects across the life course. *Z. Erziehungswiss.* 14:155. doi: 10.1007/s11618-011-0184-5

Woodcock, A., Graziano, W. G., Branch, S. E., Habashi, M. M., Ngambeki, I., and Evangelou, D. (2013). Person and thing orientations: psychological correlates and predictive utility. *Soc. Psychol. Pers. Sci.* 4, 116–123. doi: 10.1177/1948550612444320

Xu, H., and Tracey, T. J. G. (2016). Stability and change in interests. A longitudinal examination of grades 7 through college. *J. Vocat. Behav.* 93, 129–138.

Yang, Y., and Barth, J. M. (2015). Gender differences in STEM undergraduates' vocational interests: People–thing orientation and goal affordances. *J. Vocat. Behav.* 91, 65–75. doi: 10.1016/j.jvb.2015.09.007

Young, G., Tokar, D. M., and Subich, L. M. (1998). Congruence revisited: Do 11 indices differentially predict job satisfaction and is the relation moderated by person and situation variables? *J. Vocat. Behav.* 52, 208–223.

Zener, T., and Schnuelle, L. (1976). Effects of the self-directed search on high school students. *J. Couns. Psychol.* 23, 353–359.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 01 May 2019
doi: 10.3389/fpsyg.2019.00968

Pipped at the Post: Knowledge Gaps and Expected Low Parental IT Competence Ratings Affect Young Women's Awakening Interest in Professional Careers in Information Science

*Angela Schorr**

Media Psychology Lab, Institute of Psychology, Faculty II, University of Siegen, Siegen, Germany

OPEN ACCESS

Edited by:
Rebecca Lazarides, Universität Potsdam, Germany

Reviewed by:
Frances Hoferichter, University of Greifswald, Germany
Päivi Hannele Taskinen, Friedrich Schiller University Jena, Germany

***Correspondence:**
Angela Schorr
angela.schorr@uni-siegen.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *09 January 2019*
Accepted: *11 April 2019*
Published: *01 May 2019*

Citation:
Schorr A (2019) Pipped at the Post: Knowledge Gaps and Expected Low Parental IT Competence Ratings Affect Young Women's Awakening Interest in Professional Careers in Information Science. Front. Psychol. 10:968. doi: 10.3389/fpsyg.2019.00968

Although many jobs in today's information science allow favorable work-life-schedules for women, they still hesitate to enter this territory. In a study based on individual interviews with $N = 134$ students aged 14–18 years, who visited the *Deutsches Museum* in Munich, Germany, we collected data on the students' socialization in information and communication technology (ICT), on their self-rated ICT competence, their working knowledge of ICT professions, and their reaction to sexist statements. To analyze more in depth, we provided the participants with two alternative forms of vocational counseling interventions designed to modify their ICT-related attitudes (information vs. robotics condition). Analyses of variance and multiple linear regressions were administered to the data. Results: The girls in this study were socialized more than one year after the boys in using computers. While the boys received their ICT training mostly through their fathers and peers, the girls frequently had to rely on their teachers for ICT instruction. The girls rate their ICT competence lower than the boys; nevertheless, both genders share a relatively high interest in ICT professions. What's more, the girls are less convinced that men have a natural talent for computer science. Openness toward taking up jobs in the ICT industry in the case of the boys is less determined by their self-rated computer competence and the perceived ICT talent assessment by their parents. In both intervention conditions, they eagerly received and processed the new information provided. The girls' interest in an ICT career largely depends on preconditions, namely on their self-rated ICT competence, on a long-standing enthusiasm for computers, and on what they perceive their parents think about their ICT talents. Unlike the more pragmatic approach of the boys, their self-doubts, especially among the academic high school girls brings about that they are still in danger to leave the field of information/computer science before having entered it. In general, the participants' responses point to a comprehensive misdirection of young women in German middle schools and academic high schools. Fortunately, this study provides a lot of evidence on how to fix this major mishap in the interest of both sexes.

Keywords: computer competence, computer affinity, gender stereotypes, ICT socialization, ICT professions, robotics, attitude change, vocational counseling

INTRODUCTION

How is it, that so few young women are choosing information and communication technology (ICT) careers? Graduates of professions in the ICT sector (either with a middle school/MS diploma plus a vocational training in information technology or with an academic high school/AHS diploma plus a university degree in computer science) are very sought after. The job prospects are excellent. Training in media professions is in vogue, while the computer science economy in Germany has recruitment problems, which cannot be explained solely by the growing demand for IT services. Although many fields of work in the ICT area are characterized by high space-time independence and flexibility and thus can be classified as family friendly, young women in particular seldom opt for these careers.

Young Women in Computer Science: Self-Doubts and Ambivalent Feelings of Belonging

Already at the turn of the millennium German experts were able to identify potentially important factors based on representatively collected survey data, which could explain the small proportion of women in computer science studies – at that time their share was 7–8% (Schinzel et al., 1999; Ihsen et al., 2013). It turned out that the parents and the school seemed to inhibit the young women to become interested in computer science. "Accordingly, women often decide to study computer science only after graduation. . ." (Schinzel et al., 1999, p. 13). These ICT-inclined young women started their course of studies with less confidence, especially in terms of computer knowledge, programming and software. More often than their male counterparts, they said, they felt they were inferior despite good performances. Less often, as was the case with male classmates hobbies and work overlapped. Overall, the authors concluded: "The women are much more doubtful about their abilities and suitability for the subject. The conversations among the students unsettle them, especially in the early phase of their studies" (Schinzel et al., 1999, p. 13, translated by the author). These findings are consistent with studies in other countries and other STEM fields. Despite equally good performances, in computer science as well as in other STEM fields like e.g., engineering young women regularly show lower competence perceptions compared to young men (Reid, 2009; Jagacinski, 2013). That these research results continue to be relevant today is shown by a study recently published by Ito and McPherson (2018). The authors surveyed 386 high school students and their results add another piece to the big puzzle: For female students but not for male students, they could show that the self-rated likelihood of persisting in the current STEM class, to pursue the topic further in high school and college, and to pursue a career in STEM can be significantly predicted by the social belonging dimension, here defined as feeling of acceptance and fit with the STEM class.

It remained unclear how exactly the socialization processes in family, school, vocational training and studies run their course, keeping young women in Germany away from the computer science professions (Ebach, 1994; Schinzel and Ruiz Ben, 2004; Maaß and Wiesner, 2006; Ehrke et al., 2011). Two decades of research culminated in compendiums with more than 30 differentiated recommendations for gender-appropriate higher education (Ihsen et al., 2017; Resch et al., 2017). Instead of tackling the root of the problem, experts and stakeholders such as the German industry association *bitkom,* given the shortage of skilled workers, increasingly recommend mono-educational courses and pure women's study courses (Nordmann, 2016; Bitkom Press, 2018). These are approaches that the female computer science students surveyed at the turn of the millennium already rejected (Schinzel et al., 1999). In 2018, still only 17% of ICT professionals in Germany are women.

Research on the Impact of Parents, Teachers, and School Curricula

Ihme and Senkbeil (2017) carried out a scientific study with 224 students at middle schools in the 8th grade, half girls and boys. They come to the conclusion that boys rate their computer-related skills significantly higher than girls, even though there are in fact no differences in competencies. For the traditional STEM field of mathematics Gunderson et al. (2012a,b) based on a comprehensive research review conclude that parents' and teachers' expectancies for children's math competence are still often gender biased. The results of a recent representative study by Lloyd et al. (2018) on the informal learning world "family" are pointing in the same direction. They looked at 6,492 Australian children aged 3–12 and came to the sobering conclusion, "Even when parents created a supportive environment, there was little evidence indicating that girls were encouraged to pursue STEM" (Lloyd et al., 2018, p. 308). In addition to the teachers, these days especially the fathers are targeted by researchers. In the study Galdi et al. (2017) published on the family and school environment in 68 six-year-old children, they could prove that the fathers', but not the mothers', stereotypical beliefs predicted children's stereotypes.

As far as German students are concerned, Ihme and Senkbeil (2017) mainly blame the formal learning world "school." They found that the shortcomings of school education have serious consequences for girls and boys alike: Due to the missing curricular implementation of computer-related skills in Germany and the resulting lack of feedback, adolescents tend to overestimate their abilities in the area of computer-related skills. As a result, their positive self-assessment depends on the frequency of computer use and leads to a skills illusion. German schools hardly have a chance to compensate for social disparities in the acquisition of computer-related skills. In everyday life, according to the authors, the adolescents acquire these skills mainly through the help of their parents.

Research on the Impact of Gender Stereotypes, Stereotype Threat, and Domain Masculinity

Part of the current research on gender issues in the STEM domains is dedicated to researching the impact of gender stereotypes, stereotypical threat and domain masculinity on girls and young women. Both, the internalization of inferiority images and their consequences (gender stereotypes) as well as the reactions on the immediate situational threat that derives from negative stereotypes disseminated of one's group (stereotype threat) have an effect on young women in the STEM field. They display cognitive, physiological and performance-related responses like increased stress, lack of self-esteem, increased self-doubt, failure in performance, etc., (Murphy et al., 2010; Régner et al., 2016). This is particularly true of domains in which women are already negatively stereotyped, respectively, domains that are stereotyped as masculine (Schmader et al., 2004; Thoman et al., 2008; Cheryan et al., 2009; Logel et al., 2009; Cheryan, 2012; Hergatt Huffman et al., 2013; Ihsen, 2013).

Today gender parity in school or academic achievements in STEM fields, meaning that girls today are performing just as well or even better as boys in these fields, has been achieved (e.g., Kim et al., 2014). Nevertheless, women continue to choose STEM-related careers at significantly lower rates than men do (Cheryan, 2012; Hergatt Huffman et al., 2013; Régner et al., 2016). Hergatt Huffman et al. (2013) can prove in their study that not only the biological sex plays the decisive role, but rather gender roles, specifically masculinity. Cheryan (2012) concludes "that considering the extent to which math-related domains are stereotyped as masculine can help explain why women do not seek out math-related careers, even as they perform just as well in math" (Cheryan, 2012, p. 184).

Another recent finding on gender interaction in computer science studies confirms that too little progress is being made. Förtsch et al. (2018) in the *Bamberg Alumnae Tracking Study* asked the question: "Do male and female graduates differ in their level of academic achievement in computer sciences? And if not, does ICT change anything?" They found that although the academic achievements of female graduates are as good as those of male graduates, still "female graduates exhibit lower self-belief in their professional skills, partly because lower-achieving male graduates still display very high professional skills self-efficacy beliefs, irrespective of their previous academic achievements at university. (. . .) The career ambitions and career opportunities of male graduates depend less on their academic achievements at university, whereas female graduates have to be very ambitious to be able to hold a leadership position in the same field" (Förtsch et al., 2018, p. 265).

Can the Interest in ICT Counteract the Masculine Image of Computer Science? Research on Attitude Change

However, there is also evidence from research that something is changing in the generation of students studied, that both genders approach the ICT domain on their own and because of intrinsic motivation: Sáinz et al. (2016) investigated young people's gender stereotypes and attitudes about people working in the field of ICT. 900 pupils on average 15 years old were interviewed. Both boys and girls held stereotypical beliefs about ICT as a highly male-dominated field. "As expected, these stereotypical beliefs described a masculine portrayal of ICT workers. Contrary to the expectations, most of the students' portrayals of people working in ICT were either positive or neutral, not negative" (Sáinz et al., 2016, p. 154). No gender differences were observed in the type of characteristics associated with ICT professionals. Another indication of how young women are actively trying to approach the ICT domain lies in the observation of Sáinz et al. (2016), that "young females were more likely to offer feminine references about the professions where ICT is the tool rather than the object of their work" (Sáinz et al., 2016, p. 154). A similar observation is made by Lasen (2010) in focus group interviews with pupils in the 11th and 12th grade. Both, the group of ICT course participants (takers) as well as the group of those who did not attend ICT courses (non-takers), expressed an aversion to programming. But the interviews also showed that those who did not attend ICT courses misunderstood the purpose of the courses and interpreted them as geared to programming and other highly technical skills only. In the case of the female takers, it was the creative aspects of information technology systems which had attracted them to the subject and they were in fact enjoying ICTs authentic, problem-based design tasks. Lasen concludes, "Findings indicate that schoolgirls' participation in ICT pathways may well be promoted through subjects that position and call for students to engage with ICTs as 'enablers' in diverse, meaningful and creative human contexts" (Lasen, 2010, p. 1117).

Why It Is Necessary to Focus on the Topic of Career Choices and Vocational Counseling of Young People Interested in ICT

The extensive digitization of everyday life is changing the attitudes of young people toward ICT professions, but in Germany these changes are only slowly taking place. There is a lack of personal and professional support for young women to become interested in a job in the ICT sector. The well-known deficits in ICT socialization are "repaired" with more or less commitment, instead of tackling the problems at all levels, at all ages. To make progress, one needs to analyze the situation more closely.

Generally, young people between the ages of 14 and 18 for the first time want to find out what they want to do later in their lives. The interest in professional advice is high. Also for this reason, the *Deutsches Museum*, which is a technical museum, is a popular destination for young people of this age group. So far, little research has been done on this important orientation phase for the career decision of young people interested in learning something about the ICT sector.

Previous research focused on psychological dimensions such as academic self-concept, perceived parental support, stereotypical threat, sense of social belonging and others (e.g., Nagy et al., 2010; Heerwegh et al., 2016; Ito and McPherson, 2018). For about 15 years, various research teams attempted to elucidate the complex pattern of biographical, family related, and school-related factors affecting the ICT socialization of girls and young women. A number of elaborate studies guided by different research interests were published. In these studies, either primary school age children were included (Meelissen and Drent, 2008; Vekiri and Chronaki, 2008), or college freshman as well as undergraduate and graduate students (e.g., Verhoeven et al., 2010; Heerwegh et al., 2016; Ertl et al., 2017). A good example of these approaches is an exciting recent study by Ertl et al. (2017). They focus on the effects of gender stereotypes on the self-concept of female students in STEM degree programs with less than 30% females. On the topic of academic self-concept in mathematics, Nagy et al. (2010) published an insightful study, which – by the way of exception – specifically covered the age group of 13 to 18 year olds (see also Miliszewska and Sztendur, 2010).

For this study, we have specifically selected the age group of 14 to 18 year olds. The focus of the investigation is thus on the life phase, in which important decisions for the future profession are made. The focus on the ICT professions and the interest in these professions is at the center. Our central research question in this project was: Which factors have an impact on the career choice of young women and young men in terms of their interest in IT careers (e.g., ICT socialization in family and school, parental support, ICT affinity, computer competence, working knowledge of the ICT professions, gender stereotypes or others)? And: How can we change some of these learned, but "dysfunctional" attitudes toward the ICT sector? – Positive results on attitude changes as regards to the attractiveness and fit of the ICT domain for young women appear at least short-term and intervention-based possible. With regard to the research results on gender stereotypes, it is assumed that these continue to be effective. Starting with the topic of young womens' ICT socialization, the following hypotheses guided the investigation:

(1) The ICT socialization, i.e., the individual media use biography and the self-assessed computer and Internet competence have an influence on young womens' interest in computer science professions. Especially, the informal learning environment "family" has an impact on young womens' attitudes toward the ICT field.
(2) In Germany, the formal learning environment "school" currently cannot compensate for the deficits in the ICT socialization of girls and young women.
(3) Informative and targeted vocational counseling interventions can change the interest in ICT professions.
(4) Interest in ICT is still influenced by gender stereotypes.

MATERIALS AND METHODS

Design of the Study

This is an interview study in which vocational counseling interventions are integrated. It follows the tradition of educational intervention research (cf. Hunt and Hunt, 2004; Mansoori-Rostam and Tate, 2017; Shiina et al., 2017; Turetsky and Sanderson, 2017). One focus of the study is on the coverage and analysis of the participants' complex ICT biography (family/school) in connection to gender stereotypes and the participants' interest in ICT professions. The other focus is on the two vocational counseling intervention conditions we applied. For years and based on careful analyses of research, intervention measures to eliminate gender gaps in the field of computer science were proposed (e.g., Quaiser-Pohl, 2012; Ihsen et al., 2017). New to this study is that by focusing on the psychological dimension of "Interest in ICT professions," we investigated in two different intervention conditions intended to effect attitude change.

In the German Science and Technology Museum (*Deutsches Museum*) $N = 134$ students were recruited for the project. The participants were approached by the interviewers at the ICT laboratory of the Technical University of Munich (TUMLab) and motivated to participate in the project. The activities in the TUMLab can be seen from the outside through a fully glazed wall for every visitor. At a first glance it looks like a standard computer lab with 15–20 seats. Under the motto "Science to touch and experience" guided automation courses, programming courses, etc., can be booked by schools and other interested parties. However, our respondents did not take part in such courses. They passed the Lab and observed the activities inside for shorter or longer time from the outside.

The interview started with some open questions on the participants' museum experience (What was already looked at? What was particularly impressive? What was entertaining and fun?). These questions served as a warming up and were not evaluated for the investigation. Next, we applied several questionnaires (on the participant's ICT socialization at home and in school, on the self-rated ICT competence etc.). Subsequently, a questionnaire capturing the interest in ICT professions and a short scale on gender stereotypes were applied. Then the vocational counseling intervention (one of two conditions) was carried out and this was followed by a second application of the questionnaire measuring the interest in ICT professions and the gender stereotypes scale (see **Figure 1**). Overall, the interview plus intervention took two to 3 h per participant. Two alternative methods of informing the participants about the ICT field were tested in the intervention phase of the study (information condition vs. robotics condition). The participants were randomly assigned to these conditions. Each participant was accompanied by a member of the research

Recruitment →	Individual interview (all participants) →	Intervention phase →	Follow-up survey (all participants)
the participants were approached as they passed the TUMLab;	application of a questionnaire on ICT socialization in family and school	information condition	
visitors 14 to 18 years old, attending or having completed middle school or	application of a questionnaire on interest in ICT	vs.	application of a questionnaire on interest in ICT professions (time 2)
academic high school were invited to participate in an ICT vocational counseling intervention	professions (time 1) application of a questionnaire on gender stereotypes (time 1)	robotics condition	application of a questionnaire on gender stereotypes (time 2)

FIGURE 1 | Temporal sequence of the investigation.

team who guided her/him through the various stages of the investigation.

Participants

We selected the participants for the study based on three criteria: age (14–18 years), gender (half male and female participants) and education (ongoing or completed education at middle school or academic high school level). Qualifications at this level in Germany qualify for practical trainings as an IT specialists or a university education in computer science. Young people at this age and level of education know that they need to make career-oriented training choices in the near future. In terms of gender, we have strived for equal group sizes.

Concept of the Vocational Counseling Interventions

The core content of the two alternatively explored vocational counseling interventions is counseling for the computer science professions. In both interventions the perceived and actual requirements in training and at work were discussed, as well as future fields of specialization and exciting and challenging aspects of the profession (see also Ehrke et al., 2011; Conein and Schwarz, 2015). The need of the participants to make important educational decisions for their lives in the near future was addressed. All contents were presented orally, with a focus on dialogue with the participant.

A total of $N = 107$ participants in the information intervention took part in an ICT career counseling. The following topics were discussed: How to become an IT specialist based on a middle school diploma or an academic high school diploma; authentic, positive statements from computer scientists as well as from employers on the ICT working world, on supposed barriers and on little-known facts from the professional life of IT specialists. Finally, the promotion and earning opportunities and the compatibility of work and family life (flexible working hours, working from a home office) as well as professional, social and communicative skills needed as prerequisites for the profession were explained.

The total of $N = 27$ participants in the robotics intervention joined in a stimulating, knowledgeable, guided tour of the robotics exhibition, which showed the diverse applications of robots in everyday life (luggage robots, household robots, welding robots, milking robots, magnetic tape storage, medical prostheses, etc.). To arouse interest in the computer science professions among the participants, we focused on descriptive explanations based on the exhibits. They were explained in terms of their design, operation, manufacturing and uses, always referring to the tasks of the computer specialists involved. The tour was organized in groups of three people plus a guide from the research team. Due to the public traffic in the exhibition hall it was often very loud there. In the trial week before the study we selected timeslots, to which the hall during the day had little public traffic. In order to facilitate the communication of the contents and the dialogue with the participants, we waited for the implementation until there were few or no persons in the hall. As a result, only a relatively small number of participants were able to take part during the research week. This explains the low number of participants in the robotics condition.

Development of the Research Scales

All survey instruments used in this study are listed together with the items in **Tables 2**, **3**, **5–9**. Computer affinity was recorded with a short *Computer affinity scale* ($\alpha = 0.80$; $N = 134$). The self-perceived mathematical competence was captured with the short *Scale on math competence* ($\alpha = 0.68$; $N = 134$). Both items, here slightly reworded and turned positive, are based on the Self-appraised math ability scale Schmader et al. (2004) used with good results. For both dimensions of computer affinity and math competence, longer scales exist that were not used here for reasons of survey economy (see Meelissen and Drent, 2008; Heerwegh et al., 2016; Wille et al., 2018).The three questions on computer literacy and the three questions on Internet literacy were newly developed based on items found in the "ICT Familiarity Questionnaire" [see PISA, (OECD, 2015)]. From these six items the *Maintenance competence scale* (4 items) was obtained by factor analysis ($M = 15.42$, $SD = 3.41$; $\alpha = 0.76$); the remaining two items indicating the readiness to seek/accept external help and intercorrelate with $r(134) = 0.54$, $p < 0.001$.

There are a number of perceived parental beliefs/parental attitudes scales, all of which have been used in the context of research on gender stereotypes in mathematics (see Wendland and Rheinberg, 2004; Lazarides and Ittel, 2013; Lazarides et al., 2016; Lazarides and Watt, 2017). The *Perceived parental attitudes scale* used in this study consists of six items. It is a short form of the 8-item "Perceived parental support scale" by Vekiri (2010). Vekiri's scale had a good Cronbach's alpha of $\alpha = 0.84$. Due to the reduction made for reasons of survey economy, the homogeneity of our shortened scale has decreased, but is still acceptable ($M = 20.59$, $SD = 4.13$; $\alpha = 0.69$). Both the 6-items scale and the scales based on each item of this scale were included in the statistical analysis.

The nine statements on the qualities of School computer courses/computer science courses after factor-analytic treatment and reduction by two items led to a new scale named *Qualities of the ICT courses* ($M = 21.15$, $SD = 5.74$; $\alpha = 0.79$). These items, as well as the items of the *Perceived ICT competence of ICT teachers scale*, were newly developed for the study. The *Perceived ICT competence of ICT teachers scale* in the factor analysis proved to be one-factorial and has a high degree of homogeneity ($M = 22.60$, $SD = 8.55$; $\alpha = 0.92$).

In the past, questions about interest in ICT have been asked mainly in the context of large, representative surveys (e.g., PISA studies, OECD reports, JIM studies). From the start, STEM research also dealt with this topic. In this study the topic is accessed in a special way. It is not just about the interest in ICT, but about the interest in ICT professions and a career in ICT! The *Interest in ICT professions scale* consists of a total of 17 items. Originally developed for the engineering profession, six items were taken from a study by Kessels and Hannover (2004); two items were taken from a study by McLachlan et al. (2010) on middle school students' attitudes to ICT. Nine items were newly created, based on discussions with experienced computer scientists, career counselors and computer science professors as well as on research results from ICT gender research. The new questionnaire was first time applied in this study. As expected, the factor analysis confirmed a multidimensional structure of the scale. For the purposes of the present study, however, the scale was only used in its entirety (1st measurement: $M = 59.11$, $SD = 7.96$; $\alpha = 0.78$; 2nd measurement: $M = 60.05$, $SD = 8.31$; $\alpha = 0.81$).

Finally, some questions were presented to the participants in order to grasp how deeply gender stereotypes are still anchored in these young participants. The three questions used here have been used more often in gender research, most recently by Papastergiou (2008). The total scale formed from the three items called $Gender_{tot}$ *scale* reached a high Cronbach's alpha of $\alpha = 0.80$ on the 1st measurement ($M = 8.26$, $SD = 3.43$) and of $\alpha = 0.78$ on the 2nd measurement ($M = 8.37$, $SD = 3.50$) in our study.

Statistically, analyses of variance with the factors AGE (14, 16, and 18 years of age), SEX and SCHOOL (middle school vs. academic high school), *t*-tests and Chi^2-tests, factor analyses, intercorrelations and multiple linear regression analyses were administered to the data.

RESULTS

ICT Socialization, User Biography and Self-Assessment of Computer and Internet Competence

To find out what role does ICT socialization (i.e., the individual media use biography) play in the interest in ICT professions, we systematically interviewed the participants. First, we tried to determine how important computers are to their lives. This was followed by another standard question of STEM research, namely the question of the self-esteemed talent for computers. The affinity to computers was examined in the context of two questions (1 = strongly disagree, 5 = strongly agree): "*Do you agree with the statement: The computer was my favorite hobby right from the start!?*" (Hobby 1; $M = 2.77$, $SD = 1.16$) and "*Do you agree with the statement: The computer is my favorite hobby today!?*" (Hobby 2; $M = 2.84$, $SD = 1.12$). In both statements, the participants signaled a barely average interest. They were grouped into the short *Computer affinity scale* ($hobby_{tot}$), as already explained in the "Materials and Methods" section. A two-way analysis of variance with the factors SEX × SCHOOL revealed no significant main effects for the *Computer affinity scale*, but a significant interaction, $F(1,130) = 8{,}32$, $p = 0.005$, $\eta_p^2 = 0.06$. The highest approval was given by the male academic high school students ($M = 6.14$), the lowest by the female academic high school students ($M = 5.00$).

In the self-assessed math competence, recorded with the statement "*I am good at solving math problems*," the two-way analysis of variance SEX × SCHOOL ($M = 3.41$, $SD = 1.19$) showed a significant main effect for the factor SEX, $F(1,131) = 4{,}12$, $p = 0.044$, $\eta_p^2 = 0.034$, i.e., the girls rated themselves significantly less competent than the boys. The second statement "*It is important to me that I perform well in math problems.*" the two-way analysis of variance SEX × SCHOOL revealed equally high middle-level approval in both genders ($M = 3.78$, $SD = 1.10$), i.e., there were no significant differences for the main effects and the interaction. Since both items intercorrelate with $r(134) = 0.52$, $p < 0.001$, they were grouped into a short *Scale on math competence* ($math_{tot}$, $M = 7.19$, $SD = 2.00$), as already explained in the "Materials and Methods" section.

Asked about the current computer/laptop ownership and the frequency of Internet use, the participants do not differ. However, their user biography differs amazingly in terms of the first PC use and the first Internet use by age and gender. These two biographical questions can be found in the ICT Familiarity Questionnaire for PISA 2015.

Table 1 shows, based on a three-way analysis of variance AGE × SEX × SCHOOL, how the first laptop/PC use and

TABLE 1 | Small media use biography (*N* = 134).

First PC use	14 years	16 years	18 years
(in age; months)	8;1	9;3	9;4
	Male		Female
	8;5		9;5
First Internet use	14 years	16 years	18 years
(in age; months)	9;6	10;8	10;9

There were no gender differences with regard to the first Internet use.

TABLE 2 | Computer and Internet competence.

	M	***SD***	**Three-way analyses of variance** **Age × Sex × School (*N* = 133)**
Keeping *the computer* running (e.g., download virus scanner): Is it true that . . .?			
(1) I can keep my computer running all alone in all important areas!	3.71	1.09	Sex: $F(1, 122) = 11.57$, $p = 0.000$, $\eta_p^2 = 0.087$, $M^{boys} = 4.07$, $M^{girls} = 3.28$
(2) If I'm unsure or do not know something about the operation of my computer, I have no trouble getting help!	4.31	1.09	n.s.
(3) If I'm unsure or do not know something about the operation of my computer, I try to solve the problem myself!	3.68	1.16	Sex: $F(1, 122) = 17.30$, $p = 0.000$, $\eta_p^2 = 0.124$, $M^{boys} = 4.08$, $M^{girls} = 3.20$
Using *the Internet* (for example, searching for information from a search engine): Is it true that. . .?			
(4) I can do everything that matters on the Internet (visit chat rooms, download sound files, maintain the homepage, Skype, use SchülerVZ, post photos/videos, etc.)!	4.45	0.91	n.s.
(5) If I'm unsure or do not know something about the security, individual programs, offers etc., on the Internet, I have no problems getting help!	4.13	1.10	n.s.
(6) If I'm unsure or do not know something about the security, individual programs, offers etc., on the Internet, I try to solve the problem myself!	3.58	1.23	Sex: $F(1, 122) = 11.73$, $p = 0.001$, $\eta_p^2 = 0.088$, $M^{boys} = 3.96$, $M^{girls} = 3.13$

n.s.: no significant differences between the groups; scale: 1 = strongly disagree; 5 = strongly agree.

the first Internet use with falling age of the respondents drops significantly: the younger the respondents, the earlier the onset (first computer use: $F(2, 121) = 3.64$, $p = 0.03$, $\eta_p^2 = 0.057$, first Internet usage: $F(2, 121) = 5.25$, $p = 0.006$, $\eta_p^2 = 0.075$). Relevant for the ICT socialization in the family is also the statistically significant main effect for the factor SEX at the first PC use. Compared to the boys, the girls in all three age groups first time used a PC with a delay of a whole year, $F(1, 121) = 6.79$, $p = 0.01$, $\eta_p^2 = 0.053$)! In addition, the two-way analysis of variance shows a significant interaction of the factors SEX and SCHOOL at the first PC use, $F(1, 121) = 7.34$, $p = 0.008$, $\eta_p^2 = 0.06$). The discrepancy is particularly high for the first use of computers among academic high school students. While the male academic high school students started comparatively early at an average age of 8;5 years, the female academic high school students only started on average 10;0 years!

Related to the results of the self-assessment of the participants' computer or Internet competence – here the three items each for computer literacy and Internet literacy are involved – there are further differences between the sexes. The statement of being able to run and maintain a computer in all important areas (item1) by running a three-way analysis of variance AGE × SEX × SCHOOL produced a main effect on the gender variable. The girls rated their competence in this area significantly lower. This also applies to the statements 3 and 6, that the participants would in the case of computer problems or Internet problems first try to seek a problem solution on their own. Again, the girls estimate their computer and Internet skills in this area significantly lower than the boys. On the other hand, with regard to the willingness to seek external help in the event of computer or Internet problems, both sexes equally have no problem accessing external help (for the results, see **Table 2**).

For the purpose of data reduction, a factor analysis with varimax rotation was performed on the six items. The result was a clear two-factorial structure with a variance resolution of 66.8%. From the four variables loading on the first factor (items 1, 3, 4, 6), the *Maintenance competence scale* was obtained for further statistical analyses ($M = 15.42$, $SD = 3.41$, $N = 133$, $\alpha = 0.76$). The remaining Items 2 and 5, loading on the second factor (indicating the readiness to seek/accept external help) intercorrelate with $r(134) = 0.54$, $p < 0.001$. As in the case of the individual analyses of the items, the result of the three-way analysis of variance AGE × SEX × SCHOOL of the *Maintenance competence scale* shows a significant main effect for the factor SEX, $F(1, 212) = 15.93$, $p = 0.000$, $\eta_p^2 = 0.12$. The girls ($M = 13.95$) rated their computer and Internet maintenance skills significantly lower than the boys ($M = 16.67$). On the other hand, for the case of being in need of expert help due to computer or Internet problems (items 2 and 5), all participants will see no or little problems in finding and accepting help (computer problems: $M = 4.3$, $SD = 1.09$, $N = 134$, Internet problems: $M = 4.1$, $SD = 1.10$, $N = 134$).

Put Briefly

These first results show that the participants differ significantly, be it in terms of their ICT user biography – the girls were given access much later, especially the academic high school girls – be it in terms of their self-assessment of ICT-related skills, preferences and deficits (i.e., computer affinity, ICT maintenance competence, math competence). In all three areas the male participants rated themselves more competent than the female participants. When it comes to computer affinity, it is noticeable that the female academic high school students even achieve the lowest scores, while the male academic high school students achieve the highest scores.

Informal Learning Environment "Family"

Trying to have a closer look at family socialization processes, we asked our participants what they believe their parents think about their computer competence and their suitability for an IT job. This was explored with six statements on parental assessment of their skills and abilities on the computer. The first three questions deal with the perceived importance parents attribute to computer literacy among the participants, and the second group of three questions deals with the skills that their parents deem their youth in the ICT field.

While the mean values for items 1, 2, 4 and 5 are between $M = 3.7$ and almost $M = 4.0$ (**Table 3**), the mean for Item 3 "*My parents encourage me to learn more about computers*" is at a much lower value of $M = 2.3$, i.e., the participants believe their parents are not really interested to influence them to improve their computer skills. However, they still feel supported by them: Statement 5 "*My parents think I'm smart enough to improve my knowledge of computers*" achieves the highest mean, regardless of age, gender and school type.

In statement 4 "*My parents are convinced that I'm good at the computer*," which most clearly expresses how the participants perceive their parents to assess their computer competence, there is a main effect for the factor SEX, $F(1,130) = 4.28$, $p = 0.040$, $\eta_p^2 = 0.032$ within the framework of the two-way analysis of variance SEX × SCHOOL. The girls ($M = 3.70$, $SD = 1.01$) compared to the boys ($M = 4.04$, $SD = 0.84$) significantly assume a more critical competence assessment by their parents.

When directly asked, *how their parents would assess their suitability for an education in computer science*, the two-way analysis of variance SEX × SCHOOL finds main effects for both factors as well as a significant interaction: Again there is a main effect for the factor SEX $F(1, 130) = 10,21$, $p = 0.002$, $\eta_p^2 = 0.073$, i.e., the girls ($M = 2.56$, $SD = 1.37$) expected their parents to assess their suitability for an education in ICT significantly lower than the boys ($M = 3.30$, $SD = 1.30$). The factor SCHOOL, $F(1, 130) = 9.35$, $p = 0.003$, $\eta_p^2 = 0.067$ is also statistically significant: the middle school students ($M = 3.31$) expect their parents to assess their suitability for an education in computer science significantly higher than the academic high school students ($M = 2.70$). Additional information on this can be found in the interaction SEX × SCHOOL, $F(1, 130) = 4.24$, $p = 0.041$, $\eta_p^2 = 0.032$: Here the female academic high school students again reach the lowest value of all ($M = 2.03$), i.e., they assume that their parents tend to deny their suitability for an education in computer science!

To further clarify the ICT socialization of the participants, it is also important to find out who has guided them in terms of computer and Internet use as a child. The answers to the question below about the persons in the participants' social environment (family, school), who were important for teaching her/him ICT skills, provide interesting differences between young women and young men: When asked, "*By whom did you learn the most about computers, the Internet and digital media?*" the Chi^2 test shows statistically significant differences between female and male respondents (Chi^2 = 24.85, $df = 3$, $p = 0.000$, $\phi_c = 0.40$)! **Table 4** shows that as a matter of course many male participants were introduced to the handling of computers by the father and by the friends, while the female respondents are referred to a very high proportion for the acquisition of computer literacy to teachers.

TABLE 3 | Perceived parental attitudes.

	M	*SD*
"Is it true that. . .?"		
(1) My parents believe that it is important for my future that I'm well versed in computers.	3.69	1.12
(2) My parents think that having computer skills is good for me.	3.77	1.05
(3) My parents encourage me to learn more about computers.	2.31	1.20
(4) My parents are convinced that I'm good at the computer.	3.89	0.93
(5) My parents think I'm smart enough to improve my knowledge of computers.	3.97	0.93
(6) My parents think I'm suitable for an education in computer science.	2.96	1.38

Scale: 1 = strongly disagree; 5 = strongly agree.

Put Briefly

These results confirm that the familial ICT socialization of female and male participants in our study has been different. From their parents, the girls expect significantly less compared to the boys that they classify them on the computer as gifted. They themselves doubt their talent for this field. The female academic high school students here again reach the lowest value. Fewer girls than boys were mentored by the father in terms of computer and Internet usage. With the girls, the peers also play only a minor role. All the more, formal education has an important compensatory role to play. In fact, teachers play an important role in ICT socialization of girls. But are they doing justice to this task?

ICT Socialization in School

Looking at the formal learning environment "school," there are several construction sites. To find out what the school's ICT socialization is all about, and what this means for a potential career choice of the students in ICT, the analysis must be in content and personnel. Pedagogically, the legitimate question arises as to how far the school ICT curriculum meets the requirements of a digitized society. Socially, the question arises to what extent the schools are able to compensate for deficits in informal digital socialization of the students (Toomey Zimmermann and Bell, 2012). With regard to central actors, the

TABLE 4 | Persons in the participants' social environment (family, school) important for teaching them ICT skills.

	Males	Females
Father	46.6%	29.5%
Mother	–	8.2%
Peers	42.5%	24.6%
Teachers	11.0%	37.7%

question arises as to what extent todays teachers are competent to train children and adolescents for a digitized society.

Already in the flashback to the elementary school time the participants' information point to substantial deficits with equipment and curricula. The difference between boys and girls in the participation in ICT-relevant lessons is striking: 41.1% of the boys, but only 26.2% of the girls ($Chi^2 = 3.26$, $df = 1$, $p = 0.07$, $\phi_c = 0.16$), affirmed to have participated in PC courses in primary school. The difference is especially pronounced among academic high school students. Of these, 41.9% of the boys, but only 18.2% of the girls, participated in such courses as early as in primary school ($Chi^2 = 4.58$, $df = 1$, $p = 0.028$, $\phi_c = 0.25$).

When asked about the frequency of laptop/notebook classes in school lessons at the middle school or at the academic high school, the participants in our study explained to 73.1% that they had (so far) never participated in such classes. Asked about the use of interactive whiteboards (IWBs) in school lessons (scale: 1 = never to 5 = regularly), 55.2% of respondents said that IWBs were never used. Only 9% of respondents stated that they are used regularly. A significant main effect was found in the two-way analysis of variance SEX × SCHOOL for the factor SCHOOL, $F(1, 130) = 19.83$, $p = 0.000$, $\eta_p^2 = 0.132$. While in middle schools IWBs are used more often ($M = 2.64$), they are rarely used in academic high schools ($M = 1.61$).

This difference in the use of digital media between middle schools and academic high schools is systematic. Still, as the author herself knows from her many years of cooperation with secondary schools in Germany, the faculty members in academic high schools are more likely to assume that their students "do not need" digital media to support school learning. The question of whether they had already participated in special computer/ICT/computer science courses in their school (middle school or academic high school) was confirmed by 94.8% of all participants, 5.2% said no. The course content was mostly office packages, Internet security and Internet research. 13.8% of middle school students and 28% of academic high school students reported having participated in programming courses ($Chi^2 = 3.87$, $df = 1$, $p = 0.049$, $\phi_c = 0.17$). This difference is also statistically significant.

Next, we asked how the participants experienced the computer courses/computer science courses at school. Nine statements on the quality of these courses were presented (later summarized to the *Qualities of the ICT courses scale*). Only the statements 2, 8 and 9 are formulated negatively and include criticism. The responses, analyzed using two-way analysis of variance (SEX × SCHOOL), show that participants appreciated courses that allowed them to run their own projects, courses that related to everyday life, that were dealing with multimedia and web design, and those courses that allowed collaboration and discussions with the peers. For such courses, moderate to medium approval ratings between $M = 2.9$ and $M = 3.4$ were achieved. The female participants more frequently stated that they value collaboration with the peers. The male respondents describe the courses more frequently as boring, dull and repeating content (see **Table 5**). To prepare further analyses, the nine items were analyzed by factor analysis with varimax rotation. The result was a clear two-factor solution with 50.6% explained variance. The items 1, 3, 4, 5, 6, 7, (-) 8 loaded on the factor *Courses A* "*exciting, everyday-oriented, team-oriented content*" with a homogeneity of $\alpha = 0.79$. On the factor *Courses B* "*course contents too narrowly designed (programming, technical skills)*" loaded the Items 2 and 9, which intercorrelate slightly with $r(133) = 0.20$, $p = 0.023$. The scale *Courses A*, based on the participants' very low average values, was identified as suitable for later regression analyses.

TABLE 5 | Opinions about the school computer courses/computer science courses.

	M	*SD*	Two-way analyses of variance Sex × School
"How is it/was it in the school courses (computer courses/computer science courses)? What are/what were your experiences?"			
(1) What I like/always liked in these courses is/was that I can create my own things.	3.08	1.23	n.s.
(2) The topic of programming takes too much space in such courses.	2.30	1.19	n.s.
(3) The tasks in these courses have a connection to my everyday life.	2.86	1.21	n.s.
(4) In such courses, I can use my creativity and imagination.	2.74	1.24	n.s.
(5) I've always found topics like multimedia and web design exciting!	3.39	1.19	n.s.
(6) What I like about these courses is that I can work and discuss with my classmates.	3.21	1.23	Sex: $F(1, 129) = 4.79$, $p = 0.030$, $\eta_p^2 = 0.036$, $M^{boys} = 3.00$, $M^{girls} = 3.47$
(7) The tasks in these courses/in computer science correspond to my interests.	2.71	1.23	n.s.
(8) What we do/have done in the computer/computer science course is usually boring and consists of banal, repetitive tasks.	2.84	1.26	Sex: $F(1, 129) = 7.48$, $p = 0.007$, $\eta_p^2 = 0.055$, $M^{boys} = 3.08$, $M^{girls} = 2.55$
(9) What we do/have done in the computer/computer science course is often aimed at learning high technical skills, far from issues that are significant and creative.	2.50	1.03	n.s.

n.s.: no significant differences between the groups; scale: 1 = strongly disagree; 5 = strongly agree.

As for the ICT competence of the teachers, a total of nine items were presented to the respondents. Overall, the participants rated four negative and four positive statements about the ICT competence of ICT teachers. A critical statement on the ICT equipment of teachers (Item 4) was added. The critical items 1–5 were rated by the respondents with relatively low values between $M = 2.2$ and $M = 2.4$ (1 = strongly disagree to 5 = strongly agree). The four positive statements (items 6–9) had higher values between $M = 3.3$ and $M = 3.6$. Only Item 8 "*My ICT* *teachers encourage and inspire me*" falls well short of these values with $M = 2.8$ (see **Table 6**).

Overall, participants have a moderately positive opinion on the competence and teaching performance of their ICT teachers. In comparison to the male students, the female students were significantly less critical in the negative statements on the ICT competence of the teachers (Items 3, 4). In the positive statements on competence/teaching behavior (6, 7, 8, 9), they rated their teachers significantly more positive. With regard to the critical statement on the technical competence of the teachers, both sexes agreed with a low positive value of $M = 2.37$. For all items with a significant result on the main factor SCHOOL, the academic high school students were significantly more critical than the middle school students.

For later evaluations, these nine items were also analyzed by factor analysis and the *Perceived ICT competence of ICT teachers scale* was formed. Principal component analysis revealed a one-factorial solution with 62.7% variance ($\alpha = 0.92$). As with the individual items, the overall scale confirmed the statistically significant results for the two main factors SEX, $F(1, 130) = 12.30$, $p = 0.001$, $\eta_p^2 = 0.086$, boys = 24.84, girls = 19.93 and SCHOOL, $F(1, 130) = 8.53$, $p = 0.004$, $\eta_p^2 = 0.062$, $M^{MS} = 20.22$, $M^{AHS} = 24.42$. The girls judge their teachers significantly less critically than the boys. The academic high school students assess them significantly more critical than the middle school students.

Put Briefly

It should be noted that the majority of participants in this study experienced less interesting, less inspiring school-based courses on ICT topics. In the classroom, digital media are rarely used, again striking the difference between middle school students (higher use of digital media) and academic high school students. The female participants value those ICT courses significantly higher, which enable cooperation with the peers. They evaluate their ICT teachers significantly more positively than the male participants. This is conclusive, since these teachers are important partners for them in acquiring necessary ICT skills. However, the female participants do not rate the content of the ICT courses more positively! And consistent with the male participants, they evaluated the teachers' technical skills relatively low.

TABLE 6 | Perceived ICT competence of ICT teachers.

	M	*SD*	Two-way analyses of variance Sex × School
"My teachers in this field . . ."			
(1) usually have/had little experience	2.43	1.33	n.s.
(2) are/were technically poorly versed	2.37	1.15	School: $F(1, 130) = 4.54$, $p = 0.035$, $\eta_p^2 = 0.034$, $M^{MS} = 2.14$, $M^{AHS} = 2.55$
(3) do not know enough or less than me	2.26	1.37	Sex $F(1, 130) = 22.97$, $p = 0.000$, $\eta_p^2 = 0.15$, $M^{boys} = 2.74$, $M^{girls} = 1.69$
(4) are/were not properly equipped	2.29	1.19	Sex: $F(1, 130) = 13.30$, $p = 0.000$, $\eta_p^2 = 0.093$, $M^{boys} = 2.62$, $M^{girls} = 1.90$; School: $F(1, 130) = 8.07$, $p = 0.005$, $\eta_p^2 = 0.058$, $M^{MS} = 1.97$, $M^{AHS} = 2.54$
(5) cannot help	2.19	2.24	School: $F(1, 130) = 3.76$, $p = 0.055$, $\eta_p^2 = 0.028$, $M^{MS} = 1.95$, $M^{AHS} = 2.28$
(6) are/were very knowledgeable and fully qualified	3.28	1.24	Sex: $F(1, 130) = 8.46$, $p = 0.004$, $\eta_p^2 = 0.061$, $M^{boys} = 3.04$, $M^{girls} = 3.57$; School: $F(1, 130) = 10.03$, $p = 0.002$, $\eta_p^2 = 0.072$, $M^{MS} = 3.62$, $M^{AHS} = 3.03$
(7) can/could give much information	3.55	1.05	Sex: $F(1, 130) = 6{,}12$, $p = 0.015$, $\eta_p^2 = 0.045$, $M^{boys} = 3{,}34$, $M^{girls} = 3{,}80$; School: $F(1, 130) = 5.95$, $p = 0.016$, $\eta_p^2 = 0.044$, $M^{MS} = 3.81$, MGym = 3.36
(8) encourage and inspire me	2.77	1.23	Sex: $F(1, 130) = 7.35$, $p = 0.008$, $\eta_p^2 = 0.062$, $M^{boys} = 2.51$, $M^{girls} = 3.08$; School: $F(1, 130) = 6.06$, $p = 0.015$, $\eta_p^2 = 0.049$, $M^{MS} = 3.07$, $M^{AHS} = 2.54$
(9) know their way around	3.34	1.17	Sex: $F(1, 130) = 11.81$, $p = 0.001$, $\eta_p^2 = 0.083$, $M^{boys} = 3.04$, $M^{girls} = 3.70$; School: $F(1, 130) = 4.48$, $p = 0.036$, $\eta_p^2 = 0.033$, $M^{MS} = 3.59$, $M^{AHS} = 3.16$

n.s.: no significant differences between the groups; scale: 1 = strongly disagree; 5 = strongly agree.

Interest in Careers in ICT

In order to find out how interested the participants are in careers in the ICT sector, all participants were asked about their interest in ICT professions, once before the intervention (information condition vs. robotics condition) and once afterward. The 17-item *Interest in ICT professions scale* contains statements that should provide insight into the personal assessment of the costs and benefits of education and professional practice in ICT. As mentioned earlier, the concept of the scale incorporates findings from research, but statements from discussions with ICT experts in research and practice have also been incorporated. Following the survey using the *Interest in ICT professions scale*, a short survey (three items) on gender stereotypes were administered (*Gender*$_{tot}$ *scale*; see Section "Gender Stereotypes and Their Importance for the Interest in ICT Professions").

As can be seen from **Table 7**, there were higher approval ratings for items 1, 3, 5, 7, 8, 9, 12 and 15, which also take up many new developments in the occupational field of ICT. The dimensionality of the *Interest in ICT professions scale* was clarified by means of a factor analysis with varimax rotation: A three-factorial solution was obtained, which showed 46.4% explained variance. The dimension "Classically modern ICT profession" includes the items 1, 3, 5, 7, 8, 9, 14,15, summarizing the profession's positive characteristics as diverse fields of application, creative possibilities, good pay, good promotion prospects, the potential for change, and the chance to work independently ($\alpha = 0.71$). The dimension "Flexibility, diversity,

TABLE 7 | Interest in ICT professions scale.

	M	*SD*
"I find interesting in ICT professions . . . "		
(1) . . . that they offer good opportunities to become active in many areas (e.g., media informatics, environmental informatics, medical informatics, bioinformatics, business informatics).	3.60	1.10
(2) . . .that mathematics is not a central element of education	3.04	1.26
(3) . . .that they are creative.	3.60	1.00
(4) . . .that they are not difficult professions.	2.72	1.13
(5) . . .that they are respected and well paid.	3.82	0.93
(6) . . .that previous knowledge in programming is not a prerequisite for education.	3.10	1.12
(7) . . .that they are occupations that are constantly evolving.	4.19	0.82
(8) . . .that they are occupations that will always be important for society.	4.06	0.84
(9) . . .that they are occupations with a variety of tasks.	3.90	0.84
(10) . . .that they are occupations where you have to adapt to many different people with many different ICT-needs.	3.37	0.93
(11) . . .that they are occupations in which I can reconcile work and family well.	3.25	0.95
(12) . . .that I can use many different abilities.	3.61	0.99
(13) . . .that I am expected to be good with people.	3.25	1.09
(14) . . .that I also work alone a lot.	3.39	1.14
(15) . . .that I have good promotion prospects.	3.88	0.85
(16) . . .that these professions help solve many social problems.	3.50	1.01
(17) . . . that skills in mathematics are not important for the practice of the profession.	2.87	1.40

time1 = 1st measurement; total score time1: M = 59.11, SD = 7.96; scale: 1 = strongly disagree; 5 = strongly agree.

social competence" includes items 10, 11, 12, 13,16, the content of dealing with the need for high social competence, the need for many different skills, the good compatibility of work and family ($\alpha = 0.67$). The dimension "Previous knowledge in mathematics and programming are not a prerequisite" includes items 2, 4, 6, 17 ($\alpha = 0.66$).

Differences between groups became only apparent in the dimension "Previous knowledge in mathematics and programming are not a prerequisite." The three-way analysis of variance AGE × SEX × SCHOOL has a significant effect on the AGE factor, i.e., the optimistic belief that missing math and programming skills are no barrier to the different ICT educational paths decreases with age ($N = 134$, AGE: $M^{14years} = 12.53$, $M^{16years} = 11.83$, $M^{18years} = 10.81$; $F(2, 122) = 3.67$, $p = 0.028$, $\eta_p^2 = 0.070$). The 17-item comprehensive *Interest in ICT professions scale* was also subjected to a three-way analysis of variance AGE × SEX × SCHOOL. Once again the factor AGE proves to be statistically significant $F(2, 122) = 6.65$, $p = 0.002$, $\eta_p^2 = 0.098$, i.e., also the interest in ICT professions decreases in general with increasing age ($N = 134$, AGE: $M^{14years} = 60.04$, $M^{16years} = 61.78$, $M^{18years} = 55.91$). It is noteworthy that the factors SEX and SCHOOL have no significant results, i.e., that the interviewed girls and boys as well as the students/graduates of the two school types do not differ in their interest in the ICT professions. After all, it is encouraging and should not be overlooked that this interest reaches on average medium high values of 59 and 60 out of 85 possible points in the first and second measurements!

The three subscales measuring different dimensions of the interest in ICT professions have acceptable homogeneity values for research purposes. As a criterion for regression analysis, however, the 17-item comprehensive *Interest in ICT professions scale* ($M = 59.11$, $SD = 7.96$, $\alpha = 0.78$) is used. On second use after intervention, the scale proved to be stable ($M = 60.5$, $SD = 8.31$, $\alpha = 0.81$). In terms of statistics and content, the *Interest in ICT professions scale* is therefore well suited to clarify which aspects of the media biography of the participants best predicts this ICT interest on a correlative basis.

Examination of Hypotheses by Multiple Linear Regression Analyses

In order to further determine the importance of personal media competence, the informal learning environment "family" as well as the formal learning environment "school" for the interest in careers in the ICT sector, the previously developed indexes and scales were subjected to an inspection of the intercorrelations. It turned out that only few scales are suitable as predictors for the criterion scale *Interest in ICT professions*.

The scales, that have statistical significant correlations with the scale *Interest in ICT professions*, are: the *Computer affinity scale* (*hobby*$_{tot}$), $r(134) = 0.28$, $p = 0.001$, the *Maintenance competence scale*, $r(133) = 0.42$, $p = 0.000$, which measures the computer and Internet competence skills, the *Perceived parental attitudes scale, Item 4 (Parents4 scale)*, $r(134) = 0.39$, $p = 0.000$, which captures how the participants perceive their parents judgment on their computer competence, and the scale *Quality of the ICT courses*, $r(133) = 0.34$, $p = 0.000$, which summarizes the participants' evaluation of the school computer or computer science courses. The scales on the self-assessed mathematics competence, on the perceived ICT competence of the ICT teachers, and surprisingly also the gender stereotypes' scale (to be explained in the next section) do not show statistically significant correlations to the criterion.

Next, standard and stepwise versions of linear multiple regression analyses were carried out and led to identical results. Here the results of the standard versions are reported. The linear multiple regression analysis was performed for the whole group ($N = 134$) and also separately for the female and male participants. It showed that the scale *CoursesA* with its previously documented very low participants' ratings, calculated for the whole group did not contribute significantly to the regression result. In the subsequent analyses of the subgroups, it was completely eliminated from the final regression equation. In a next step, only the scales *Computer affinity/hobby*$_{tot}$, *Maintenance competence* and *Parents4* were included in the regression analysis. The result of the multiple linear regression analysis indicates that there is a collective significant effect between the three predictors. For the overall group of participants, the prediction of the *Interest in ICT professions* (criterion) with these three predictors yields a good variance explanation of *27%* (R^2 = 0.27, R^2adj = 0.25, $F(3, 129) = 15.62$, $p = 0.000$). It turns out that the three predictors *Maintenance competence* (ß = 0.27, $p = 0.0024$), *Parents4* (ß = 0.25, $p = 0.0037$) and *Computer affinity/ hobby*$_{tot}$ (ß = 0.21, $p = 0.0059$) almost equally contribute to the prediction of the criterion.

When performing separate regression analyses for the sexes, the young womens' interest in ICT professions is better predicted than that of the young men by the three predictors: For the young women, the three predictors yield a high explained variance of *45%* (R^2 = 0.45, R^2adj = 0.42, $F(3, 57) = 15.58$, $p = 0.000$ with the scales *Maintenance competence* (ß = 0.37, $p = 0.0036$), *Computer affinity /hobby*$_{tot}$ (ß = 0.29, $p = 0.0074$) and *Parents4* (ß = 0.24, $p = 0.042$). Thus, as hypothesis 1 assumes, these three variables explain the interest of young women in computer science careers. Not only the self-assessment of one's own computer and Internet competence, which is based on personal learning history, plays a decisive role here. Rather, the enthusiasm for everything that has to do with computers plays an important role. Added to this, the belief of the parents in the daughter's computer science talent is a very important variable. All three variables predict the interest of female participants in computer science professions.

For the young men, the predictors yield a variance explanation of *20%* (R^2 = 0.20, R^2adj = 0.16, $F(3, 68) = 5.60$, $p = 0.002$), whereby their interest in careers in IT can be significantly predicted with the two predictors *Maintenance competence* (ß = 0.26, $p = 0.02$) and *Parents4* (ß = 0.25, $p = 0.03$) only. The predictor *Computer affinity/hobby*$_{tot}$ is not significant in the standard version of the multiple linear regression analysis, and in the stepwise version this predictor is omitted in step 2 of the analysis. This is an interesting result, calling for further psychological clarification of the background of the male participants' interest in ICT professions in future research.

Put Briefly

The results of the regression analyses and the preceding statistical comparisons lead to the conclusion that the participants' self-assessed computer and Internet literacy and the parents' belief in their childrens' computer science talent decisively determines their interest in IT professions. For the female participants, their enthusiasm for everything that has to do with computers plays an important role. All skills and motivational variables relevant to develop an interest in a career in the ICT sector are presently based on support measures that are predominantly offered or not offered in the private sphere. The results confirm hypotheses 1 and 2.

Effect of Short-Term Interventions on the Interest in ICT Professions

We asked ourselves: Can a short-term vocational counseling intervention – here called information intervention vs. robotics intervention – increase the participants' interest in ICT professions? Which attitude patterns are activated by these two interventions? After having completed either the information intervention or the robotics intervention, the participants for the second time answered the *Interest in ICT professions scale* as well as, also again, the three items of the gender stereotype scale. In paired t-tests, the total group of participants as well as the gender-separated subgroups were analyzed separately for the two interventions. We looked for significant changes, item per item and for the scale as a whole.

Calculated separately for the intervention groups, the participants in the information condition ($N = 107$) showed a slightly higher, statistically significant, interest in ICT occupations ($M = 11.59$ to $M = 12.57$, $t = -3.57$, $df = 106$, $p < 0.001$, $d = 0.33$), i.e., the positive rating of ICT occupations was increased by the intervention. For the small group of participants of the robotics intervention, the values decreased slightly after the intervention, but not statistically significant ($M^{time1} = 12.22$, $M^{time2} = 11.89$). This result should be treated with caution given the small sample size of this intervention. Nevertheless revealing, if only purely exploratory, is the pattern of attitude changes to ICT occupations in both interventions.

The results show that the male participants of the information condition have benefited from the enlightening content. They prove with significant higher values in the second measurement that they can change their attitudes in the short term. The intervention increased their interest in key areas. After the intervention, they were more convinced that ICT occupations are not insurmountably difficult, that they do not need any programming background as a prerequisite for education and that the ability to deal with people is necessary. However, less than before, they were convinced that computer science degrees lead to many different jobs. With the girls of the information condition, no statistically significant changes caused by the intervention can be detected.

The boys of the robotics condition rated the variety of own abilities that can be applied after the intervention higher than before, while the girls rated the role of social skills as less important after the intervention than before. All in all, although only a small sample was recorded in the case of the robotics intervention, the constructive processing of new positive information about the training and the profession of IT specialists seems to be more difficult for the female participants in both intervention conditions than for the male participants. There is even a slight tendency for the girls to call in question previously identified positive professional characteristics (see **Table 8**, robotics intervention).

Put Briefly

Although both interventions led to a change in individual attitude dimensions, only the information intervention increased the interest in the ICT sector statistically significantly. Informative and targeted advisory interventions can thus change the interest in ICT professions in the short term as stated in Hypothesis 3. But this result should be treated with caution, because primarily the male participants benefited from this intervention. They constructively processed the new information on the ICT domain. Also, the long-term after-effects and the sustainability of our interventions could not be recorded in the absence of follow-up opportunities. Nevertheless, the results of these short-term interventions give important indications that we have to decide carefully about what

TABLE 8 | Attitude change after interventions[a].

	M^{time1} M^{time2}	Paired *t*-test results
"I find interesting in ICT professions. . ."		
1. Information condition: males (*N* = 57)		
(4) that they are not difficult professions. ↑	M^{time1} = 2.68, M^{time2} = 3.18	$t = -3.77$, $df = 56$, $p < 0.001$, $d = 0.45$
(6) that previous knowledge in programming is not a requirement for education. ↑	M^{time1} = 3.05, M^{time2} = 3.39	$t = -2.10$, $df = 56$, $p < 0.05$, $d = 0.27$
(9) that they are occupations with a variety of tasks. ↓	M^{time1} = 4.04, M^{time2} = 3.68	$t = 3.35$, $df = 56$, $p < 0.01$, $d = 0.41$
(13) . . .that I am expected to be good with people. ↑	M^{time1} = 3.04, M^{time2} = 3.44	$t = -3.30$, $df = 56$, $p < 0.01$, $d = 0.40$
2. Information condition: females (*N* = 50)		no statistically significant attitude changes
3. Robotics condition: males (*N* = 16)		
(12) that I can use many different abilities. ↑	M^{time1} = 3.44, M^{time2} = 3.94	$t = -2.24$, $df = 15$, $p < 0.05$, $d = 0.50$
4. Robotics condition: females (*N* = 11)		
(13) that I am expected to be good with people. ↓	M^{time1} = 3.36, M^{time2} = 3.00	$t = 2.39$, $df = 10$, $p < 0.05$, $d = 0.60$

[a]*Statistically significant change items only; N^{Inf} = 107 (participants information condition), N^{Rob} = 27 (participants robotics condition); scale: 1 = strongly disagree; 5 = strongly agree; The arrows (↑↓) signal the direction of change.*

content is taught to whom on the subject of information and communication technologies.

Gender Stereotypes and Their Importance for the Interest in ICT Professions

The question how gender stereotypes influence the interest in ICT training is a focus of current studies (Galdi et al., 2017; Ihme and Senkbeil, 2017; Förtsch et al., 2018; Lloyd et al., 2018). In this study, the extent to which common gender stereotypes are still effective was assessed with three questions before and after the intervention. The three items and the overall scale formed from these items ($Gender_{tot}$/time 1, α = 0.80; follow-up: $Gender_{tot}$/time2, α = 0.78) were tested with a three-way analysis of variance, using the already known factors SEX and SCHOOL and the new variable Learned_F(learned from the father), which records whether the participants have learned their knowledge about computers and digital media primarily from the father or from other people. After all, almost half of the male participants and almost a third of the female participants have their computer knowledge acquired from the father.

The results of the first variance analysis listed in **Table 9** show that the boys who acquired their computer knowledge primarily

TABLE 9 | Gender Stereotypes in ICT.

	M	*SD*	Three-ways analyses of variance Sex × School × Learned_F(time1)
"My opinion on ICT professions is . . . "			
(1) Computer science suits men better than women.	2.68	1.44	School: $F(1,126) = 3.06$, $p = 0.039$, $\eta_p^2 = 0.024$, $M^{MS} = 2.36$, $M^{AHS} = 2.92$ Sex × Learned_F: $F(1, 126) = 5.72$, $p = 0.018$, $\eta_p^2 = 0.043$; boys who have learned from the father achieved the highest value ($M = 3.18$), girls who have learned from the father had the lowest value ($M = 2.17$)
(2) Men have more chances to succeed in the IT/ICT sector than women.	2.84	1.36	School: $F(1,126) = 5.58$, $p = 0.020$, $\eta_p^2 = 0.042$, $M^{MS} = 2.45$, $M^{AHS} = 3.14$
(3) By nature men are more IT/ICT oriented than women.	2.84	1.34	Sex: $F(1,126) = 4.68$, $p = 0.032$, $\eta_p^2 = 0.036$, boys = 3.07, girls = 2.57
$GENDER_{tot}$ Scale/time2	8.37	3.50	Sex: $F(1,126) = 5.56$, $p = 0.020$, $\eta_p^2 = 0.042$, boys = 8.90, girls = 7.72 School: $F(1,126) = 4.36$, $p = 0.039$, $\eta_p^2 = 0.033$, $M^{MS} = 7.41$, $M^{AHS} = 9.09$ Sex × Learned_F: $F(1, 126) = 3.65$, $p = 0.058$, $\eta_p^2 = 0.027$; boys who have learned from the father achieved the highest value ($M = 9.59$), girls who learned from the father achieved the lowest value ($M = 7.17$)

M^{MS} = means of middle school students, M^{AHS} = means of academic high school students; $Gender_{tot}$ Scale = total scale on Gender Stereotypes in ICT (time1 = 1st measurement; α = 0.80, time2 = 2nd measurement; α = 0.78); scale: 1= strongly disagree; 5= strongly agree.

from the father quite strongly internalized the gender stereotype "*Computer science suits men better than women.*" By contrast, the girls, who learned the most about computers, the Internet and digital media by their father, are the least convinced of this gender stereotype. Also, all female participants in the study are significantly less convinced than the male participants of the statement, "*By nature men are more IT/ICT oriented than women.*" These are signals for change.

An unexpected result is: The gender scales in this study, tested for the three single items and the overall *Gender*$_{tot}$ *scale*, show no significant correlations (values $< r = 0.10$) with the *Interest in ICT professions scale*. Also, with the other scales, they are just as low or uncorrelated!

However, as already seen in previous results, academic high school students are more "conservative" than middle school students: in our sample girls and boys who attend academic high school or graduated from academic high school have internalized common gender stereotypes significantly stronger. The effects of the two vocational counseling interventions on the gender stereotypes was analyzed by a paired *t*-test for the measurement times before and after the intervention, each separated by intervention and gender, at the three single items and the *Gender*$_{tot}$ *scale*. With one exception, gender attitudes have not changed after the interventions. The exception concerns the group of male participants in the information condition. They agreed significantly less on the statement "*Computer science suits men better than women*" after the intervention ($M^{time1} = 2.84$, $M^{time2} = 2.60$ ($t = 2.03$, $p = 0.047$, $d = 0.26$).

It should be noted with these results that the approval of the gender stereotypes, be it on the three single items or on the GENDER$_{tot}$ scale, is in the lower middle range for all participants. Addressed directly to common gender stereotypes, the female and male participants in the study show no "strong" reactions. The description of ICT professions as a typical "male" domain did not meet with the female participants' approval.

Put Briefly

Overall, from these previously described results, it can be concluded that the implicit core of Hypothesis 4 "Interest in ICT is influenced by gender stereotypes" is consistent. The ICT socialization conditions in informal (family) and formal (school) education are gendered and designed to hinder gender-equitable access to ICT education and access to ICT careers for girls and young women. Frequently important players (fathers, teachers) in the lives of girls contribute to this situation. However, interest in the ICT sector guided by personal convictions is also present among girls. It is noteworthy that the early introduction to computers by the father reinforces gender stereotypical attitudes in the boys, while significantly attenuates these attitudes in the case of father-trained girls.

The male participants in this study approach the ICT sector with much more self-confidence. In all ICT competency dimensions, they achieve average to higher scores in their self-assessment. They have internalized gender stereotypes, albeit on the whole rather weak. They constructively develop this masculine domain and turn out to be versatile when it comes to processing new information about the ICT sector. Both developments confirm hypothesis 4, but there are also signs of change. In terms of their interest in ICT professions, the surveyed young women and young men, as shown before, do not differ! *Both sexes* show a medium-high interest in this professional field.

DISCUSSION

All four hypotheses presented in the introduction could be confirmed. The learning history of the participants shows that still today it requires a lot of motivation and initiative on the part of young women, but also of young men, in order to be confident in and strive for future professional activities in the ICT domain. By focusing the investigation on a selected group of well researched variables that might influence young peoples' career choice in ICT it becomes clear that the time is right to move forward and to pursue this topic strictly solution-oriented (see Meelissen and Drent, 2008; Miliszewska and Sztendur, 2010; Murphy et al., 2010; Nagy et al., 2010; Jagacinski, 2013; Heerwegh et al., 2016; Lazarides et al., 2016; Sáinz et al., 2016; Ertl et al., 2017; Ihsen, 2017; Förtsch et al., 2018).

The results of the regression analyses and the preceding statistical comparisons confirm *hypothesis 1* that the dimension of self-assessed computer and Internet literacy decisively determines the interest of participants in ICT professions. This is especially true for young women interested in a career in IT. It also applies to the family environment and the support measures provided by families. In the family context, the girls experience significantly less support in learning ICT skills, both factually and psychologically. Thus, the individual variable "computer affinity" plays an important role for their interest in ICT professions. The late ICT socialization, the low level of guidance in the private sphere and the reliance on compensatory educational offers by the schools put them at a disadvantage. These facts must also be seen in the context of the much lower self-confidence of female participants assessing their own knowledge and skills in the field of ICT. Developing high computer affinity under such conditions is difficult and only achieved by a fraction of them.

Hypothesis 2 is also confirmed in numerous results of statistical group comparisons and in the results of the regression analyses: The educational offerings of the schools, which are hardly anchored in the school curriculum, are too rare, too unsystematic and qualitatively insufficient to provide solid ICT knowledge in order to stimulate an interest in ICT professions. The milder judgment of female participants on the competence of ICT teachers and the quality of their courses does not mean that the teachers' efforts have led to a compensatory achievement, – it signals gratitude. According to the regression analyses, these school offers have no lasting influence on the young womens' interest in ICT professions. Overall, these results show that, considering the low quality of ICT course content and the participants' low ratings of their ICT teachers' ICT competence, todays schools' influence is too weak to motivate male and female students to seek an education and a career in the ITC sector. The task of compensating socialization deficits among girls and young women can't be fulfilled by the schools.

Although both vocational counseling interventions led to a change in individual attitude dimensions, only the information intervention increased the interest in the ICT sector statistically significantly. Informative and targeted advisory interventions can thus change the interest in ICT professions in the short term as stated in *hypothesis 3*. The fact is: in this case, in this investigation, the male participants primarily benefited from this intervention.

It is noteworthy that the early introduction to computers by the father reinforces gender stereotypical attitudes in the boys, while significantly attenuates these attitudes in the case of father-trained girls. It is also noteworthy that *both sexes* show a medium-high interest in this professional field. From the previously described results on gender stereotypes, it can be concluded that the implicit core of *hypothesis 4* "Interest in ICT is influenced by gender stereotypes" is consistent. The ICT socialization conditions in informal (family) and formal (school) education are gendered and designed to hinder gender-equitable access to ICT education and access to ICT careers for girls and young women. Important players (fathers, teachers) in the lives of girls still contribute to this situation. At the same time, they are, at least temporarily, a key to solving these problems.

The Interest in ICT Professions and Gender Stereotypes

The *Interest in ICT professions scale* addresses important aspects of the requirements for a career in the ICT industry. Due to the scale's dimensionality, it is advisable to further develop it based on the structures found. Both sexes showed no statistically significant differences before and after the interventions. Both genders consistently show a medium-high interest in ICT careers. Intervention-related changes indicate that it is important which content is passed on to young people at this age in the context of vocational counseling. For this it is important to clarify what the psychological processes involved in career choice are all about, especially, how professional interests develop. A new approach to the development of interests (in general) could be helpful: O'Keefe et al. (2018) recently published the results of several experiments on peoples' implicit theories of the development of personal interests. The authors distinguish two groups of implicit/commonsense theories they have found in their investigations. The first group includes implicit theories that are relatively fixed ("fixed theory"), i.e., the person assumes that a once found interest/passion must be pursued with the highest motivation (excluding other interests); the resulting problems are often underestimated. The second group includes individuals' implicit theories based on the assumption that personal interests are developing ("growth theory"). O'Keefe and colleagues could show in a series of experiments that people with a fixed theory more quickly lose interest in the pursued passion if problems arise compared to people with an implicit growth theory of interest development (O'Keefe et al., 2018).

A research approach that incorporates these findings and relates them to career choice processes in young women could provide important information. As our research results show one key ingredient in the young women for a positive attitude toward the ICT domain is a long-standing enthusiasm for computers/computer science. This distinguishes them from the male participants in the study, who approach the topic in a more pragmatic way. In particular, the development of effective vocational counseling intervention could benefit from this new research on interest development.

Clayton et al. (2009, p. 153) in a research review on the role of gender stereotypes in ICT came to the conclusion that gender stereotypes "provide misleading ideas about ICT as a career discouraging both girls and boys." In a more recent study published by Banchefsky and Park (2018, p. 1), the authors researching gender-science stereotypes in male-dominated academic disciplines could prove that their female participants were "significantly less likely to endorse the gender-science stereotype." In the introduction to this article the question was asked "*Can the interest in ICT counteract the masculine image of computer science?*" Based on the lessons learned from their projects, the two research teams of Lasen (2010) and Sáinz et al. (2016) agreed that a new understanding of the tasks and professions in the ICT business already show a positive impact on the interest of girls and young women in ICT careers. Not only statistically, also noteworthy in content is the missing correlation of the reliably measuring gender stereotypes scale ($\text{Gender}_{\text{tot}}$) with the scale "Interest in ICT professions." As in the project of Papastergiou (2008) both sexes' approval of the gender stereotypes is in the lower middle range. What role do these gender stereotypes actually play today? These results probably indicate a step in the right direction: Apparently, the female respondents in our study are interested in the ICT area regardless of how they assess this domain, which they nevertheless perceive as gendered. They realistically value the ICT work environment and the problems that await them there, as triggered by domain masculinity. In any case, they separate between their interest in ICT professions and the perceived masculinity of the ICT domain.

Limitations

As the results of the survey show, the participants in the study were motivated to deal with the subject of ICT professions. Their interest in this professional field was well above average. This can be interpreted as a limitation of the relevance of the test results due to the selection of the participants. After all, the *Deutsches Museum* in Munich is a technology museum. By contrast, it can be contended that the museum is visited by families, groups of students and tourists from all over Germany and from abroad, usually as a half or full day trip. It may also be argued that the study identifies shortcomings in the education system in Germany that apply only to a specific region. However, anchoring the study locally in the *Deutsches Museum* had the advantage that students from all over Germany were interviewed. That the study sample is lacking in representativeness is already indicated by the sample size.

Another limitation concerns the vocational counseling intervention part of the study: In the absence of follow-up opportunities, we were unable to detect whether our interventions had long-term effects and how they might have turned out. Our interventions served as a test. It was proven that the information provided during the interventions was processed differently by the female and male participants in the study and, to a limited extent, produced short-term effects. More research in this area is needed to follow up on the effects of the interventions over a longer period of time (e.g., several months).

In general, it is recommended that interventions at the current level of gender and STEM research are used more frequently. They lead to revealing test phases, in this study as well as in the actual study by Wheeler and Blanchard (2019). Beyond the hoped-for effects (e.g., Boaler et al., 2018), interventions provide information about the (possibly changing) validity of already investigated important factors that influence education and professional choice in STEM subjects (Wang et al., 2013). Generational, cultural and age-related topics and preferences come into play too (Han, 2019; Wheeler and Blanchard, 2019). Also, interventions offer the opportunity to reveal one-sidedness in research. In today's view, the assessment of Hsieh et al. (2019) is to be agreed that it should be avoided to study the STEM disciplines as if they were a single subject. In fact, the later professions, which build on the various training paths in the STEM disciplines, have very different profiles.

Surprising Results

The study also showed some surprising results: as the survey involved interviewing participants with two different (desired or already achieved) school qualifications, it seemed appropriate to look at these groups separately. Surprisingly, it turned out that the female academic high school students had the biggest gap in their ICT socialization/training. As far as these students are concerned, it can be assumed that the socialization disadvantages identified in this study affect them strongly. Apparently, they are still educated and trained according to an outdated "humanistic" educational ideal. In this study, they are the ones whose first use of computers took place very late; they critically consider their mathematical knowledge and their ICT-maintenance competence. They give the lowest ratings in the parental assessment of their computer skills. They are also the most critical of parents' attitude to their suitability for an ICT profession. While their male classmates have the highest computer affinity despite the low level of computer science training at school, their average computer affinity is lowest. So it is not surprising that the number of female computer science students is growing very slowly. But there is a ray of hope: measures of informal ICT education (family) – currently mostly mediated by the father – seem to be quite effective in establishing computer skills and computer affinity in girls. In terms of gender stereotypes, an increased involvement of the fathers can have an immunizing effect.

Although the results of this small study can only claim a limited scope of validity, they nevertheless offer many starting points for gender-sensitive access to ICT education at all ages (from primary school pupils to high school graduates). As early as possible, digital media have to be firmly anchored as educational media in the consciousness of the younger generation. Something has to change in the informal, family driven as well as the formal education at school. For years, the universities have enthusiastically organized "Boy's Days" and "Girl's Days" to encourage children to become interested in STEM fields at an early age. In fact, it requires "Parents' Days" for parents of children of all ages! The schools, in this study especially the middle schools and the academic high schools have deficits in their curricula and in relation to the ICT competence of the teaching staff. These deficits must be reduced rapidly in the coming years. Digital education should start early and be firmly anchored in the primary school curriculum in order to counterbalance early deficits in access and knowledge of information and communication technologies. But even in adolescence, the ICT competence of students can still be successfully promoted if the contents are clearly anchored and chosen wisely. Much can be achieved on every school-based training level, if one offers a high-quality, lasting and interesting ICT curriculum in the schools and at the same time does not forget to additionally get the parents on board.

Parents, fathers as well as mothers, play an important role in communicating ICT-relevant skills, and above all in educational digital agenda setting. Almost equally important is the emotional support of parents for the children and adolescents interested in ICT education. The fact that fathers are currently better able to provide gender-appropriate ICT guidance for girls is probably a time-bound result that can be overtaken by future social developments. But that is why it is no less relevant! This is a research result that we share with the authors Gunderson et al. (2012a,b), Galdi et al. (2017), Ihme and Senkbeil (2017), and Lloyd et al. (2018).

Today's students and graduates have a new need for gendered "normality." They learn self-directed, and with the teaching of ICT skills, the peers play an important role in both sexes. Gender stereotypes have statistically little or no relevance to the other ICT-relevant variables in this study, although they are still present in the gendered training biographies of the girls and boys in this study. After all, according to our results, gender-appropriate, target group-oriented, high-quality school education was not available to our male participants too. However, they approach the topic with a higher self-esteem. An important step toward easing and objectifying this imbalance would be to make the subject of computer science in Germany a standard subject in the curriculum, as is the case with the subjects of physics, chemistry and biology.

ETHICS STATEMENT

The study was approved by the *ad hoc* Ethics Committee of Faculty II of the University of Siegen. The commission follows the ethics rules and the rules of procedure for local ethics committees of the German Psychological Society. In the study, 14–18-year-old participants were approached in the context of their visit to the Deutsches Museum in Munich, Germany, to participate and were informed about the purpose and content of the survey. The participants came to the museum with their families. Oral informed consent was obtained from the parents of the participants; those participants who were of age decided on their own to participate. The consent procedure was approved by the ethics committee.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

The publication of this research was funded by the Ministry of Culture and Science of North Rhine-Westphalia (by University of Siegen).

ACKNOWLEDGMENTS

The author would like to thank André Bertels and Aleksandr Gorovoj for their help in carrying out the study and/or helpful comments on the manuscript, and also wishes to express gratitude to the participants for their voluntary participation in the study.

REFERENCES

Banchefsky, S., and Park, B. (2018). Negative gender ideologies and gender-science stereotypes are more pervasive in male-dominated academic disciplines. *Soc. Sci.* 7:27. doi: 10.3390/socsci7020027

Bitkom Press (2018). *Frauenanteil in der ITK-Branche Wächst Langsam*. Available at: https://www.bitkom.org/Presse/Presseinformation/Frauenanteil-in-der-ITK-Branche-waechst-langsam.html (accessed March 7, 2018).

Boaler, J., Dieckmann, J. A., Pérez-Núnez, G., Sun, K. L., and Williams, C. (2018). Changing students minds and achievement in mathematics: the impact of a free online student course. *Front. Educ.* 3:26. doi: 10.3389/feduc.2018.00026

Cheryan, S. (2012). Understanding the paradox in math-related fields: why do some gender gaps remain while others do not? *Sex Roles* 66, 184–190. doi: 10.1007/s11199-011-0060-z

Cheryan, S., Plaut, V. C., Davies, P. G., and Steele, C. M. (2009). Ambient belonging: how stereotypical cues impact gender participation in computer science. *J. Pers. Soc. Psychol.* 97, 1045–1060. doi: 10.1037/a0016239

Clayton, K. L., von Hellens, L. A., and Nielsen, S. H. (2009). "Gender stereotypes prevail in ICT; a research review," in *Proceedings of the SIGMIS-CPR'09*, (New York, NY: ACM), doi: 10.1145/1542130.1542160

Conein, S., and Schwarz, H. (2015). IT-Berufe auf dem Prüfstand. *BWP* 44, 58–59.

Ebach, J. (1994). Der rückgang des frauenanteils in der informatik – überlegungen zu möglichen ursachen aus psychologischer Sicht. *Zeitschrift für Frauenforschung* 12, 16–27.

Ehrke, M., Hageni, K.-H., and Heimann, K. (2011). *Die duale IT-Berufsausbildung in Deutschland im Kontext der Globalisierung*. München: ISF München.

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Förtsch, S., Gärtig-Daugs, A., Buchholz, S., and Schmid, U. (2018). "Keep ICT going, girl!" An empirical analysis of gender differences and inequalities in computer sciences. *Int. J. Gender Sci. Technol.* 10, 265–286.

Galdi, S., Mirisola, A., and Tomasetto, C. (2017). On the relations between parents' and children's implicit and explicit academic gender stereotypes. *Psicologia Sociale* 2, 215–238.

Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012a). New directions for research on the role of parents and teachers in the development of gender-related math attitudes: response to commentaries. *Sex Roles* 66, 191–196. doi: 10.1007/s11199-011-0100-8

Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012b). The role of parents and teachers in the development of gender-related math attitutes. *Sex Roles* 66, 153–166. doi: 10.1007/s11199-011-9996-2

Han, F. (2019). Self-concept and achievement in math among Australian primary students: gender and culture issues. *Front. Psychol.* 10:603. doi: 10.3389/fpsyg.2019.00603

Heerwegh, D., De Wit, K., and Verhoeven, J. C. (2016). Exploring the self-reported ICT skill levels of undergraduate science students. *J. Inform. Technol. Educ.* 14, 19–47.

Hergatt Huffman, A., Whetten, J., and Huffman, W. H. (2013). Using technology in higher education: the influence of gender roles on technology efficacy. *Comput. Hum. Behav.* 29, 1779–1786. doi: 10.1016/j.chb.2013.02.012

Hsieh, T., Liu, Y., and Simpkins, S. D. (2019). Changes in United States Latino.A high School students' science motivational beliefs: within group differences across science subjects, gender, immigrant status, and perceived support. *Front. Psychol.* 10:380. doi: 10.3389/fpsyg.2019.00380

Hunt, C. S., and Hunt, B. (2004). Changing attitudes toward people with disabilities: experimenting with an educational intervention. *J. Manag. Issues* 16, 266–280.

Ihme, J. M., and Senkbeil, M. (2017). Warum können Jugendlichen ihre eigenen computerbezogenen Kompetenzen nicht realistisch einschätzen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 49, 24–37. doi: 10.1026/0049-8637/a000164

Ihsen, S. (2013). Der ingenieurberuf. von der traditionellen monokultur zu aktuellen gender- und diversity-relevanten perspektiven und anforderungen. arbeit: zeitschrift für arbeitsforschung. *Arbeitsgestaltung und Arbeitspolitik* 3, 236–246.

Ihsen, S. (2017). *Dialog MINT-Lehre. Mehr Frauen in MINT-Studiengänge. Handlungsempfehlungen zur Integration von Gender in der MINT-Lehre. Abschlussbericht und Transferkonzept*. Available at: https://www.kompetenzz.de/Unsere-Projekte/A-Z/Dialog-MINT-Lehre (accessed December 18, 2018).

Ihsen, S., Höhle, E. A., and Baldin, D. (2013). *Spurensuche! Entscheidungskriterien für Natur- und Ingenieurwissenschaften und mögliche Ursachen für frühe Studienabbrüche von Frauen und Männern an TU9-Universitäten.TUM Gender- und Diversity-Studies*. Münster: LIT Verlag.

Ihsen, S., Mellies, S., Jeanrenaud, Y., Wentzel, W., Kubes, T., Reutter, M., et al. (2017). *Weiblichen Nachwuchs für MINT-Berufsfelder gewinnen. Bestandsaufnahme und Optimierungspotenziale. TUM Gender- und Diversity-Studies, Bd.* 3. Münster: LIT Verlag.

Ito, T. A., and McPherson, E. (2018). Factors influencing high school students' interest in pSTEM. *Front. Psychol.* 9:1535. doi: 10.3389/fpsyg.2018.01535

Jagacinski, C. M. (2013). Women engineering students: competence perceptions and achievement goals in the freshman engineering course. *Sex Roles* 69, 644–657. doi: 10.1007/s11199-013-0325-9

Kessels, U., and Hannover, B. (2004). Was bewirkt die Werbung für ingenieurwissenschaftliche Fächer? Evaluation eines universitären Angebots für Schülerinnen und Schüler. *Empirische Pädagogik* 18, 228–251.

Kim, H.-S., Kil, H.-J., and Shin, A. (2014). An analysis of variables affecting ICT literacy levels of Korean elementary School students. *Comput. Educ.* 77, 29–38. doi: 10.1016/j.compedu.2014.04.009

Lasen, M. (2010). Education and career pathways in information communication technology. What are Schoolgirls saying? *Comput. Educ.* 54, 1117–1126. doi: 10.1016/j.compedu.2009.10.018

Lazarides, R., and Ittel, A. (2013). Mathematics interest and achievement: what role do perceived parent and teacher support play? A longitudinal analysis. *Int. J. Gender Sci. Technol.* 5, 207–231.

Lazarides, R., Rubach, C., and Ittel, A. (2016). Motivational profiles in mathematics: what roles do gender, age, and parents' valuing of mathematics play? *Int. J. Gender Sci. Technol.* 8, 124–131.

Lazarides, R., and Watt, H. M. G. (2017). Student-perceived mothers' and fathers' beliefs, mathematics and English motivations, and career choices. *J. Res. Adolesc.* 27, 826–841. doi: 10.1111/jora.12317

Lloyd, A., Gore, J., Holmes, K., Smith, M., and Fray, L. (2018). Parental influences on those seeking a career in STEM: the primacy of gender. *Int. J. Gender Sci. Technol.* 10, 308–328.

Logel, C., Walton, G. M., Spencer, S. J., Iserman, E. C., von Hippel, W., and Bell, A. E. (2009). Interacting with sexist men triggers social identity threat among female engineers. *J. Pers. Soc. Psychol.* 96, 1089–1103. doi: 10.1037/a0015703

Maaß, S., and Wiesner, H. (2006). Programmieren, mathe und ein bisschen hardware . . . wen lockt dies bild der informatik? *Informatik Spektrum* 29, 125–132. doi: 10.1007/s00287-006-0059-y

Mansoori-Rostam, S. M., and Tate, C. C. (2017). Peering into the 'black box' of education interventions and attitude change: audience characteristics moderate the effectiveness . . . and then only towards specific targets. *J. Soc. Psychol.* 157, 1–15. doi: 10.1080/00224545.2016.1152211

McLachlan, C., Craig, A., and Coldwell, J. (2010). "Student perceptions of ICT: a gendered analysis," in *Proceedings of the 12th Australasien Computing Education Conference (ACE 2010), Conferences in Research and Practice in Information Technology*, Vol. 103, (Brisbane), 127–136.

Meelissen, M. R. M., and Drent, M. (2008). Gender differences in computer attitudes: does School matter? *Comput. Hum. Behav.* 24, 969–985. doi: 10.1089/g4h.2014.0056

Miliszewska, I., and Sztendur, E. M. (2010). Interest in ICT studies and careers: perspectives of secondary School female students from low socioeconomic backgrounds. *Interdiscip. J. Inform. Knowl. Manag.* 5, 237–260. doi: 10.28945/1162

Murphy, M. C., Steele, C. M., and Gross, J. J. (2010). Signaling threat. How situational cues affect women in math, science, and engineering settings. *Psychol. Sci.* 18, 879–885. doi: 10.1111/j.1467-9280.2007.01995.x

Nagy, G., Watt, H. M. G., Eccles, J. S., Trautwein, U., Lüdtke, O., and Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: different countries, different trajectories? *J. Res. Adolesc.* 20, 482–506. doi: 10.1111/j.1532-7795.2010.00644.x

Nordmann, S. (2016). *MINT-Frauenstudiengänge in Deutschland. Review Created on Behalf of Bitkom*. Available at: https://www.bICTkom.org/sICTes/default/files/file/import/160419-MINT-Frauenstudiengaenge-Uebersicht.pdf (accessed December 18, 2018).

OECD (2015). *ICT Familiarity Questionnaire for PISA 2015 (International Option), Main Survey Version*. Available at: http://www.oecd.org/pisa/data/CY6_QST_MS_ICQ_Final.pdf (accessed December 18, 2018).

O'Keefe, P. A., Dweck, C. S., and Walton, G. M. (2018). Implicit theories of interest: finding your passion or developing it? *Psychol. Sci.* 29, 1653–1664. doi: 10.1177/0956797618780643

Papastergiou, M. (2008). Are computer science and information technology still masculine fields? High School students' perceptions and career choices. *Comput. Educ.* 51, 594–608. doi:10.1016/j.compedu.2007.06.009

Quaiser-Pohl, C. (2012). "Mädchen und Frauen in MINT: Ein Überblick," in *Mädchen und Frauen in MINT: Bedingungen von Geschlechtsunterschieden und Interventionsmöglichkeiten*, eds H. Stöger and M. Heilmann (Münster: LIT Verlag), 19–39.

Régner, I., Selimbegovic, L., Pansu, P., Monteil, J.-M., and Huguet, P. (2016). Different sources of threat on math performance for girls and boys: the role of stereotypic and idiosyncratic knowledge. *Front. Psychol.* 7:637. doi: 10.3389/fpsyg.2016.00637

Reid, C. (2009). Technology-loving luddites? Declining participation in high School computing studies in Australia. *Br. J. Sociol. Educ.* 30, 289–302. doi: 10.1080/01425690902812562

Resch, D., Graf, I., Dreiling, A., and Konrad, J. (2017). *Attraktivität von ICT-Berufen – Synthesebericht. Hochschule für Wirtschaft*. Available at: https://www.fhnw.ch/de/die-fhnw/hochschulen/hsw/pmo/forschung-und-dienstleistung/attraktivitaet-von-ict-berufen (accessed March 7, 2018).

Sáinz, M., Meneses, J., López, B.-S., and Fàbregues, S. (2016). Gender stereotypes and attitudes towards information and communication technology professionals in a sample of Spanish secondary students. *Sex Roles* 74, 154–168. doi: 10.1007/s11199-014-0424-2

Schinzel, B., Kleinn, K., Wegerle, A., and Zimmer, C. (1999). Das studium der informatik: studiensituation von studentinnen und studenten. *Informatik-Spektrum* 22, 13–23. doi: 10.3205/zma000738

Schinzel, B., and Ruiz Ben, E. (2004). Softwareentwicklung als profession? professionalisierungstendenzen und implikationen für die beteiligung von frauen. *Informatik Spektrum* 18, 441–447. doi: 10.1007/s00287-004-0408-7

Schmader, T., Johns, M., and Barquissau, M. (2004). The costs of accepting gender differences: the role of stereotype endorsement in women's experience in the math domain. *Sex Roles* 50, 835–850. doi: 10.1023/b:sers.0000029101.74557.a0

Shiina, A., Niitsu, T., Sato, A., Omiya, S., Nagata, T., Tomoto, A., et al. (2017). Effect of educational intervention on attitudes toward the concept of criminal responsibility. *World J. Psychiatry* 7, 197–206. doi: 10.5498/wjp.v7.i4.197

Thoman, D. B., White, P. H., Yamawaki, N., and Koishi, H. (2008). Variations of gender-math stereotype content affect women's vulnerability to stereotype threat. *Sex Roles* 58, 702–712. doi: 10.1007/s11199-008-9390-x

Toomey Zimmermann, H., and Bell, P. (2012). "Everyday expertise. Learning within and across formal and informal settings," in *Theoretical foundations of learning environments*, eds D. Jonassen and S. Land (New York, NY: Routledge), 224–241.

Turetsky, K. M., and Sanderson, C. A. (2017). Comparing educational interventions: correcting misperceived norms improves college students' mental health attitudes. *J. Appl. Soc. Psychol.* 48, 46–55. doi: 10.1111/jasp.12489

Vekiri, I. (2010). Socioeconomic differences in elementary students' ICT beliefs and out-of-School experiences. *Comput. Educ.* 54, 941–950. doi: 10.1016/j.compedu.2009.09.029

Vekiri, J., and Chronaki, A. (2008). Gender issues in technology use: perceived social support, computer self-efficacy and value beliefs, and computer use beyond School. *Comput. Educ.* 51, 1392–1404. doi: 10.1016/j.compedu.2008.01.003

Verhoeven, J. C., Heerwegh, D., and De Wit, K. (2010). Information and communication technologies in the life of university freshmen: an analysis of change. *Comput. Educ.* 55, 53–66. doi: 10.1016/j.compedu.2009.12.002

Wang, M., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice: individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937

Wendland, M., and Rheinberg, F. (2004). "Welche motivationsfaktoren beeinflussen die mathematikleistung? eine längsschnittanalyse," in *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung*, ed. J. Doll (Münster: Waxmann Verlag), 309–328.

Wheeler, S. R., and Blanchard, M. R. (2019). Contextual choices in online physics problems: promising insights into closing the gender gap. *Front. Psychol.* 10:594.

Wille, E., Gaspard, H., Trautwein, U., Oschatz, K., Scheiter, K., and Nagengast, B. (2018). Gender stereotypes in a children's television program: effects on girls' and boys' stereotype endorsement, math performance, motivational dispositions, and attitudes. *Front. Psychol.* 9:2435. doi: 10.3389/psyg.2018.02435

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 07 May 2019
doi: 10.3389/fpsyg.2019.00996

Young Spanish People's Gendered Representations of People Working in STEM. A Qualitative Study

Milagros Sáinz[1]*, *José-Luis Martínez-Cantos*[2], *María Rodó-de-Zárate*[1], *María José Romano*[1], *Lidia Arroyo*[1] *and Sergi Fàbregues*[1,3]

[1] *Internet Interdisciplinary Institute, Universitat Oberta de Catalunya, Barcelona, Spain,* [2] *Department of Education, Universidad Complutense de Madrid, Madrid, Spain,* [3] *Department of Psychology and Education, Universitat Oberta de Catalunya, Barcelona, Spain*

The present qualitative study analyzes how a group of young people already involved in STEM fields perceive the prototypical person working in STEM. Gender differences between participants in technological and non-technological STEM fields were analyzed. A total of 27 young people (59.3% women) took part in the interviews (Mean Age = 25.48 years). Of them, 16 participants were working in STEM professions, and 11 were enrolled in the final courses of STEM degrees. The results of the content analysis were examined in light of social role theory and the multidimensional structure of gender stereotypes. Men in these fields were therefore attributed an unappealing and weird physical appearance. Some female participants linked STEM professionals' intellectual abilities to the stereotype that men have higher abilities in these fields. Whereas females attributed effort and perseverance to STEM professionals' intellectual aptitudes, males referred to the development of soft skills. Participants in technological STEM fields connected the stereotype of being a 'weirdo' to a boring job, whereas those in non-technological fields linked it to their unconventional character. Some participants were disappointed by a lack of correspondence between expectations and the actual job STEM professionals do. Moreover, females in technological STEM fields commented on the job's low social impact, while males mentioned low attainment of technical qualifications. Most referents in STEM fields were masculine, some of whom were present in the mass media. The practical implications of the findings are discussed.

Keywords: gender stereotypes, role models, portrayals, STEM, under-representation

OPEN ACCESS

Edited by:
M. Gail Jones,
North Carolina State University,
United States

Reviewed by:
Sarah J. Carrier,
North Carolina State University,
United States
Marco Salvati,
Università degli Studi di Roma "La Sapienza", Italy

Correspondence:
Milagros Sáinz
msainzi@uoc.edu

Specialty section:
This article was submitted to Gender, Sex and Sexuality Studies, a section of the journal Frontiers in Psychology

Received: *06 January 2019*
Accepted: *15 April 2019*
Published: *07 May 2019*

Citation:
Sáinz M, Martínez-Cantos J-L, Rodó-de-Zárate M, Romano MJ, Arroyo L and Fàbregues S (2019) Young Spanish People's Gendered Representations of People Working in STEM. A Qualitative Study. Front. Psychol. 10:996. doi: 10.3389/fpsyg.2019.00996

INTRODUCTION

Our society has experienced important advances in terms of equality thanks to the efforts deployed to achieve an egalitarian education among young people. However, we continue observing a marked gender gap in the academic and professional aspirations that young people develop during secondary education (Wang and Degol, 2013; Sáinz and Müller, 2018). In addition, and although they have nearly attained equality with men in several formerly male-dominated fields, women remain underrepresented in several fields of science, technology, engineering and mathematics (STEM), (Wang and Degol, 2013; UNESCO, 2017).

Nowadays, many girls are reluctant to choose STEM disciplines related to engineering, computer science, and physical science. In fact, and according to the Spanish Ministry of Education (MECD, 2018), during the 2016–2017 academic year women represented only 17, 17.39, and 11.83% of the student enrollments in computer science, electrical and energy engineering, and electronics and automation technologies, respectively. However, women are significantly represented in scientific disciplines such as biology, mathematics, and chemistry, accounting for 61.78, 37.66, and 53.20%, respectively, of student matriculation in these university degrees. Above all, women outnumbered men in disciplines related to the provision of healthcare, such as medicine or pharmacy, representing 65.8 and 69.58% of total enrollments in these studies, respectively.

This phenomenon can be also observed in many Western countries. According to data from UNESCO (2017), within the female student population in higher education globally, only around 30% choose STEM studies. As in Spain, differences can be observed by disciplines. Female students' enrollment is particularly low in information and communication technologies (ICT) (3%), natural sciences, mathematics and statistics (5%), and engineering, manufacturing and construction (8%); but the highest participation is in health and welfare (15%) studies (UNESCO, 2017). These data highlight the importance of analyzing gendered representations of people working in technological STEM fields versus those in non-technological STEM fields. Thus, technological STEM fields include people graduate in areas like engineering, computer science, or architecture. These disciplines are mainly oriented to the design of technological appliances and services and in most of the cases women remain remarkably underrepresented (Sáinz and Müller, 2018). In addition, the non-technological STEM fields group comprises people graduate in science disciplines like biology, pharmacy, medicine, or mathematics, where technologies are frequently the tool rather than the object of their work and women are in general highly represented (Sáinz and Müller, 2018).

Gender Stereotypes of People Working in STEM

Gender-role stereotyping of careers might be an important reason why women are staying away from many STEM careers (Wang and Degol, 2013; Sáinz et al., 2016a; Steinke, 2017). According to the multidimensional structure of gender stereotypes (Deaux and Lewis, 1984) and Eagly's (2001) social role theory, gender stereotypes have a multidimensional structure because they comprise features associated with the ideal person working in a particular field (i.e., physical appearance, role behaviors, personality traits, and occupations). People take these characteristics as a reference to make inferences about the ideal man or woman working in different occupations.

According to social role theory, women are thought to behave in a communal fashion—that is, concerned about other people, friendly, and expressive. In contrast, men are thought to behave in an agentic manner—independent, assertive, and instrumental (Eagly, 2001). People in highly male-dominated STEM fields (such as engineering) will therefore be more likely to behave in an agentic way, whereas people in highly female-dominated fields (such as education or nursing) will be more likely to behave in a communal fashion. Thus, people in highly male-dominated STEM fields like engineering will be depicted as having several attributes (such as being weird, possessing high intellectual abilities, developing technical tasks, or earning lots of money) congruent with the masculine agentic gender role rather than with the feminine communal gender role. For the goals and roles congruity theory (Diekman et al., 2010)—a theoretical framework stemming from social role theory—the underrepresentation of women in STEM careers is associated with the perception that STEM careers are less likely than careers in other fields (such as psychology) to fulfill communal goals (e.g., working with or helping other people). Consistent with this theory, these perceptions might disproportionately affect young women's career decisions in many STEM fields, because women are more likely to endorse communal goals than men (Diekman et al., 2010).

In this regard, research on young people's portrayal of a typical person working in STEM shows that, when asked to draw a scientist using the Drawing a Scientist Test (DAST) (Chambers, 1983) or describe STEM professionals, adolescents tend most often to depict these professionals as male, as well as unattractive, white, middle-aged or elderly, dressed in a lab coat, wearing glasses, 'geeky,' or 'nerdy,' socially awkward, and being people who work alone (Barker and Aspray, 2006; Steinke et al., 2007; Cheryan et al., 2013; Sáinz et al., 2016b). Girls were more likely than boys to report the counter gender-stereotyped perception of scientists and STEM professionals as female (Steinke et al., 2007; Sáinz et al., 2016b). In addition, recent studies have corroborated the assumption that women are more likely than men to be underrepresented in many STEM fields because women are stereotyped as being less likely to possess a sort of 'raw' talent than men (Meyer et al., 2015). Most of these portrayals are related to people working in scientific or technological fields, such as physical scientists or engineers (Steinke et al., 2007; Sáinz et al., 2016b). Moreover, several studies show that male-dominated jobs such as engineering and other technology-related occupations are associated with a high status and well-paying stereotype (Eagly, 2001; Sáinz et al., 2016b). That is, young people's portrayals of STEM professionals include different features that make reference to the person's physical appearance and other several gender role behaviors (Eagly, 2001; Sáinz et al., 2016b). These portrayals are important to examine because they shape young people's interest in pursuing STEM courses and occupations (Steinke et al., 2007). For this reason, in the present research we aim to study the gendered representations that a group of young people (some of them already in STEM) have about a typical person working in various STEM fields beyond engineering and physical science. Given the disparity of women's representation across STEM fields, differences in gender and discipline will also be analyzed.

Schools, families, and popular media such as TV series and Hollywood movies play a crucial role in the construction, representation, reproduction, and transmission of stereotypes of these STEM professionals (Steinke, 2017). In this regard, gender stereotypes of STEM professionals in the media influence students' stereotyped perceptions of STEM (Steinke et al., 2007; Steinke, 2017). This information is particularly salient and relevant for girls and boys during and beyond adolescence, as young people actively consider their future personal and professional identities not only before selecting any concrete field, but also after having selected it and deciding to develop professionally in that area. Incidentally, it can be expected that the stronger the correspondence between young people's self-portrait and the archetypal person working in a given field—for instance a science teacher—the more likely the young person is to choose this field (Kessels and Taconis, 2012).

Gender -Marking Language to Express Stereotypical Portrayals of People in STEM

The use of masculine, feminine, and neutral gender marks provides researchers with interesting additional information about the way young people depict the typical image of someone working in the different STEM fields (Gabriel and Gygax, 2016). As demonstrated in other studies, the use of the masculine generic in Indo-European-origin languages with grammatical gender, such as Spanish, Catalan, and French, denotes a high underrepresentation and undervaluing of women in many STEM fields (Gabriel and Gygax, 2016; Sáinz et al., 2016b). The existence of semantic gender markers in Spanish and Catalan (languages used in the context of the present research) activates gender categories and the perpetuation of differing expectations for men and women. It also reinforces existing gender stereotypes (Gabriel and Gygax, 2016). For instance, the use of the feminine singular enfermera or infermera to refer to a female nurse in Spanish and Catalan, or the usage of the masculine singular to refer to a doctor in both languages as médico or metge, or the generic plural masculine to refer to different professions such as engineers, physical scientists, or scientists either in Spanish—ingenieros, físicos, científicos— or in Catalan—enginyers, físics, científics—. These gender markers are not only limited to nouns, but also apply to pronouns and adjectives. In a recent research study conducted in the context of Spain, secondary students associated more masculine than non-masculine references to a person working in a highly male-dominated field such as information and communication technologies (Sáinz et al., 2016b). These workers were associated with masculine characteristics through the use of adjectives and other markers.

The Present Study

There is a lack of research focused on gendered portrayals of people working in STEM fields with high numbers of women, mainly with a non-purely technological orientation such as biology, biomedicine, or chemistry. In the present research we thereby simultaneously examine the opinion that a group of participants belonging to highly male-dominated STEM fields (with a high technological component, such as engineering or computer science) and highly female-dominated STEM fields (with a less technological orientation) have about the typical person working in the STEM field. Similarly, there is no research about the image that young people already in STEM hold regarding the typical person working in STEM. In this regard, most research looking at young Spanish people's portrayals of professionals has been conducted with secondary students (Sáinz et al., 2016a,b). In addition, most of the research on these aspects has been conducted via surveys and using various mixed methods. Qualitative research delving into the type of stereotypical gender role portrayals of people already in STEM fields is scarce. For this reason, the present study applies a novel qualitative approach to examine young people's gendered representations of people working or studying in different STEM fields. Through this research, we therefore attempt to cover the aforementioned research gaps.

The research questions and hypotheses were therefore formulated as follows:

(RQ1) What are the main features that participants highlight as portraying the typical person working in—technologically and non-technologically oriented—STEM fields?

H1: Participants are expected to report more masculine than feminine characteristics (i.e., physical appearance, intellectual abilities, personality traits, or social position) when describing the prototypical person working in technologically, and non-technologically–oriented STEM fields. (RQ2) To what extent do male participants differ from female participants in their portrayals of the prototypical person working in STEM?

H2: More male than female participants are expected to provide masculine features when portraying the prototypical person working in STEM.

(RQ3) To what extent do participants studying or working in technological STEM fields express their portrayal of the prototypical person working in STEM in the same terms as participants in non-technological STEM fields?

H3: More participants in technological STEM fields than in non-technological STEM fields are expected to use masculine features when describing the prototypical person working in the field.

(RQ4) To what extent do male and female participants from technological STEM fields express their portrayal of the ideal of the person working in STEM in similar terms as male and female participants from non-technological STEM fields?

H4: In comparison to male participants in non-technological STEM fields, male participants in technological STEM fields will use more masculine features to portray the person working in STEM. The same would be true for female participants in technological STEM fields, in comparison to female participants in non-technological STEM fields.

MATERIALS AND METHODS

Study Design

A qualitative descriptive design (Sandelowski, 2010) was used to generate an accurate and in-depth account of how young people already in STEM perceive the prototypical person working in this field. This type of design is especially suited to research situations where researchers want to use a low level of interpretation of the events studied. In contrast to more interpretative qualitative approaches such as grounded theory, phenomenology, or ethnography, in which "a conceptual or otherwise highly abstract rendering of the data" (Sandelowski, 2010, p. 335) is required, in qualitative descriptive studies researchers stay close to the data by presenting the facts in the everyday language of the participants. Therefore, this approach enabled us to ensure descriptive validity (Maxwell, 1992), that is, to gain an accurate understanding of participants' thoughts and beliefs, expressed in their own words and, as a result, minimize researcher bias. Furthermore, the adoption of a qualitative descriptive approach was consistent with the primary goal of describing and understanding the subjective nature of the perceptions conveyed by the participants.

Sample

Purposive sampling was used to select 11 students in the second or higher year of their bachelor's degree and 16 STEM professionals employed in the private sector for 1–5 years.

We aimed for heterogeneity in both groups in terms of gender and type of degree program. Potential participants were identified using formal and informal strategies, including the following: (a) Asking acquaintances if they knew of any potential participants; (b) contacting student associations, professors, and companies in the STEM field; and (c) snowballing from previous contacts. We continued to interview until data saturation was achieved, that is, new data generated no further insights. Saturation was assessed by analyzing the interview transcripts. Consequently, 27 participants were included in the study. The sample size was consistent with recommendations suggested in the literature (Kuzel, 1999; Guest et al., 2006). Before being interviewed, participants were individually screened by telephone or email to ensure eligibility criteria were met.

The characteristics of the 27 study participants are displayed in **Table 1**. Participants included 11 males and 16 females, either finishing the last course of a STEM university degree (five males and six females) or working in a STEM field with a maximum of 5 years' experience in private companies (6 males and 10 females). Participants were living in the metropolitan areas of Barcelona (11 students and 8 professionals) and Madrid (8 professionals). The mean age of participants was 22.6 (SD = 1.4) for students and 27.4 (SD = 2.9) for professionals. The students were enrolled on degree courses in physical sciences (n = 3), computer science engineering (n = 2), telecommunications engineering (n = 2), mathematics (n = 1), medicine (n = 1), pharmacy (n = 1), and physics engineering (n = 1), whereas the professionals had completed degrees in industrial engineering (n = 3), architecture (n = 2), biology (n = 2), pharmacy (n = 2), physical sciences (n = 2), telecommunications engineering (n = 2), aeronautical engineering (n = 1), mathematics (n = 1), and mining engineering (n = 1). All the participants were born in Spain.

Data Collection

Semi-structured interviews were conducted from April to September 2016. All interviews took place in Spanish in locations chosen by the participants, such as campuses, workplaces, and coffee shops. The interviews lasted from 40 to 90 min and were conducted by four members of the research team. The majority of the interviews were conducted in Spanish, but six of them were held in Catalan. Follow-up prompts were used to allow interviewees to expand on their answers. Before the interviews, we obtained informed consent and authorization to record the responses.

The interview guide, based on the research questions and a review of the literature, included the following three questions: (1) Why did the participants decide on a STEM degree? (2) How would they characterize a prototypical STEM professional? (3) What do they consider to be the significant barriers to and facilitators of women's access to the STEM field? Each question had the same weight and allotted time in the interview. However, only the findings related to the second question are reported and discussed in this article.

Data Analysis

Interviews were transcribed verbatim, imported into QSR NVivo software, and analyzed using qualitative content analysis (Schreier, 2012). This method allowed us to focus on categories of interest in the interview data set and systematically analyze them in a flexible way. Qualitative content analysis was implemented in three phases. First, the coding scheme was drafted based on Deaux and Lewis (1984) theory of the multidimensional structure of gender stereotypes. The coding scheme included various codes, including the following: intellectual aptitudes, personality traits, social position, and role models associated with the various STEM disciplines. In addition, we added the code of Spanish gender-marked terminology referring to professionals in these disciplines (e.g., *enfermera*, *médico*, *ingeniero*, and *arquitecto*). Second, to test the coding scheme, two researchers applied it to the same 30% of the data using NVivo. Results were compared and the researchers discussed those cases in which the same segments of text were assigned different codes. The coding comparison ensured that the two coders interpreted the codes similarly and facilitated evaluation of the consistency and validity of the coding scheme. Disagreements were discussed and arbitrated by a third member, when necessary. A few changes were made as a result of this test. These included merging similar codes and eliminating those that were found to be irrelevant. In the third phase, we applied the coding scheme to the interviews. After all the data were coded, NVivo matrix coding query was performed in order to compare responses with the characteristics of the interviewees (i.e., gender, STEM field, bachelor's degree).

TABLE 1 | Characteristics of the 27 study participants.

	Student (*n* = 11)	Professional (*n* = 16)	Total (*n* = 27)
Mean age (*SD*)	22.64(1.50)	27.44(2.99)	25.48(3.43)
Gender, *n* (%)			
Male	5(45.5)	6(37.5)	11(40.7)
Female	6(54.5)	10(62.5)	16(59.3)
Place of residence, *n* (%)			
Barcelona	11(100)	8(50)	19(70.4)
Madrid	0(0)	8(50)	8(29.6)
STEM field, *n* (%)			
Technological	5(45.5)	9(56.3)	14(51.9)
Non-technological	6(54.5)	7(43.8)	13(48.1)
Bachelor's degree, *n* (%)			
Technological STEM fields			
Aeronautical Engineering	0(0)	1(6.25)	1(3.7)
Architecture	0(0)	2(12.5)	2(7.4)
Telecommunications Engineering	2(18.2)	2(12.5)	4(14.8)
Computer Science Engineering	2(18.2)	0	2(7.4)
Industrial Engineering	0(0)	3(18.75)	3(11.1)
Physics Engineering	1(9.1)	0	1(3.7)
Mining Engineering	0(0)	1(6.25)	1(3.7)
Non-technological STEM fields			
Medicine	1(9.1)	0	1(3.7)
Pharmacy	1(9.1)	2(12.5)	3(11.1)
Physical Sciences	3(27.3)	2(12.5)	5(18.5)
Mathematics	1(9.1)	1(6.25)	2(7.4)
Biology	0(0)	2(12.5)	2(7.4)
Mean years since degree completion (*SD*)	–	4.06(2.35)	–
Mean years of work experience (*SD*)	–	3.94(1.48)	–

RESULTS

Physical Appearance

Several instances regarding the physical appearance of people working in the different STEM fields were identified. However, whereas the prototypical image associated with most people in these STEM fields had a positive formal look (with descriptions such as 'a person with glasses and wearing a white coat'), computer and physical scientists were mainly associated by some participants with a 'weird' and sometimes negative unattractive physical image ('untidy,' 'careless,' 'with uncombed hair,' or 'pale skin'). Consistently, most of these prototypical people were explicitly associated with men. The use of gender marks (masculine nouns, attributes, or complements) was evidence of this masculine portrayal. The next description of a male computer science student exemplifies that masculine portrayal.

> With dark clothes, a bit heavy metal-looking (masculine). A bit pale (masculine), spending all the time confined to a room under a florescent light without daylight, [. . .] the typical freak spending hours on the computer with a bag of Cheetos by his side. (Participant 1)

No differences were observed among the participants with technological and non-technological STEM backgrounds. In addition, both male and female participants expressed similar stereotypes about the physical appearance of STEM professionals. However, a gender bias emerged since the stereotype about the unkempt appearance of professionals in highly male-dominated fields was exclusively related to masculinity. In general, both male and female participants considered that women take much more care of their physical appearance, which according to them could discourage women from entering fields where physical appearance is not important or which involve dealing with raw materials and wearing coveralls. The following testimony of a female pharmacist working in a lab refers to those aspects.

> In the production department everything was very dirty, you handle lots of materials, raw material [. . .] it's more for boys. (Participant 2)

In this regard, for a female engineer the notion that only 'intelligent and ugly women enter these masculine fields' discouraged many young girls from entering these professions. She also explained how at university she changed her physical appearance (abandoning the use of make-up and high heels) in order to adapt to the way her female engineering university colleagues dressed and looked.

> I entered university wearing high heels. I used to wear make-up, but none of my classmates did. Then I started to wear low-heeled shoes, dress more casually, comfortably, [. . .] engineering [. . .] is something very *macho*, it's like people considered that only the

ugliest and brightest women entered the field [. . .], which could be a reason for women not studying engineering, you get your hands dirty. (Participant 3)

In fact, a couple of female scientists believed that they did not fit into the stereotype of someone who does not take care of their physical appearance. In general terms, the results are in line with previous studies (Cheryan et al., 2013; Sáinz et al., 2016b) and show participants' use of features regarding the physical appearance of STEM professionals when referring to their prototypical image.

Intellectual Aptitudes

In general, different attributes associated with intelligence were reported by many participants when describing the prototypical person working in STEM fields. No differences between participants from technological and non-technological STEM fields were identified. However, differences between male and female participants were observed. On the one hand, male participants (like the engineer quoted below) linked intelligence to technical, spatial, mathematical, and/or physical science abilities.

I imagined a man with good mathematical or physical science abilities, with a lot of technical knowledge and good spatial abilities. (Participant 4)

Some male participants with work experience (like the next engineer) also underlined the need to possess certain soft-skills (being open-minded or having good managerial and business skills) or good personal qualities (especially if working with clients) as a complement to technical skills.

Everyday good managers are in demand. It is not only a matter of being technically qualified. You also have to understand the fiscal and economic implications of your work. (Participant 5)

On the other hand, some female participants (such as the following telecommunications engineering student) considered that hard work, perseverance, and effort were basic dimensions of the intelligence associated with STEM professionals.

It is obvious that having the ability is essential, but effort is also important. (Participant 6)

Likewise, some female engineers remarked that having high intellectual abilities did not mean being educated; it could be connected with STEM professionals' lack of social skills. The next female telecommunications engineer suggests a lack of cultural knowledge among engineers, despite their high intellectual abilities.

They are intelligent people, who know to compute a partial derivative in 20 s, but maybe they don't know what the capital city of Kuwait is. (Participant 7)

Moreover, some female participants like the following graduate in biomedicine talked about further aspects of intelligence related to STEM professionals' personality traits (e.g., being methodical, capable of resolving problems, rigorous, or highly creative). That is, people with flexible intellectual aptitudes, strong analytical skills, and logical reasoning.

With an analytical vision—not narrow-minded, but analytical, objective—of how to plan things with sound logical reasoning; but this does not exclude a more intuitive side. (Participant 8)

Finally, it is important to note that a couple of female participants highlighted a relationship between the stereotype of STEM professionals' high intellectual abilities and sexism in the field. For them, intelligence tends mainly to be considered a masculine characteristic. Another female engineer commented on the common assumption that women have less technological abilities than men, and for this reason, women were supposed to stand out because of their good communication and organizational skills.

I don't know why it is supposed that women have less knowledge about technologies. In my field, when you are doing an interview, unconsciously, they think that you have fewer abilities. Maybe they expect you to make up for that gap in your technological abilities with other qualities such as being more organized, getting on well with other people, having more fluid communication skills. (Participant 7)

A female interviewee in the life sciences explained that she had held the prejudice that men were the best and most outstanding scientists.

At the research level, I had the mindset that women also did research, but I always believed a certain [. . .] cliché that men were better. (Participant 9)

As observed, and in line with the theoretical background (Deaux and Lewis, 1984; Eagly, 2001), many students referred to several aspects of the intellectual aptitudes associated with people working in STEM.

Personality Traits

Participants alluded to aspects related to STEM professionals' character or personality traits. No gender differences emerged regarding their view of STEM professionals' personality traits. In this sense, both male and female participants belonging to technological STEM settings (a total of 14 interviewees) characterized professionals in these fields in terms of being 'freaks' or 'weirdos.' In the same fashion, the term *freak* was used by some participants in the fields of math and physical science to describe people working in these fields. However, both male and female participants mainly referred to engineers, computer scientists, or physical scientists as people (normally men) lacking communication skills (i.e., a grumpy male, confined to his room, or a person lacking empathy). Remarkably, one male participant even alluded to physical scientists as male heterosexuals.

Moreover, computer and physical scientists were described by participants from both genders as males with a clear focus on activities that could be boring for other people (i.e., obsessed, lunatic, or a 'bookworm'). Interestingly, many of these participants (like the following male computer engineering student) also referred to STEM professionals as lacking team-building abilities.

A bookworm, a grumpy male, a person confined to his/her room. As the machine does not allow human interaction [. . .]. Little empathy. (Participant 10)

In addition, whereas participants in technological STEM fields (like the male aeronautical engineer referred in the next first quote) described those working in these fields as people who found it difficult to establish social relations ('grumpy') or had an analytical mind or little empathy, participants in non-technological STEM fields (like the male physical science student mentioned in the following second quote) portrayed people working in scientific STEM fields as 'lively,' 'amusing,' 'spontaneous,' 'extroverted,' and also 'weird,' but in terms of being independent and unconventional.

> Yes, with difficult personal relations. (. . .) like Sheldon Cooper, who thinks that his work is more important than what others do; a man or woman who is passionate about what he/she is studying or working on (. . .) obsessed about this. (Participant 11)
>
> I imagined physical scientists a bit like mathematicians, stereotypically more spontaneous. They do not follow social conventions. (Participant 12)

Remarkably, in comparison to the two participants from the field of architecture (who portrayed architects as being bohemian or artists), a participant in the field of engineering portrayed engineers as being serious, entrepreneurial, and practical because they were supposed to get straight to the point. On the other hand, participants in STEM health-related fields (mostly women such as the pharmacy student mentioned in the next quote) commented on these STEM professionals' kindness and predisposition to help people.

> A very serious person, [. . .], kind. A person who can give you a hand, an honest male, very upright. (Participant 13)

Likewise, and aligned with their vision of the intellectual aptitudes associated with STEM professionals, female participants (like the biomedicine worker cited in the following quote) placed greater emphasis on personality traits related to dedication, perseverance, and seriousness. Conversely, male participants were more focused on describing STEM professionals as independent people who do not follow social conventions.

> A well-considered person, a hard-working person, with intellectual capacity, a serious male. Strict and dogmatic people. Willing to work many hours without being paid. (Participant 9)

Moreover, most male and female participants believed that the stereotype regarding STEM professionals' lack of communication skills fitted more with male rather than female examples, given that women were supposed to be more communicative and empathetic. However, some female participants complained about how the stereotype of women's poor technical competences led them to assume tasks congruent with this stereotype. That is, to join teams to develop social and communication skills rather than technical skills. In the following testimony, a male engineering student attempts to dismantle the stereotype that women have more communicative skills than men.

> Women, [. . .] empathize more easily given the work that they have unconsciously achieved, but [. . .] I have also seen disastrous presentations given by girls; girls who do not know how to communicate. (Participant 14)

Some participants (as illustrated in the testimony of the next male physical science student) also acknowledged that the image they had of a person working in the field had changed after having entered into contact with real people either at university or work.

> Physical scientists are not actually like I imagined; they are a bit crazy and extroverted, but a bit serious and in this regard a bit different from what I expected. (Participant 1)

All these testimonies inform us about the importance of considering personality traits when tackling the portrayal of a typical person working in STEM.

Social Position

Explicit reference was made to aspects related to the status or social position of people working in STEM fields. In this regard, some male and female participants stated that STEM jobs were generally well considered because of the associated prestige (or social importance, as represented in the own words of a male architect in the next first quote), salary, and respect (these two aspects are commented by a male engineer in the following second quote), or the content of the tasks to be developed. Equally, there is the belief that people do not fully understand the type of work carried out by many of these STEM professionals.

> With high social importance. (Participant 15)
>
> It is well-paid [. . .] you suffer for some years, but then people outside began to respect me. (Participant 3)

However, the two male architects participating in the study complained about the low salary and lack of stability associated with jobs in the field, mainly in comparison to the years prior to the economic crash (that took place in Spain between years 2008 and 2011 with negative effects particularly on the real-state sector) and to other participants in the fields of engineering and computer science who did not mention any of these aspects.

> My salary is basically my main barrier because I consider that with my background I should have a higher salary. I have a technical degree, architecture, I speak four languages, I have an international career, I have done international projects and I believe that people in Spain do not value this. (Participant 16)

On the other hand, women in non-technological fields (mainly in health) felt that their jobs were very well considered, as long as they were associated with respect and admiration. Nevertheless, in comparison to technological STEM fields, participants in biology-related jobs like the next woman complained of the low pay and poor labor conditions, in spite of being well considered socially.

> On a labor level, a low-paid person. [. . .] Socially speaking, quite the opposite, that is, a well-considered person, in terms of being hard-working, and with intellectual capabilities. (Participant 9)

Some participants also mentioned that a pharmacist's (male) status was lower than that of a medical doctor (male). Likewise as defended by the following female medicine student, a medical doctor (male) had more 'knowledge' than a nurse (female). Interestingly, the singular masculine was associated with a person working as a medical doctor (*médico* or *metge*) or a pharmacist (*farmacéutico* or *farmacèutic*), whereas the feminine singular was used to refer to a person working as nurse (*enfermera* or *infermera*).

> The knowledge acquired by a medical doctor (masculine) is greater than that of a female nurse [. . .] The female nurse has a lot of contact with the patient [. . .], but for making decisions and knowing why things happen, there are things that the (male) medical doctor knows that the (female) nurse has not studied. (Participant 17)

Finally, some participants in the field of health also mentioned a change in the current condition of the pharmacy profession, given that in the past it used to be associated with 'very wealthy' people, who had the resources to open a pharmacy. The next female pharmacy student also commented that medical doctors have a higher social consideration than pharmacists.

> I cannot see differences between a medical doctor and a pharmacist (both in masculine), but the medical doctor has higher status than the pharmacist. (Participant 13)

In general terms, the social standing of people working in the different STEM professions was part of the stereotypical view of the different STEM fields.

Type of Tasks STEM Professionals Do

This category refers to participants' views of the tasks carried out by people working in the different STEM fields. Most participants in the fields of engineering and physical science believed that, at the time of choosing their university degree, they did not know much about the actual activities and tasks performed by professionals working in each particular field. In fact, for those with work experience, the actual job had little to do with the previous image they had had. In this regard, some of these participants (like the male aeronautical engineer referred in the next quote) reported that their expectations revolved around a professional more dedicated to the manufacturing process (including aspects such as design, calculus, or analysis), and less involved in the performance of managerial or business-related tasks.

> In 2009 or 2010 the view that I had of engineering was related to design, manufacturing, calculus, analysis, and maintenance [. . .]. I thought of the tasks that I like to do as an engineer [. . .]. But I know that aeronautical engineering is more than that. (Participant 11)

On the other hand, few male participants such as the following mining engineer admitted to having an erroneous previous image of the tasks carried out by engineering professionals, given that they had an idealization of these professionals.

> Sometimes when I say that I am a mining engineer, people associate me with being underground, in the dark, breathing dust. (Participant 18)

Interestingly, two engineering students (one male and one female) reported being deeply disappointed with the gap between the expectations they had about the concrete tasks carried out in each STEM field and what they actually experienced at university. In this regard, a gender difference also emerged, since the female student felt she had been 'deceived' because of the lack of social impact of the professional activity (i.e., doing something that changes humanity).

> I imagined some genuine work, to investigate something that changes humanity, but then you say: 'but this, they've conned me.' (Participant 6)

The male student was disappointed because he expected to have attained more technical knowledge (e.g., understanding computer-related processes).

> I got upset [. . .] I used to believe that computer science was [. . .] people who knew how to handle a computer, understood it and could work with it. (Participant 18)

However, some female participants in health occupations like the following pharmacy student also indicated a certain degree of knowledge about the tasks to be performed in these occupations.

> I imagined what I wanted to become, what I have done in my master's degree, that is, developing new products in different settings, chemistry, cosmetics, food. (Participant 2)

In line with expectations, most participants referred to the specificities of the tasks that are performed in the different STEM fields when thinking about someone working in the field.

Role Models

This category involves the concrete references (significant people like family members, secondary teachers, neighbors, or characters in TV series or movies) that participants mentioned with regard to people working in STEM fields. Both women and men agreed on the idea that people working in the fields of engineering and physical science were predominantly men. The historically higher visibility of men in scientific and technological fields in the media or other public spaces has contributed to this underrepresentation of women. However, some female references such as participants' mothers or female teachers were mentioned by the interviewees from life science disciplines such as biology, pharmacy, or medicine. Moreover, female referents were thought to take care of their physical appearance, as well as being more responsible, empathetic, predisposed to help others, and with more social skills than men. Male referents, however, were associated with high intellectual and research capabilities, along with greater physical force, and roughness.

Similarly, most male participants in the fields of engineering, architecture, and physical science more frequently spoke about pre-eminent male role models such as Stephen Hawking, Albert Einstein, or Richard Feynman in the field of physical science. Equally, in the field of computer science, several outstanding male figures (such as Mark Zuckerberg or Edward Snowden) were mentioned, as well as some characters in TV series and movies (such as *Mr. Robot*, or Sheldon Cooper from *The Big Bang Theory*), cartoons, or science books. The following testimony of a male computer engineering student revolved around those famous media characters.

> When I was I child I used to watch cartoons where inventions, inventors making machines appeared [. . .] I admired the creator of Facebook [. . .] I also paid attention to Snowden [. . .] Right now

> I am following a TV series about people in Silicon Valley [. . .] *Halt and Catch Fire* [. . .] *Mr. Robot* [. . .] is about security and hackers. (Participant 1)

But also they referred to close male family members and role models such as uncles, parents, teachers, university professors, company leaders, etc. The next quote of one male engineer illustrates this idea.

> I had a female professor teaching on engines, she taught very well [. . .] one of the managers in my company [. . .] he works very well. (Participant 3)

Similarly, among most female engineers and physical scientists, male role models also predominated. However, and in line with the next female engineer's statement, these male figures were the participants' fathers or secondary school teachers.

> My father had an electronics and telecommunications company [. . .] since I was a child [. . .] I wanted to become like my father when I grew up. (Participant 19)

Interestingly, female interviewees rarely mentioned famous figures as their role models. Only one engineer referred to Leonardo Da Vinci as a male role model and to Marie Curie as a female role model. In addition, one male architect complained about having not had any idea during his university training about Zaha Hadid, one of the most inspiring modern architects.

> Zaha Hadid, the most famous female architect in the last years, who died recently [. . .] I did not know anything about her, practically until I left university. Participant 13)

In this regard, only a few female engineers mentioned close female role models (such as cousins) working in STEM fields who had inspired them. The next female telecommunications engineer talked about her cousin as her closest female role model.

> It was a female cousin [. . .]. When I was a child everybody said 'Ah, Maria Jesús is doing an engineering degree, it is very difficult [. . .]' Everybody admired her, [. . .] I somehow aspired to be like her. (Participant 3)

Furthermore, and unsurprisingly, female participants in life science disciplines predominantly identified female referents (i.e., relatives, friends, or teachers) as their role models or mentors. They did not mention experts or characters present in the media.

DISCUSSION

In general terms and congruent with the multidimensional structure of gender stereotypes (Deaux and Lewis, 1984; Eagly, 2001), the results of the present study confirm the co-existence of various stereotypical attributes about people working in different STEM fields, both technologically and non-technologically oriented. These features reinforce traditional masculine views of people working in STEM, but they also reflect current changing roles of women in the STEM workplace and our society. Some of these stereotypical portrayals are very much associated with gender roles more or less congruent with STEM fields (Diekman et al., 2010). In this regard and in line with expectations (Cheryan et al., 2013; Sáinz et al., 2016b), several marked gender stereotypical portrayals of people working in male-dominated STEM fields such as engineering and computer science were observed. However, in light with expectations more male than female participants provided masculine features when portraying the prototypical person working in STEM. In addition, some intergroup differences with regard to participants' gender and belonging to technological and non-technological STEM fields have been identified. This is one of the novel contributions of the present research.

As regards the features selected to describe professionals belonging to different STEM occupations, the association of many of these professionals with their physical appearance is evident. These characteristics are aligned with other studies conducted in light of social role theory (Cheryan et al., 2013; Sáinz et al., 2016b). Consistent with expectations (Steinke et al., 2007; Cheryan et al., 2013; Sáinz et al., 2016b), both male and female participants agreed on labeling the stereotypical image of people working in STEM through the use of masculine physical features typically associated with these professionals. In some STEM professions (e.g., physical science or computer science) these physical features were negative and unattractive (uncombed hair, careless, ugly, working in the dark) and opposed to the typical portrayal associated with women, who according to some participants are more likely to take care of their physical appearance. Interestingly, no differences were observed in how participants from technological and non-technological STEM occupations referred to masculine attributes associated with the physical appearance of people working in STEM. For some female participants in technological STEM fields, the lack of possibilities for women to develop their feminine identity in highly masculine STEM occupations such as engineering may discourage them from entering these fields.

In addition and congruent with other studies (Cheryan et al., 2013; Meyer et al., 2015; Sáinz et al., 2016a), attributes related to the possession of intellectual abilities were a major feature mentioned by many participants. No differences between people from technological and non-technological fields emerged. However, for some female participants, these intellectual abilities were associated with other skills related to personality traits that revolve around effort, hard work, or perseverance. But for some male participants, such intellectual abilities were mainly associated with technical capabilities (e.g., mathematical or spatial thinking) or the development of complementary soft skills (e.g., good managerial skills). These findings confirm the hypothesis that STEM fields are above all associated with a high level of raw intellectual capabilities (Meyer et al., 2015). This could also be interpreted as a recognition of communal goals taking an important role in the development of intellectual skills across the different STEM fields (Eagly, 2001; Diekman et al., 2010).

In relation to STEM professionals' personality traits and in agreement with predictions (Cheryan et al., 2013; Sáinz et al., 2016b), some negative characteristics (such as being antisocial, 'freaks,' and only interested in machines) were attributed mainly to people working as scientists, engineers, or computer scientists; that is, highly male-dominated occupations. However, with the exception of physical scientists and mathematicians, other non-technological STEM professionals (that is, highly female-dominated occupations) were less likely to be associated with weird personality characteristics. In line with expectations, more male than female participants provided more masculine features related to personality traits when portraying the prototypical person working in STEM. In addition, being a 'weirdo' or a 'freak' meant something different for professionals belonging to technological and non-technological STEM fields. Whereas 'being a weirdo' in non-technological STEM fields was identified with having a peculiar, independent character, in technological STEM fields it was identified with lacking social skills and performing boring tasks. This result illustrates how according to predictions, more participants in technological STEM fields referred to masculine features when depicting the prototypical person working in the field. Interestingly, whereas male participants in physical science and engineering fully identified with the prototypical negative weird image of these professionals, female participants in these male-dominated STEM fields did not identify with it. In agreement with empirical research (Cheryan et al., 2013; Steinke, 2017), this was considered by some female participants as a factor that detracts many girls from entering in these fields. These findings also suggest that in comparison to male participants in non-technological STEM fields, male participants in technological STEM fields use more masculine features related to physical appearance when portraying the person working in STEM. The same is true for female participants in technological STEM fields, in comparison to female participants in non-technological STEM fields.

Participants from different STEM fields highlighted some stereotypical requirements demanded by the specific field they belong to. For instance, possessing a bohemian character was associated with architecture, whilst being a practical and serious person was part of engineering's portrayal. Moreover, health professionals were linked to a high predisposition to help others. Interestingly, and with regard to gender differences, some female participants in male-dominated STEM fields such as engineering assumed the stereotype that women are more socially, but less technically, skillful. In addition and in line with predictions, for many female participants, the stereotype regarding STEM professionals' lack of social skills fitted more with males than with females (Cheryan et al., 2013; Sáinz et al., 2016b). For this reason, female participants did not identify with the stereotypical portrayal of STEM professionals as lacking social skills. In fact, female participants tended to highlight the link between intelligence and personality traits, such as perseverance, dedication, or seriousness. Interestingly, women already in STEM contradict the stereotype regarding people in this field as possessing a high level of raw intellectual ability (Meyer et al., 2015). These women highlight the need to possess complementary attitudes, such as effort or interest in developing the supposed technical intellectual mindset.

With regard to the prestige of STEM occupations, most of the participants coincided in highlighting that STEM degrees were prestigious because of the highly competitive entry prerequisites, difficulty, the social contribution they make, or the high salary associated with some engineering programs (Sáinz et al., 2016b). In harmony with expectations, many of these prestige-related features are congruent with the attainment of masculine agentic goals (Eagly, 2001; Diekman et al., 2010; Sáinz et al., 2016b). However, interesting differences were observed across the various STEM fields. Some architects and biologists complained that these occupations were significant in terms of social recognition, but lacked a decent salary and job stability. Within the health professions, different reputational features emerged, for instance between a medical doctor (in masculine terms) and a nurse (in feminine terms), or a medical doctor (in masculine terms) and a pharmacist (in masculine terms). These findings suggest a certain devaluation of STEM occupations with a high presence of women.

In reference to the type of tasks STEM professionals perform, many participants complained about the lack of information they had when they were in secondary school, at a time when they were deciding on their future studies, regarding the specific work that STEM professionals carry out. Similarly, all the participants with work experience (regardless of the STEM discipline) stressed their disillusionment with the lack of correspondence between the expectations they had regarding the work they would be doing and the actual job. However, whereas male participants in technological STEM fields complained about their disillusionment with regard to the development of technical tasks, female participants in these technical fields complained about their disenchantment in terms of social impact. This last aspect confirms the adherence of some participants' expectations and professional goals to existing gender roles in highly male-dominated fields (Eagly, 2001; Diekman et al., 2010). That is, whereas women expressed high expectations of attaining communal goals through STEM occupations with a high technical orientation, men expressed high expectations of attaining agentic goals through these STEM fields. These findings also confirm how in comparison to male participants in non-technological STEM fields, male participants in technological STEM fields use more references to the attainment of masculine agentic goals when portraying the typical person working in STEM.

Finally, and congruent with predictions and other research findings (Cheryan et al., 2013; Sáinz et al., 2016a,b), most role models in STEM fields were mainly masculine (i.e., close family members, university professors, secondary teachers, and workmates), particularly in male-dominated fields such as physical science, computer science, or engineering (Steinke, 2017). However, participants in life science occupations mentioned that, although the presence of women in these professions was considerably high, men's contribution to the field was more salient and men therefore had a better reputation than women in these occupations. Furthermore, while some male participants from technical STEM fields and physical science underlined that some male characters in the media, along with male scientists and science writers, had inspired them, female participants made no such mention. This is further evidence of the lack of female STEM role models in the media. In fact, some participants pointed out how the contribution of women to some STEM fields was not made visible in different STEM fields' education programs.

Contributions

The present study contributes to the literature on young people's portrayals of several STEM fields beyond computer science and technological fields. It is a novel attempt to bring disciplines with a high presence of women, such as medicine and other life science occupations, into the loop of what must be considered as a STEM field. Most studies about these issues do not refer to biological sciences in STEM nor incorporate the prototypical ideal of someone working in biological sciences (Sáinz, 2017). In addition, it provides empirical evidence on how young people already in STEM university studies and jobs perceive the typical person working in STEM. The participants in this research had the opportunity to see the extent to which the stereotypes they had about people working in STEM before entering the STEM field as students or workers had been accurate.

Limitations and Future Research

This study has four relevant limitations. First, the study used a small sample and, therefore, the findings may not be representative of the views of all the STEM students and young professionals residing in Spain. Variations in the educational and cultural backgrounds of the participants may lead to different representations of people working in STEM. Future cross-cultural research could be conducted with a larger sample of STEM students and workers that incorporate their ethnical and educational backgrounds. Second, the precise prevalence of theme endorsement by the interviewees was not possible to determine. The representations expressed by the participants were identified from interview questions and not from a predetermined list of representations. Consequently, the fact that a representation was not mentioned by a participant does not mean that this representation was not endorsed by the participant. Forthcoming international research could incorporate a systematic way of collecting via surveys a list of representations typically associated with the prototypical person working in STEM. Third, the lack of diversity in the participants' belongingness to STEM could be also mentioned as a limitation. Future studies should target more people representing those students and workers from the different STEM fields. Fourth, the different life experiences of people working and studying in STEM fields could be also noted as a constraint when comparing both types of participants' opinions on the topic. Upcoming research could overcome this limitation by asking through closed-ending statements more concrete topics to participants with different life trajectories.

Breaking down stereotypes about people working in STEM is crucial. These stereotypical portrayals tend to discourage many girls, but also some boys to enroll in several STEM fields. Future research on this topic should therefore continue investigating how young people already in STEM perceive the typical person working in STEM occupations. Furthermore, a survey could be designed in order to delve further into the way young people already in STEM perceive the different STEM fields. In addition, these portrayals could inspire young people (particularly girls) and encourage them to enter the field.

Ideas for Intervention

There is a lack of recognition of women's contributions to STEM fields and a dearth of female role models making outstanding contributions. It is therefore necessary for teachers and didactic materials to raise awareness of women's contributions to the STEM field across the different stages of the educational system. In addition, afterschool curricular STEM activities could be designed to increase both boys' and girls' interest in STEM fields by making them more accessible and familiar (Koch and Gorges, 2016). Inviting STEM professionals into schools to talk to students about the day-to-day work they carry could be a good strategy for this purpose. Moreover, more interventions aimed at changing how mass media reproduce the stereotypical portrayal of the ideal person working in STEM are required.

ETHICS STATEMENT

At the time of realization of fieldwork, ethical approval was not required as per local legislation. In addition, participants were not exposed to any physical, psychological, professional or otherwise risk, as they only were posed questions regarding their vision on certain aspects of their personal experience.

All procedures performed in it studies were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed written consent was obtained from all individual participants included in the study.

AUTHOR CONTRIBUTIONS

MS conceptualized, designed, and coordinated the realization of the study and wrote the first draft of the manuscript. SF wrote most of the method section and make insightful comments with regards the justification of the study. JM-C, MR-d-Z, MJR, and LA contributed to the development of the results section. All authors contributed to the development of fieldwork and data analysis and manuscript revision, read, and approved the submitted version.

FUNDING

The present study is part of a broader study funded to the corresponding author by Fundación Telefónica (ES, 2016FT).

ACKNOWLEDGMENTS

We are grateful to all participants who enthusiastically took part in the study.

REFERENCES

Barker, L. J., and Aspray, W. (2006). "The state of research on girls and IT," in *Women and Information Technology: Research on Underrepresentation*, eds J. M. Cohoon and W. Aspray (Cambridge, MA: The MIT Press), 3–54.

Chambers, D. W. (1983). Stereotypic images of the scientist: the draw-a-scientist test. *Sci. Educ.* 67, 255–265. doi: 10.1002/sce.3730670213

Cheryan, S., Play, V. C., Handron, C., and Hudson, L. (2013). The stereotypical computer scientist: gendered media representations as a barrier to inclusion for women. *Sex Roles* 69, 58–71. doi: 10.1007/s11199-013-0296-x

Deaux, K., and Lewis, L. L. (1984). Structure of gender stereotypes: interrelations among components and gender level. *J. Pers. Soc. Psychol.* 46, 991–1004. doi: 10.1037//0022-3514.46.5.991

Diekman, A. B., Brown, E. R., Johnston, A. M., and Clark, E. K. (2010). Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol. Sci.* 21, 1051–1057. doi: 10.1177/0956797610377342

Eagly, A. H. (2001). "Social role theory of sex differences and similarities," in *Encyclopedia of Women and Gender*, ed. J. Worell (San Diego, CA: Academic Press).

Gabriel, U., and Gygax, P. (2016). "Gender and linguistic sexism," in *Advances in Intergroup Communication (Language as Social Action)*, Vol. 21, eds H. Giles and A. Maass (Bern: Peter Lang), 177–192.

Guest, G., Bunce, A., and Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 18, 59–82. doi: 10.1177/1525822x05279903

Kessels, U., and Taconis, R. (2012). Alien or alike? How the perceived similarity between the typical science teacher and a student's self-image correlates with choosing science at school. *Res. Sci. Educ.* 41, 1–23.

Koch, M., and Gorges, T. (2016). Curricular influences on female afterschool facilitators' computer science interests and career choices. *J. Sci. Educ. Technol.* 25, 782–794. doi: 10.1007/s10956-016-9636-2

Kuzel, A. J. (1999). "Sampling in qualitative inquiry," in *Doing Qualitative Research*, eds B. F. Crabtree and W. L. Miller (Thousand Oaks, CA: Sage), 33–45.

Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educ. Rev.* 62, 279–300.

MECD (2018). *Avance de la Estadística de Estudiantes. Curso 2016-2017. Estudios de Grado y Primer y Segundo Ciclo*. Available at: http://www.mecd.gob.es/servicios-al-ciudadano-mecd/estadisticas/educacion/universitaria/estadisticas/alumnado/2016-2017/Grado-y-Ciclo.html (accessed October 15, 2018).

Meyer, M., Cimpian, A., and Leslie, S. J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Front. Psychol.* 6:235–246. doi: 10.3389/fpsyg.2015.00235

Sáinz, M. (2017). *¿Por qué no hay más Mujeres STEM? Se Buscan Ingenieras, Físicas y Tecnólogas. [Why are not there more Women in STEM? Female Engineers, Physical Scientists, and Technologist are Wanted]*. Madrid: Ariel.

Sáinz, M., Meneses, J., Fàbregues, S., and López, B. (2016a). Adolescents' gendered portrayals of information and communication technologies occupations. *Int. J. Gender Sci. Technol.* 8, 181–201.

Sáinz, M., Meneses, J., López, B., and Fàbregues, S. (2016b). Gender stereotypes and attitudes towards ICT in a sample of Spanish secondary students. *Sex Roles* 74, 154–168. doi: 10.1007/s11199-014-0424-2

Sáinz, M., and Müller, J. (2018). Gender and family influences on Spanish students' aspirations and values in stem fields. *Int. J. Sci. Educ.* 40, 188–203. doi: 10.1080/09500693.2017.1405464

Sandelowski, M. (2010). What's in a name? Qualitative description revisited. *Res. Nurs. Health* 33, 77–84. doi: 10.1002/nur.20362

Schreier, M. (2012). *Qualitative Content Analysis in Practice*. London: Sage.

Steinke, J. (2017). Adolescent girls' STEM identity formation and media images of STEM professionals: considering the influence of contextual cues. *Front. Psychol.* 8:716. doi: 10.3389/fpsyg.2017.00716

Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., et al. (2007). Assessing media influences on middleschool-aged children's perceptions of women in science using the Draw A Scientist Test (DAST). *Sci. Commun.* 29, 35–64. doi: 10.1177/1075547007306508

UNESCO (2017). *Cracking the Code: Girls' and Women's Education in Science, Technology, Engineering and Mathematics (STEM).* Available at: https://en.unesco.org/unesco-international-symposium-and-policy-forum-cracking-code-girls-education-stem (accessed October 16, 2018).

Wang, M. T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy–value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 08 May 2019
doi: 10.3389/fpsyg.2019.01053

Traditional Gender Role Beliefs and Career Attainment in STEM: A Gendered Story?

Anna-Lena Dicke, Nayssan Safavian and Jacquelynne S. Eccles*

School of Education, University of California, Irvine, Irvine, CA, United States

Gender role beliefs (i.e., beliefs about gender-specific responsibilities) predict one's educational and occupational aspirations and choices (Eccles et al., 1983; Schoon and Parsons, 2002). Focusing on STEM careers, we aim to examine the extent to which traditional work/family related gender role beliefs (TGRB) in adolescence predict within and across gender differences in subsequent educational and STEM occupational attainment in adulthood. Using longitudinal data from the Michigan Study of Adolescent and Adult Life Transitions ($N = 744$; 58% female), participants' educational attainment and their occupations were assessed at age 42. Their occupations were then categorized into three categories: traditional STEM-related careers in the physical sciences, mathematics, engineering, and technology (PMET); life sciences (e.g., health sciences, LS); and non-STEM. For females, TGRB at age 16/18 significantly predicted lower educational attainment as well as a lower likelihood to be in PMET-related occupations in comparison to non-STEM occupations – controlling for their own educational attainment. TGRB also predicted a higher likelihood to be in LS-related in comparison to PMET-related occupations. No significant associations were found for males. However, patterns of findings for males were similar to those of females. TGRB also mediated across gender differences in educational and PMET-related occupational attainment. Findings reveal TGRB to be one underlying psychological factor influencing gender disparity in educational and STEM occupational attainment.

Keywords: traditional gender role beliefs, educational attainment, STEM, occupational attainment, gender differences

OPEN ACCESS

Edited by:
Silke Luttenberger,
University of Teacher Education Styria, Austria

Reviewed by:
Tiziana Ramaci,
Kore University of Enna, Italy
Josip Burusic,
Institut Društvenih Znanosti Ivo Pilar, Croatia

****Correspondence:***
Anna-Lena Dicke
adicke@uci.edu

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *08 January 2019*
Accepted: *24 April 2019*
Published: *08 May 2019*

Citation:
Dicke A-L, Safavian N and Eccles JS (2019) Traditional Gender Role Beliefs and Career Attainment in STEM: A Gendered Story? Front. Psychol. 10:1053. doi: 10.3389/fpsyg.2019.01053

INTRODUCTION

There has been an increasing number of students aspiring to careers in STEM (science, technology, engineering, and mathematics) in the last decade. The STEM workforce is also increasingly diversifying with respects to gender as female students outnumber male students in some STEM fields, such as biology, medicine, and chemistry (Beede et al., 2011). However, females are still underrepresented in engineering, computer science, and physical sciences (Chen and Ho, 2012). A multitude of reasons for the gender disparities in STEM participation have been investigated, including gender differences in attitudes and beliefs, such as the valuing of various STEM domains (Eccles et al., 1993; Ceci et al., 2014; Lauermann et al., 2015; Cheryan et al., 2017). One of the relevant underlying beliefs that might be driving gender differences in STEM participation are traditional gender role beliefs. These general beliefs about responsibilities and behaviors deemed appropriate for women and men (Eccles, 1987; Williams and Best, 1990) predict aspirations, choices, and occupational outcomes (Eccles et al., 1990). However, the long-term impact of traditional gender role beliefs on STEM participation is less understood.

In the current study, we address this gap in research by investigating the long-term association of traditional gender-role beliefs in adolescence with subsequent educational and STEM occupational attainment in adulthood for females and males using a longitudinal dataset spanning over 20 years. To explore the complexity of the impact of traditional gender role beliefs on these outcomes, we investigated the impact of traditional gender role beliefs within as well as across genders.

Understanding Gender Disparity in STEM Fields

To better address the gender disparities across various STEM fields, the mechanisms behind its emergence need to be better understood. Research has shown that gender differences are evident in the valuing of gender-stereotyped domains such as mathematics and physics with males showing a stronger inclination toward typically male-stereotyped domains and vice versa for females (e.g., Eccles et al., 1993). Similarly, the values underlying various career-related choices are often gendered. For example, females tend to value helping others, improving society, and giving back to their communities relative to males and to other career values such as making lots of money (Lyson, 1984; Eccles, 1987; Konrad et al., 2000). This is consistent with their prevalent interest in human services occupations (Lauermann et al., 2015). Males, on the other hand, are more likely than females to value working with tools and machines and making lots of money, as well as to aspire to careers within traditional male dominant STEM domains (Su et al., 2009; Wang and Degol, 2013; Ramaci et al., 2017).

According to the Eccles et al. (1983) Expectancy-Value theoretical framework, social and contextual factors (such as cultural values, gender role belief systems, social beliefs and behaviors, and prior aptitude and experiences) exert influence onto adolescents' self-beliefs, aspirations, choices, and attainment through their socialization experiences. Thus, gender differences in valuing and subsequent choices are likely results of internalized cultural values and social expectations linked to such belief systems as gender roles (see Eccles et al., 1983, 1990; Eccles, 2015).

Traditional Gender Role Beliefs and Educational and STEM Occupational Attainment

Amongst important internalized social and cultural values are the general beliefs about responsibilities and behaviors deemed appropriate for women and men (Eccles et al., 1983; Eagly, 1987; Williams and Best, 1990; Corrigall and Konrad, 2007): Individuals holding traditional gender role beliefs support women's role as the caretaker at home and in the family and men's role is to provide financial support as the breadwinner of the family. Research has shown that traditional gender role beliefs are more strongly endorsed by men than women (Larsen and Long, 1988; Brewster and Padavic, 2000). These beliefs are linked to greater emphasis being put on men's and husbands'/fathers' careers than on women's and wives/mothers' careers. Such beliefs are then likely to be reflected in individual women's and men's social identities, anticipated future social roles, and short-and long-term goals (Eccles and Bryan, 1994; Eccles et al., 1999). They are also key predictors of their aspirations and both educational and occupational choices (e.g., Schoon and Parsons, 2002; Webb et al., 2002).

Women who endorse traditional gender role beliefs related to family and work roles are more likely to focus on family responsibilities with consequences for the choices they make with regards to educational and occupational aspirations and attainment. For instance, the decrease in traditional work/family related gender role beliefs within society is likely related to increases in educational attainment for females (Buchmann et al., 2008). Female participation in higher education has increased as the prevalence of traditional family related gender role beliefs decreased over time (Brooks and Bolzendahl, 2004; Goldin, 2006). Furthermore, Scott (2004) found a direct link between traditional gender role beliefs and educational attainment: Using data from a National Panel study in Britain, females holding more traditional beliefs about family and work were more likely to show worse performance in their high school exams than females not endorsing traditional beliefs. As expected, given the emphasis of the males' role as a breadwinner within the traditional gender role belief system, this association was not as pronounced for males.

Past studies have also shown associations of endorsement of traditional work/family related gender role beliefs with employment and earnings for females (Cassidy and Warren, 1996; Christie-Mizell, 2006; Corrigall and Konrad, 2007; Buchmann et al., 2008). For instance, Corrigall and Konrad (2007) found that women with more traditional attitudes in their early twenties worked fewer hours and had lower income than women with more egalitarian views in their late twenties using a large nationwide United States sample. In addition, Christie-Mizell (2006) found that endorsement of traditional gender role beliefs was most strongly associated with a decrease in income for white women compared to white and black men and black women within a large-scale longitudinal United States sample.

Although traditional gender role beliefs have become less prevalent over time (Bolzendahl and Myers, 2004; Brooks and Bolzendahl, 2004; Raley et al., 2006), these core beliefs about the roles of women and men in society might help explain still existing differences in STEM occupational choices across gender. According to the Expectancy-Value theoretical framework (Eccles et al., 1983), links between gender role belief systems operate through the association of gender role beliefs with both individuals' gendered expectations for success in and the relative attainment values of various gender typed occupations. Thus, traditional gender role beliefs likely drive across gender differences in STEM-related occupational attainment. With males typically holding more traditional gender role beliefs, they are more likely to seek out high status jobs and thus, pursue STEM-related careers than females, in particular in the traditional STEM fields.

However, the impact of traditional gender role beliefs is likely to be even more complex and might be able to also explain within gender variation in STEM occupational choices. Females are overrepresented in the medical, social, and life sciences, which concern caring and helping others – a value typically endorsed by women with more traditional work/family related gender role beliefs. Females' interest in these specific STEM fields may be due to the values they attach to these specific fields and the extent to which they identify with these values more than other science disciplines. Thus, a stronger endorsement of traditional work/family related gender role beliefs might be perceived to be in accordance with the pursuit of STEM occupations in the life and medical sciences. In contrast, the more traditional STEM fields, such as physics and engineering, are perceived as male-dominated, isolated, and incompatible with the goals of helping others (Eccles et al., 1999; Cheryan et al., 2015). In other words, traditional gender role beliefs should lead those females who go into STEM to be more likely to go into careers in the medical and life sciences than into more traditional STEM fields. The extent to which traditional gender role beliefs can help explain the unequal distribution of females and males in various STEM fields has not been investigated.

Previous research, however, has shown that females with more traditional work/family related gender role beliefs are less likely than males to persist in STEM occupational aspirations than non-STEM occupational aspirations. In a study using earlier waves of the MSALT dataset used in the present study, Frome et al. (1996) found that traditional work/family gender role beliefs held at age 20 were significantly associated with changes in STEM-related occupational aspirations for females. More specifically, they found that females with more traditional gender role beliefs were more likely to change from an occupational aspiration in math, engineering or physical science in 12th grade to an occupational aspiration outside of these fields at age 20. These links were not found for males. Given the wide variation of STEM and non-STEM careers that fit with male gender roles, the association of traditional work/family related gender role beliefs with within gender variation of occupational choices for males is likely to be less pronounced. Frome et al. (2006) found that the impact of traditional work/family related gender role beliefs persisted for females in a follow-up study of a subsample of females that aspired to male-dominated occupational fields in 12th grade. A higher desire for a family flexible job reported in 12th grade was associated with a change of aspirations away from male-dominated occupational fields by age 25.

In sum, gendered beliefs about suitable social roles inform both the pathways and opportunities that are perceived as accessible or socially desirable, as well as the related educational and occupational choices that young people make along the way toward professional attainment. However, despite some exceptions (Frome et al., 1996; Corrigall and Konrad, 2007), more longitudinal studies investigating the long-term associations of traditional gender role beliefs are needed. In addition, there is a lack of studies investigating the associations of traditional gender role beliefs with gendered patterns of STEM-related occupational attainment using a differentiated conceptualization of traditional STEM fields and medical and life sciences.

Current Study

In the current study, we address these gaps in existing research by examining a developmental model spanning over 20 years investigating the association of traditional work/family related gender role beliefs in adolescence with educational and STEM-related occupational attainment in adulthood. Furthermore, to accurately capture the representation of males and females in various STEM fields, we created and used a classification of STEM occupations that differentiates the classic STEM disciplines (i.e., physical sciences, engineering, mathematics, and technology, PMET) from the life and medical sciences (LS). In addition, to account for the impact of participants' socio-demographic family background, we included mother's educational background as a predictor of participant's educational and occupational attainment. Using a longitudinal dataset and building on work from Frome et al. (1996), we asked the following research questions:

RQ1: To what extent do traditional gender-role beliefs held in adolescence (age 16/18) predict subsequent educational and STEM occupational attainment in adulthood (age 42) for females and males?

Taking into account the within gendered pattern of occupational choices found in previous work by Frome et al. (1996), we first investigated the associations of traditional gender role beliefs and subsequent educational and STEM occupational attainment separately for male and female adolescents. Based on previous research (e.g., Scott, 2004), we hypothesized that stronger endorsements of traditional gender role beliefs during adolescence would be associated with lower levels of education in adulthood (as measured by years of formal education) amongst females, but not males. We hypothesized that stronger endorsement of traditional gender role beliefs in adolescence would be associated with a reduced likelihood of occupational attainment within male-typed STEM domains (i.e., PMET) compared to non-STEM occupations amongst females, but not males. We also hypothesized that stronger endorsement of traditional gender role beliefs in adolescence would increase the likelihood to be in less male-typed STEM domains (i.e., LS) compared to non-STEM careers. Lastly, we hypothesized that traditional gender role beliefs in adolescence would decrease the likelihood of occupational attainment in less male-typed STEM domains (i.e., LS) relative to male-typed STEM domains (i.e., PMET).

RQ2: Are gender differences in educational and STEM occupational attainment in adulthood (age 42) mediated by traditional gender role beliefs in adolescence (age 16/18)?

We hypothesized that across gender differences in educational and STEM occupational attainment will be mediated by traditional gender role beliefs. Given previous research (Brewster and Padavic, 2000), we hypothesized that males will hold more traditional gender role beliefs than females. Thus, we hypothesized that gender differences in the endorsement of traditional gender role beliefs by males than females will explain differences in rates of educational attainment and STEM-related occupational attainment between males and females. More specifically, we hypothesized that stronger endorsement of traditional gender role beliefs by males will explain a higher rate of attainment of PMET-related compared to non-STEM occupations. In contrast, we hypothesized that a higher rate of attainment of LS-related compared to non-STEM occupations of females will be explained by males' higher levels of traditional gender role beliefs. The same holds true for the comparison of LS-related and non-STEM related careers.

MATERIALS AND METHODS

Sample

The current study used data from the large scale longitudinal Michigan Study of Adolescent and Adult Life Transitions (MSALT) that followed 2,474 participants over a time span of 30 years from the end of elementary school at age 11 into adulthood at age 42. Participants were from largely middle-income communities located within a large industrial Midwestern city in Michigan, United States and largely from European American descent (91%). We used parent reported data from Wave 1 (participants in grade 6/age 12) and participant self-reported data from Waves 5 and 6 (grade 10/age 16 and grade 12/age 18), and Wave 10 (age 42). In Wave 10, data was collected through surveys via mail, via phone interviews and via web search using social media profiles (i.e., LinkedIn, Facebook). For participants located through web search, educational and occupational attainment was assessed using the information presented in online profiles. All Wave 10 participants with valid data for occupational attainment were included in the current study (n = 744; 58% female; 93% European American). This subsample constituted 89% of the overall Wave 10 sample and 30% of the original sample. Attrition analyses using the original sample of 2,474 participants showed that Wave 10 participants differed significantly from the participants that had dropped out of the study: The Wave 10 sample had a significantly higher rate of females than the original sample [$t(2,470) = 3.435$, $p = 0.001$], participants reported significantly lower levels of traditional gender beliefs at age 16/18 [$t(1,840) = 3.240$, $p = 0.001$], and their mothers reported significantly higher educational background [$t(1,927) = -6.524$, $p = 0.000$].

Measures

All measures were assessed using survey questionnaires. Up to Wave 6 of data collection, participants received and filled out surveys at school. Parents filled out surveys at home. In Wave 10, surveys were mailed to prior participants. In addition, four percent of Wave 10 data were collected through phone interviews and 33 percent of Wave 10 were collected via web search.

Traditional Gender Role Beliefs

Participants' traditional gender role beliefs with regards to job responsibilities were assessed at Wave 5 (age 16) and Wave 6 (age 18) with a 5-item scale ($\alpha = 0.83/0.80$, e.g., "It is usually better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family," see **Appendix A**). This scale assesses beliefs about the relative importance of a man's vs. a woman's career and beliefs about better dispositions of men for career success. The scale was developed by Eccles et al. (1983) and validated in previous studies (Belansky et al., 1993; Frome et al., 1996). Students rated items on a 7-point Likert scale ranging from 1 = Disagree to 7 = Agree. To minimize missing data, missing students' reports from Wave 6 were supplemented by Wave 5 reports.

Participants' Educational Attainment

At Wave 10 (age 42) participants reported their highest attained educational level (Range: 1 = "12th grade or less" to 10 = "Doctorate degree"). For Wave 10 participants that were located through web search, information was coded using available information.

Maternal Educational Attainment

Participants' mothers were asked to report their highest attained educational level at Wave 1 (Range: 1 = "Grade school" to 9 = "Ph.D or professional degree"). In addition, participants were asked to report their mother's educational level at Wave 5 (age 16) with responses ranging from 1 = Grade school to 6 = Graduate school. To minimize missing data, parents' reports from Wave 1 were converted to the 1–6 response scale and supplemented by Wave 5 student reports.

Participants' Occupational STEM Attainment

At Wave 10 (age 42) participants were asked to report their current occupation. If participants were not currently working, they were asked to report their most recent occupation (n = 75).

For the present analyses, the open-ended answers were first coded using employment classification standards set by the United States Bureau of Labor Statistics and the 2010 standard occupational classification (SOC) system manual (U.S. Bureau of Labor Statistics, 2010). Next, SOC-coded occupations were further coded for STEM using U.S. Department of Labor's STEM classification recommendations and subsequently collapsed to three categories to capture the type of STEM-relatedness: traditional STEM-related careers in the physical sciences, engineering, mathematics, computing and technology (PMET; e.g., engineers, surveyors, and mapping scientists, mathematical scientists, physicists, and astronomers, etc); LS (e.g., biology, health sciences, LS; e.g., biologist, physical therapists, nurses, dentists, and veterinarians, etc); and non-STEM. The categorization of non-STEM occupations was guided by our research question and therefore comprised occupations in the social sciences as well all other occupations (including legislators, chief executives and general administrators, teachers, social workers, homemaker, etc). Three dichotomized indicator variables for each of the STEM categories (LS, PMET, and

Non-STEM) indicating membership in the respective category (e.g., 1 = LS-related occupation) were computed.

Statistical Analyses

To investigate the longitudinal associations of traditional gender role beliefs in adolescence with occupational and educational STEM attainment in adulthood for females and males, multi-group manifest path analyses by gender were conducted. In addition, we used the model comparison approach advocated by Judd et al. (1995, 2009) in our analyses, which encourages the use of specific focused comparisons to test specific theoretically derived comparisons. In this case, we conducted three separate path models comparing the different types of STEM-related careers using the following pair comparisons: LS vs. non-STEM, PMET vs. non-STEM, and LS vs. PMET. These comparisons not only allowed us to compare the differentiated STEM careers with non-STEM careers, but also with each other. In the models, educational attainment and STEM occupational attainment in adulthood (Wave 10, age 42) were regressed on traditional gender role beliefs in adolescence (Waves 5 and 6, ages 16 and 18). Educational attainment also predicted STEM occupational attainment. To take into account participants' educational family background, mother's educational attainment was included in the model as a covariate of educational attainment and STEM occupational attainment in adulthood. To address the associations with dichotomous STEM categories (LS, PMET and non-STEM), logistic regressions path analyses were estimated using mixture modeling in MPlus 7.1 (Muthén and Muthén, 2013). Separate models for females and males were estimated using the KNOWNCLASS option in MPlus. To address missing data ($\leq$14%), models were estimated using maximum likelihood estimation with robust standard errors as well as Montecarlo integration (Muthén and Muthén, 2013).

To investigate whether gender differences in educational and STEM occupational attainment were mediated by traditional gender role beliefs, separate mediation path analyses were conducted for each of the four outcomes of interest (educational attainment, LS vs. non-STEM, PMET vs. non-STEM and LS vs. PMET). For each set of analyses, gender was used to predict the outcome to test for existing gender differences in a first step. Then, a path model was estimated, in which gender and traditional gender role beliefs predict the outcome. In addition, gender predicted traditional gender role beliefs. In order for mediation to be met, four conditions had to be met: First, gender must be related to the outcome. Second, gender must be related to traditional gender role beliefs. Third, traditional gender role beliefs should be a significant predictor of the outcome. Fourth, gender should no longer significantly predict the outcome. If all four conditions are met, full mediation is supported. If only the first three conditions are met, partial mediation is supported (Hayes, 2009). To test for the significance of the indirect effect of gender on the outcome via traditional gender role beliefs, the MODEL INDIRECT command in Mplus was used. In addition, models were estimated in Mplus using maximum likelihood estimation as well as Montecarlo integration to address missing data and dichotomous outcome variables (Muthén and Muthén, 2013).

RESULTS

Descriptive analyses revealed gender differences in the endorsement of traditional gender role beliefs and educational attainment (see **Table 1**). Female participants reported lower endorsement of the traditional gender role beliefs scale [$t(638) = -13.610$, $p = 0.000$] in adolescence and higher educational attainment [$t(689) = 2.964$, $p = 0.003$] in adulthood than male participants. The majority of participants were engaged in non-STEM occupations ($n = 511$, 61%) followed by PMET-related ($n = 147$, 18%) and LS-related occupations ($n = 87$, 10%). As shown in **Figure 1**, gender differences in the distribution emerged. Females were more likely than males to be in LS- related occupations in adulthood [$t(742) = 6.328$, $p = 0.000$], whereas males were more likely than females to be in PMET-related occupations [$t(742) = -4.422$, $p = 0.000$].

Table 2 presents correlations of traditional gender role beliefs, educational and occupational attainment, and their mother's educational attainment separately for males and females. Some gender differences in correlation patterns were evident. For females, traditional gender role beliefs were statistically significantly negatively associated with their mothers' educational attainment and their own educational attainment as adults. Traditional gender role beliefs among females were also statistically significantly negatively associated with being in

TABLE 1 | Descriptive statistics for relevant study variables.

	Total (N = 744)		Female (N = 430)		Male (N = 314)		Range
	M	SD	M	SD	M	SD	
Parent report (Mother)							
Educational attainment (W1)	3.90	1.12	3.88	1.12	3.92	1.12	1–6
Participant reports							
Traditional gender role beliefs (W5/6)	3.09	1.30	2.57	1.03	3.81	1.31	1–7
Educational attainment (W10)	5.90	1.97	6.09	1.93	5.64	1.99	1–10

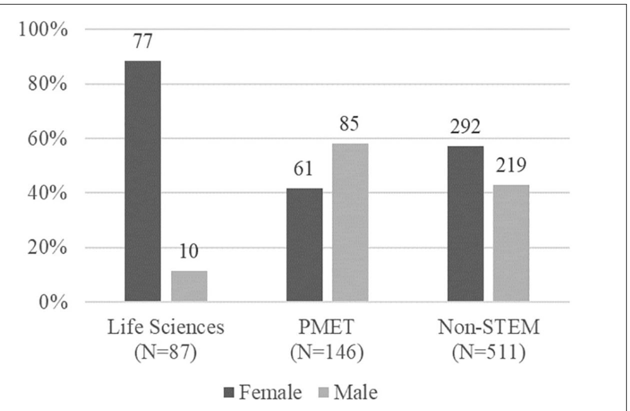


FIGURE 1 | Distribution of STEM-related careers by gender.

TABLE 2 | Bivariate correlations of relevant variables by gender.

	1	2	3	4	5	6
1. Traditional gender role beliefs	–	−0.12	−0.15*	0.02	−0.08	0.07
2. Educational attainment	−0.22***	–	0.26***	0.13*	0.17**	−0.22***
3. Mother's educational attainment	−0.23***	0.28***	–	0.12	0.02	−0.06
4. LS related occupation (=1)	−0.01	0.13**	0.1	–	–	–
5. PMET related occupation (=1)	−0.13**	0.10*	0.04	–	–	–
6. Non-STEM related occupation (=1)	0.11*	−0.18***	−0.11*	–	–	–

*Coefficients for females below diagonal/coefficients for males above diagonal. ***p < 0.001, **p < 0.01, *p < 0.05.*

a PMET-related occupation, but positively associated with being in a non-STEM occupation. Employment within a LS-related occupation (vs. any other occupation) was not associated with traditional gender role beliefs. Females' educational attainment in adulthood was positively associated with their mother's educational attainment and employment in LS- or PMET-related occupations; and negatively associated with being in non-STEM-related occupations. For males, traditional gender role beliefs were also statistically significantly negatively associated with their mother's educational attainment. Educational attainment in adulthood showed the same correlation pattern as for females: It was positively correlated with mother's educational attainment and employment in a LS- or PMET-related occupation, but negatively correlated with employment in a non-STEM occupation.

Traditional Gender Role Beliefs Predicting Subsequent Educational and STEM Occupational Attainment

Figures 2–4 present the results of the multi-group path analyses by gender for each of the STEM category comparisons in our examination of the long-term associations of traditional gender role beliefs with educational and occupational STEM attainment (RQ1).

Educational Attainment

With regards to participants' educational attainment the following pattern was found across all three models (see **Figures 2–4**): For females, traditional gender role beliefs were significantly negatively associated with mother's educational attainment and with their own educational attainment in adulthood. In other words, female participants that endorsed stronger traditional gender role beliefs were more likely to have mothers with lower educational attainment and also more likely to attain lower levels of education themselves. Moreover, their educational attainment was statistically significantly and positively associated with their mother's educational attainment. In other words, females were more likely to attain a higher degree of education when their mothers were also more highly educated. For males, traditional gender role beliefs were marginally negatively associated with mother's educational attainment. Mother's educational attainment was also statistically significantly positively associated with males' own educational attainment in adulthood. However, traditional gender role beliefs were not statistically associated with educational attainment in adulthood for males.

STEM-Related Occupational Attainment

With regards to attainment of LS-related occupations in comparison to non-STEM occupations (see **Figure 2**), traditional gender role beliefs were not associated with attainment of a LS-related occupation for either males or females after taking into account their educational attainment. Educational attainment was statistically significantly associated with a higher likelihood to be in a LS-related career for females, but not males.

With regards to attainment of PMET-related occupations in comparison to non-STEM occupations (see **Figure 3**), females with more traditional gender role beliefs in adolescence were statistically significantly less likely to be employed in PMET-related careers as adults after controlling for their educational attainment. For males, no statistically significant association of traditional gender role beliefs with the likelihood to be in PMET-related careers was found. Higher educational attainment statistically significantly increased the likelihood for being in a PMET-related career for males and females.

With regards to attainment of LS-related occupations in comparison to PMET-related occupations (see **Figure 4**), traditional gender role beliefs statistically significantly increased the likelihood of being in a LS-related career instead of a PMET-related career for females. However, higher educational attainment significantly decreased the likelihood of being in a Non-STEM related career for females. The likelihood of being in a LS- vs. a PMET-related occupation was not associated with endorsements of traditional gender role beliefs for males. Moreover, higher educational attainment did not significantly predict the likelihood of being in a LS- vs. a PMET related occupation for either gender.

Gender Role Beliefs as Mediators of Gender Differences in Educational and STEM Occupational Attainment

To examine whether traditional gender role beliefs explain the gender differences in educational and STEM occupational attainment, separate mediation path analyses were conducted for each of the relevant outcomes (RQ2). Gender was significantly related to all outcomes: Males were more likely to be in a PMET-related career in comparison to a non-STEM career [OR = 1.86, 95% CI (1.28, 2.70)]. In contrast, females were more likely to be in a LS-related career compared to a non-STEM career [OR = 0.17, 95% CI (0.09, 0.34)] as well as when compared to a PMET-related career [OR = 0.09, 95% CI (0.05, 0.20)]. Females also had more years of schooling than males [$b = -0.11$, 95% CI (−0.19, −0.04)].

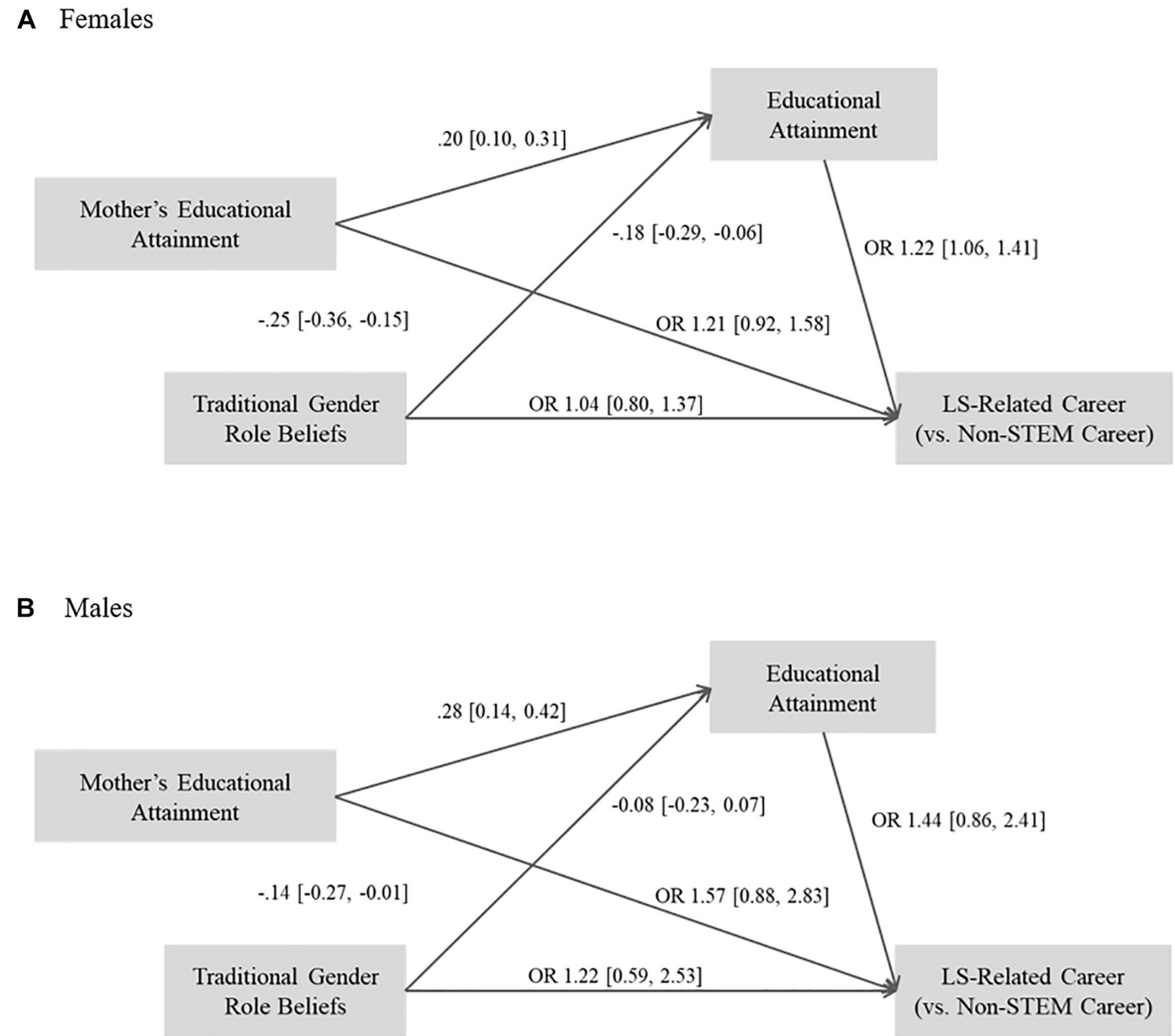


FIGURE 2 | Results of multi-group path analyses by gender for LS-related careers vs. non-STEM careers. **(A)** Results for females and **(B)** results for males. Standardized coefficients are shown for continuous variables. OR, Odds Ratio. 95% Confidence intervals in brackets. Confidence intervals excluding 0/1 for regression coefficients/odds ratios indicate statistical significance.

For educational attainment (see **Figure 5A**), gender differences in educational attainment were fully mediated by traditional gender role beliefs, as the association of gender and educational attainment was no longer significant after including traditional gender role beliefs as the mediator. In addition, results indicated that the indirect effect was significant [$b = -0.10$, 95% 95% CI (−0.14, −0.05)].

Gender differences in the likelihood to be in a LS-related career vs. a non-STEM career were not statistically significantly mediated by traditional gender role beliefs (see **Figure 5B**). The association between gender and endorsement of a LS-related career remained significant after including traditional gender role beliefs in the model and no significant association of traditional gender role beliefs with LS-related career attainment was found. Thus, the indirect effect was not significant [OR = 0.92, 95% CI (0.64, 1.20)].

However, gender differences in the likelihood to be in PMET-related career vs. a non-STEM career were partially mediated by traditional gender role beliefs (see **Figure 5C**). The higher likelihood of males to be in a PMET-related career remained statistically significant after the inclusion of traditional gender role beliefs in the model, but results indicated a statistically significant indirect effect [OR = 0.76; 95% CI (0.61, 0.92)].

Lastly, gender differences in the likelihood to be in a LS- vs. a PMET-related occupation were not mediated by traditional gender role beliefs (see **Figure 5D**). The higher likelihood of females to be in a LS-related career remained statistically significant after the inclusion of traditional gender role beliefs in the model and no significant association of traditional gender role beliefs and LS-related vs. PMET-related attainment was found. Thus, the indirect effect was not statistically significant [OR = 1.43, 95% CI (0.87, 1.99)].

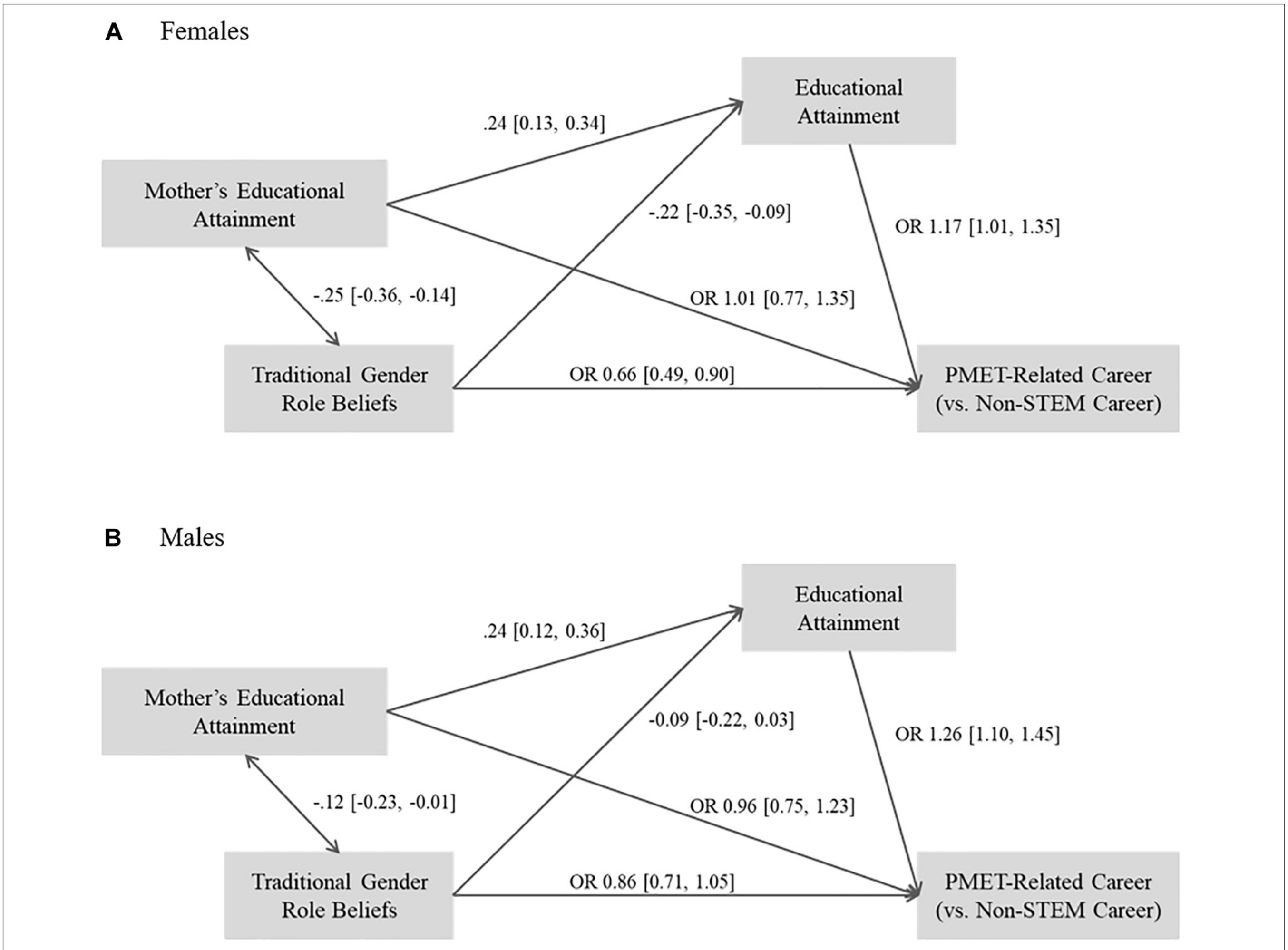


FIGURE 3 | Results of multi-group path analyses by gender for PMET-related careers vs. non-STEM careers. **(A)** Results for females and **(B)** results for males. Standardized coefficients are shown for continuous variables. OR, Odds Ratio. 95% Confidence intervals in brackets. Confidence intervals excluding 0/1 for regression coefficients/odds ratios indicate statistical significance.

DISCUSSION

The current study investigated the impact of traditional work/family related gender role beliefs in adolescence on educational and STEM occupational attainment in adulthood using a longitudinal dataset spanning 20 years. As an important determinant of life choices, traditional work/family related gender role beliefs were used to investigate impacts on educational and occupational attainment in PMET, LS, and non-STEM occupational attainment within and across gender. By doing so, we fill a need for longitudinal studies on the impact of traditional gender role beliefs as well as address the lack of STEM differentiation when investigating its impact on gendered occupational choices in previous research. This is particularly noteworthy given the misrepresentation of women in STEM when LS occupations and PMET occupations are not differentiated. By highlighting these differentiated associations we can better contribute to the conversation of how we can better represent and support females' STEM-related choices.

Impacts of Traditional Gender Role Beliefs on Subsequent Educational and STEM Occupational Attainment Within Gender

Our investigation of the impact of traditional work/family related gender role beliefs revealed a nuanced pattern of findings for females. As hypothesized, females with stronger traditional gender role beliefs in adolescence attained lower levels of education in adulthood – a finding that further supports previous work by Scott (2004). One explanation for this association could be that the endorsement of traditional gender roles during adolescence (e.g., beliefs about women's role as the caretaker at home and in the family) may be a reflection of young women's expectations for marriage and child bearing early on and their reliance on men's role to provide financial support as the breadwinner of the family. If so, this explanation would be in congruence with findings by Corrigall and Konrad (2007) that found that women with more traditional attitudes worked fewer

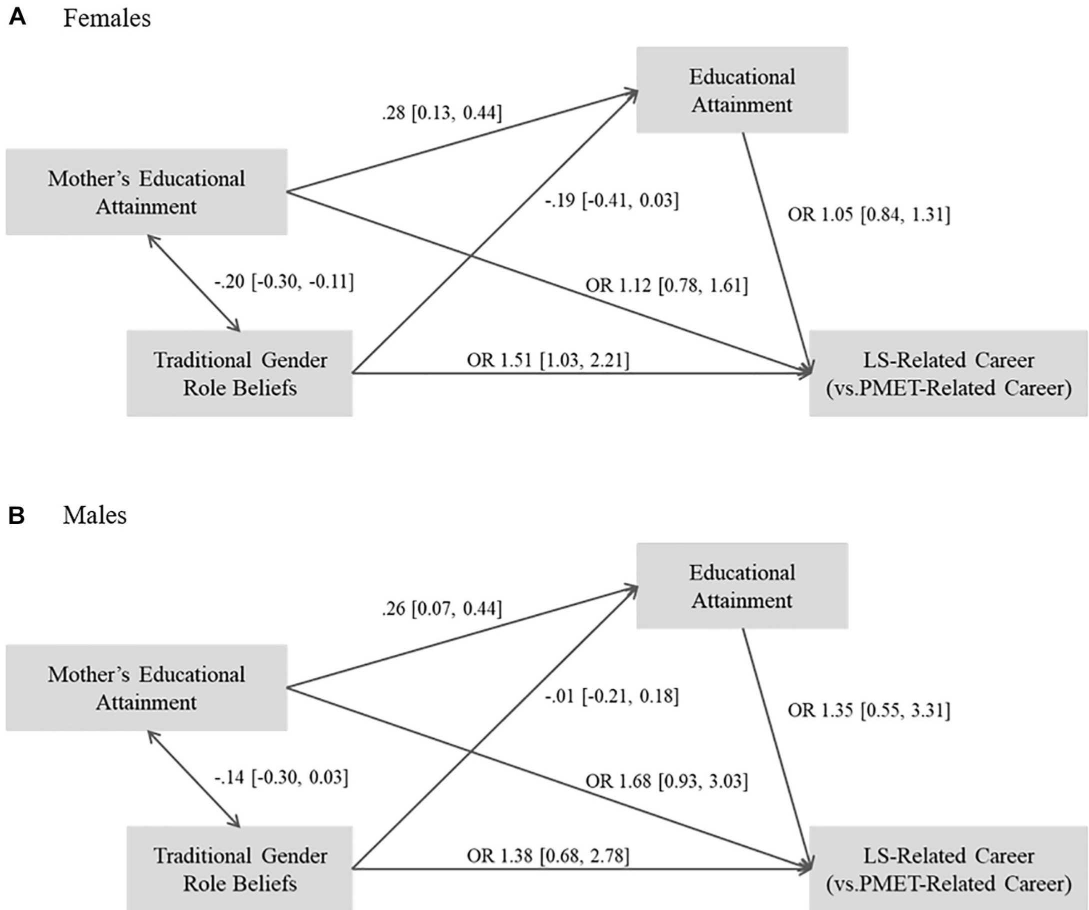


FIGURE 4 | Results of multi-group path analyses by gender for LS-related careers vs. PMET-related careers. **(A)** Results for females and **(B)** results for males. Standardized coefficients are shown for continuous variables. OR, Odds Ratio. 95% Confidence intervals in brackets. Confidence intervals excluding 0/1 for regression coefficients/odds ratios indicate statistical significance.

hours and had lower income than women with more egalitarian views in their late twenties.

By using a differentiated classification of STEM-related occupations, we also found, as hypothesized, that females' endorsement of traditional gender role beliefs in adolescence reduced the likelihood of occupational employment within PMET domains, but was not associated with their occupational attainment within LS domains. In addition, more traditional gender role beliefs actually predicted occupational attainment within LS domains over PMET domains. The reduced likelihood of occupational attainment in a PMET domain among females that endorse traditional gender role beliefs lends further support to research that has documented male and female value of gender-stereotyped domains in alignment with their respective gender (e.g., Eccles et al., 1993). However, our nuanced findings with regards to the effects on occupational attainment in PMET- and LS-related careers underline the importance of using a differentiated conceptualization of STEM domains. Endorsement of traditional gender role beliefs did not affect females' occupational attainment in LS domains negatively. Thus, to truly capture and understand the origins of gender differentiation in the STEM field, a broader conceptualization of STEM-related occupations that is fully inclusive of LS such as health and medicine is needed. This will not only allow for a better scope of STEM-related or, more broadly speaking, science-related occupations, but it will also more accurately represent the participation of women in STEM.

However, it is important to note that our models accounted for females' educational attainment in adulthood; and for females, their endorsement of traditional gender role beliefs were negatively associated with their educational attainment. It may be that there is an indirect link between traditional gender role beliefs and STEM-typed occupational attainment that is mediated by educational attainment. This might be especially relevant as STEM occupations generally require a higher degree of educational attainment and technical training relative to non-STEM occupations.

Traditional gender role beliefs did not significantly associate with educational or STEM-related occupational attainment for male participants. However, interestingly, associations of traditional gender role beliefs and STEM occupational attainment were in the similar direction as for females, pointing to

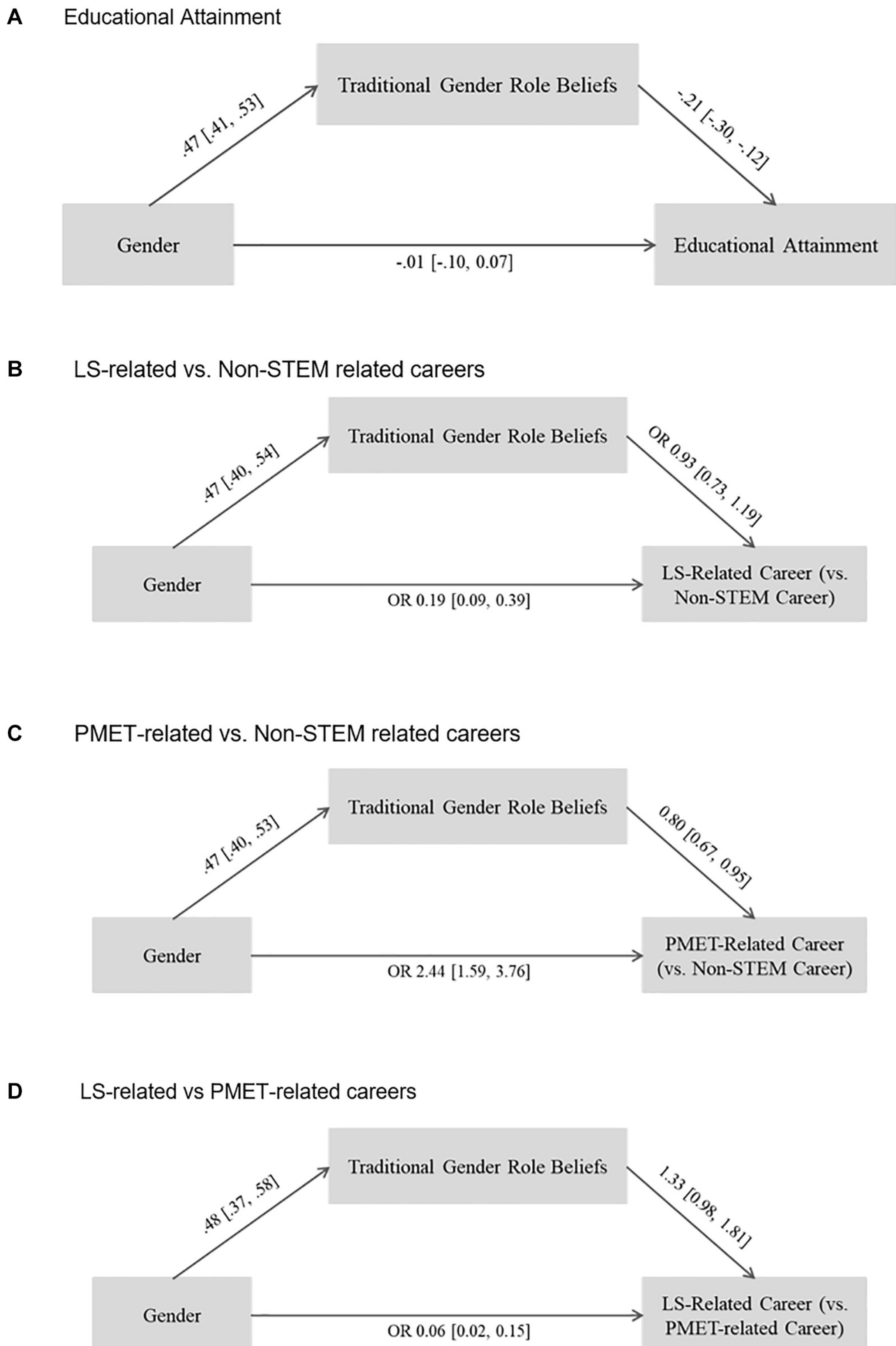


FIGURE 5 | Results of path analyses investigating the mediation of the association of gender with educational and STEM occupational attainment outcomes via traditional gender role beliefs. Gender coded 1, male. **(A)** Results for educational attainment, **(B)** results for comparison of LS-related vs. non-STEM related careers, **(C)** results for comparison fo PMET-related vs. non-STEM related careers, and **(D)** results for comparison of LS-related vs. PMET-related careers. Standardized coefficients are shown for continuous variables. OR, Odds Ratio. 95% Confidence intervals in brackets. Confidence intervals excluding 0/1 for regression coefficients/odds ratios indicate statistical significance.

a similar pattern of impact for males as for females, only less pronounced. Particularly with regards to STEM-related occupational attainment, one reason for the non-significance of the effects for males might be the small sample size of males in LS-related careers. It also needs to be noted that coefficients for females and males were not statistically significantly different from each other.

Gender Differences in Educational and STEM Occupational Attainment: Impact of Traditional Gender Role Beliefs

Our investigation of whether traditional work/family related gender role beliefs are related to across gender differences revealed that gender differences in the endorsement of traditional gender role beliefs explain differences in the rates of educational attainment and STEM-related occupational attainment of males and females. More specifically, as expected higher educational attainment by females was mediated by lower endorsement of traditional gender role beliefs by females. In addition, as expected stronger endorsement of traditional gender role beliefs by males partially explained a higher rate of attainment of PMET-related careers compared to non-STEM careers. However, gender differences in attainment of LS-related occupations in comparison to non-STEM occupations and PMET-related occupations were not mediated by traditional gender role beliefs. The found effects were, however, in the expected direction and might have been affected by the low sample size of males in LS occupations in the current sample. Thus, in accordance with the Expectancy-Value theoretical framework (Eccles et al., 1983), our study provides some evidence that traditional gender role beliefs are one potential underlying psychological factor that can help explain gender disparity in attainment. This finding further highlights that it is important to have a differentiated conceptualization of STEM occupations, as STEM occupations encompass a variety of occupations with differential values attached to them by males and females.

Given our findings, one potential way to address the existing gender disparity in the traditional STEM fields could be to better contextualize the human applications of these fields to attract more females. It would be equally prudent to address the stereotype of PMET-related occupations as male-typed domains, that are isolating and incompatible with the goals of helping others (Cheryan et al., 2015). This might be deterring females from aspiring to such occupations. On the other hand, our findings indicate that changes in the socialization of societal gendered expectations with a movement to more egalitarian gender role beliefs, as currently ongoing (Brooks and Bolzendahl, 2004), will ultimately help ease gender disparities in educational and STEM occupational attainment.

Limitations and Future Research

While the longitudinal dataset used in the present study allowed for an investigation of the long-term impact of traditional gender role beliefs, it needs to be kept in mind that the present longitudinal sample was biased toward lower levels of traditional gender role beliefs due to attrition. As a result, the present study did not present the full variation in traditional gender role beliefs that likely exist in the general population. Our present sample was also biased toward having mothers with a higher level of education. Our results, thus, do not represent the full spectrum with regards to participant's socioeconomic background. Given these constraints with regards to variation in traditional gender role beliefs and socio-economic background, our findings likely underestimate the effects of traditional gender role beliefs on educational and STEM occupational choices. Lastly, the present sample also consisted of a higher rate of females than males due to attrition. As a result, the sample size for individual STEM categories (e.g., LS) was small for male participants. This means that these particular findings need to be interpreted with caution due to the lack of power. To address the bias in our present sample, future research should replicate the findings using a more gender balanced sample capturing effectively the whole spectrum of traditional gender role beliefs, STEM occupations, and socio-economic backgrounds to test generalizability.

Our findings illustrate how general beliefs about societal norms, i.e., traditional gender role beliefs, can affect specific life choices in important life domains, i.e., educational and occupational attainment. Our findings did, however, not look into the educational and occupational trajectories of the participants to see how educational and occupational aspirations and choices developed over time. This important future avenue for research would allow us to better understand the educational and occupational pathways taken by females and males. Such analyses might shine a light on whether females and males differ in the timing or variation of educational and occupational choices, which might, in turn, affect their eventual educational, and occupational attainment.

Future research should also examine the mechanisms through which traditional gender role beliefs affect educational and occupational choices. As previously discussed, traditional gender role beliefs are likely to inform valuing of education and particular STEM domains, which, in turn, determine occupational choices (Wigfield and Eccles, 2000). They might also inform gender-specific stereotypes about women's lack of competencies in STEM majors and occupations, which have been found to negatively influence STEM choices for women (Nosek and Smyth, 2011; Cundiff et al., 2013). These possible ways through which traditional gender role beliefs might differentially affect educational and occupational choices for females and males, particularly in STEM, need to be empirically tested.

In addition, apart from exploring the processes driving the impact of traditional gender role beliefs on career choices, future analyses should explore how other important life choices (e.g., marriage, children) mediate or moderate the impact of traditional gender role beliefs on educational and occupational attainment. More importantly, particularly life choices with regards to the timing of marriage and child bearing very likely affect educational and occupational pathways differentially for females and males. As such, another significant avenue of research will also be to examine actual, and perceived, opportunities for employment and lifestyle affordances (i.e., number of hours worked, work-life balance) of STEM-related domains by men and women that could contribute to gender-differentiated choices and pathways as a function of their gender role beliefs. For example, women might gravitate more toward LS-typed careers if there are a greater number of opportunities for work in non-academic settings as opposed to traditional science domains (Ceci et al., 2014). Research is beginning to examine the congruence of perceived affordances and desired goals in explaining gender-differentiated STEM occupational choices (e.g., Diekman et al., 2016). It will be imperative to continue this avenue of research and examine how gender roles beliefs inform a socially constructed narrative of perceived abilities, affordances, and anticipated goals and resultant choices, if we are to support continued opting into these STEM fields.

Overall, our findings showcase the importance of culturally socialized general beliefs about society, in this case traditional work/related gender role beliefs, in influencing the specific life choices women and men make, and specifically their potential in explaining disparate gender participation in STEM.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Institutional Review Board of the University of Michigan and the University of California, Irvine, Irvine, CA, United States with written informed consent from all subjects. Written informed consent for subjects under the age of 16 was obtained from parents. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Institutional Review Boards of the University of Michigan and the University of California, Irvine, Irvine, CA, United States.

AUTHOR CONTRIBUTIONS

A-LD and NS conceived the idea of the study. JE was the architect of the data used in the study. A-LD conducted the analyses and wrote the manuscript with feedback and assistance from NS and JE.

FUNDING

This research was supported in part by grants from the National Science Foundation (# BNS 85-10504 and # 1108778), the National Institute of Mental Health (# MH31724), and the National Institute of Child Health and Human Development (# HD17296) to JE.

REFERENCES

Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., and Doms, M. E. (2011). *Women in STEM: A Gender Gap to Innovation. Economics and Statistics Administration Issue Brief No. 04-11*. Available at: https://ssrn.com/abstract=1964782 (accessed November 15, 2018).

Belansky, E. S., Early, D. M., and Eccles, J. S. (1993). "The impact of mothers and peers on adolescents' gender role traditionality and plans for the future," in *Poster Session Presented at the Meeting of the Society for Research on Child Development*, New Orleans, LA.

Bolzendahl, C. I., and Myers, D. J. (2004). Feminist attitudes and support for gender equality: opinion change in women and men, 1974–1998. *Soc. Forces* 83, 759–789. doi: 10.1353/sof.2005.0005

Brewster, K. L., and Padavic, I. (2000). Change in gender-ideology, 1977–1996:the contributions of intracohort change and population turnover. *J. Marriage Fam.* 62, 477–487. doi: 10.1111/j.1741-3737.2000.00477.x

Brooks, C., and Bolzendahl, C. (2004). The transformation of US gender role attitudes: cohort replacement, social-structural change, and ideological learning. *Soc. Sci. Res.* 33, 106–133. doi: 10.1016/S0049-089X(03)00041-3

Buchmann, C., DiPrete, T. A., and McDaniel, A. (2008). Gender inequalities in education. *Ann. Rev. Sociol.* 34, 319–337. doi: 10.1146/annurev.soc.34.040507.134719

Cassidy, M. L., and Warren, B. O. (1996). Family employment status and gender role attitudes: a comparison of women and men college graduates. *Gend. Soc.* 10, 312–329. doi: 10.1177/089124396010003007

Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Interest* 15, 75–141. doi: 10.1177/1529100614541236

Chen, X., and Ho, P. (2012). *STEM in Postsecondary Education: Entrance, Attrition, and Coursetaking Among 2003–04 Beginning Postsecondary Students (NCES 2013-152)*. Washington, DC: National Center for Education Statistics.

Cheryan, S., Master, A., and Meltzoff, A. N. (2015). Cultural stereotypes as gatekeepers: increasing girls' interest in computer science and engineering by diversifying stereotypes. *Front. Psychol.* 6:49. doi: 10.3389/fpsyg.2015.00049

Cheryan, S., Ziegler, S. A., Montoya, A. K., and Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychol. Bull.* 143, 1–35. doi: 10.1037/bul0000052

Christie-Mizell, C. A. (2006). The effects of traditional family and gender ideology on earnings: race and gender differences. *J. Fam. Econ. Issues* 27, 48–71. doi: 10.1007/s10834-005-9004-5

Corrigall, E. A., and Konrad, A. M. (2007). Gender role attitudes and careers: a longitudinal study. *Sex Roles* 56, 847–855. doi: 10.1007/s11199-007-9242-0

Cundiff, J. L., Vescio, T. K., Loken, E., and Lo, L. (2013). Do gender–science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Soc. Psychol. Educ.* 16, 541–554. doi: 10.1007/s11218-013-9232-8

Diekman, A. B., Steinberg, M., Brown, E. R., Belanger, A. L., and Clark, E. K. (2016). A goal congruity model of role entry, engagement, and exit: understanding communal goal processes in STEM gender gaps. *Pers. Soc. Psychol. Rev.* 21, 142–175. doi: 10.1177/1088868316642141

Eagly, A. H. (1987). *Sex Differences in Social Behavior: A social-role interpretation.* Hillsdale, NY: Lawrence Erlbaum.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychol. Women Q.* 11, 135–172. doi: 10.1111/j.1471-6402.1987.tb00781.x

Eccles, J. S. (2015). Gendered socialization of STEM interests in the family. *Int. J. Gend. Sci. Technol.* 7, 116–132.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). "Expectancies, values, and academic behaviors," in *Achievement and Achievement Motivation*, ed. J. T. Spence (San Francisco, CA: W. H. Freeman), 75–146.

Eccles, J. S., Barber, B., and Jozefowicz, D. (1999). "Linking gender to educational, occupational, and recreational choices: applying the Eccles et al. model of achievement-related choices," in *Sexism and Stereotypes in Modern Society: The Gender Science of Janet Taylor Spence*, eds W. B. Swann, J. H. Langlois, and L. A. Gilbert (Washington, DC: American Psychological Association), 153–192. doi: 10.1037/10277-007

Eccles, J. S., and Bryan, J. (1994). "Adolescence: critical crossroad in the path of gender-role development," in *Gender Roles Through the Life Span*, ed. M. R. Stevenson (Muncie: Ball State University Press), 110–147.

Eccles, J. S., Jacobs, J. E., and Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *J. Soc. Issues* 46, 183–201. doi: 10.1111/j.1540-4560.1990.tb01929.x

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., et al. (1993). Development during adolescence: the impact of stage-environment fit on young adolescents' experiences in schools and in families. *Am. Psychol.* 48, 90–101. doi: 10.1037/0003-066x.48.2.90

Frome, P. M., Alfeld, C. J., Eccles, J. S., and Barber, B. L. (2006). Why don't they want a male-dominated job? An investigation of young women who changed their occupational aspirations. *Educ. Res. Eval.* 12, 359–372. doi: 10.1080/13803610600765786

Frome, P. M., Alfeld-Liro, C. J., and Eccles, J. S. (1996). Why don't young women want to pursue male-typed occupational aspirations? A test of competing hypotheses. *Paper Presented at the Biennial Meeting of the Society for Research on Adolescence*, Boston, MA. doi: 10.1080/13803610600765786

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *Am. Econ. Rev.* 96, 1–21. doi: 10.1257/000282806777212350

Hayes, A. F. (2009). Beyond baron and kenny: statistical mediation analysis in the new millennium. *Commun. Monogr.* 76, 408–420. doi: 10.1080/03637750903310360

Judd, C., McClelland, G., and Ryan, C. (2009). *Data Analysis: A Model Comparison Approach*. New York, NY: Routledge.

Judd, C. M., McClelland, G. H., and Culhane, S. E. (1995). Data analysis: continuing issues in the everyday analysis of psychological data. *Annu. Rev. Psychol.* 46, 433–465. doi: 10.1146/annurev.ps.46.020195.002245

Konrad, A., Ritchie, J. E., Lieb, P., and Corrigall, E. (2000). Sex differences and similarities in job attribute preferences: a meta-analysis. *Psychol. Bull.* 126, 593–641. doi: 10.1037/0033-2909.126.4.593

Larsen, K. S., and Long, E. (1988). Attitudes toward sex-roles: traditional or egalitarian? *Sex Roles* 19, 1–12. doi: 10.1007/BF00292459

Lauermann, F., Chow, A., and Eccles, J. S. (2015). Differential effects of adolescents' expectancy and value beliefs about math and english on math/science-related and human services-related career plans. *Int. J. Gend. Sci. Technol.* 7, 205–228.

Lyson, T. A. (1984). Sex differences in the choice of a male or female career line: an analysis of background characteristics and work values. *Work Occup.* 11, 131–146. doi: 10.1177/0730888484011002001

Muthén, L. K., and Muthén, B. O. (2013). *Mplus user's Guide*, 7th Edn. Los Angeles, CA: Muthén & Muthén.

Nosek, B. A., and Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *Am. Educ. Res. J.* 48, 1125–1156. doi: 10.3102/0002831211410683

Raley, S. B., Mattingly, M. J., and Bianchi, S. M. (2006). How dual are dual-income couples? documenting change from 1970 to 2001. *J. Marriage Fam.* 68, 11–28. doi: 10.1111/j.1741-3737.2006.00230.x

Ramaci, T., Pellerone, M., Ledda, C., Presti, G., Squatrito, V., and Rapisarda, V. (2017). Gender stereotypes in occupational choice: a cross-sectional study on a group of Italian adolescents. *Psychol. Res. Behav. Manag.* 10, 109–117. doi: 10.2147/PRBM.S134132

Schoon, I., and Parsons, S. (2002). Teenage aspirations for future careers and occupational outcomes. *J. Vocat. Behav.* 60, 262–288. doi: 10.1006/JVBE.2001.1867

Scott, J. (2004). Family, gender, and educational attainment in Britain: a longitudinal study. *J. Comp. Fam. Stud.* 35, 565–589. doi: 10.1016/j.alcr.2013.02.001

Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychol. Bull.* 135, 859–884. doi: 10.1037/a0017364

U.S. Bureau of Labor Statistics (2010). *Standard Occupational Classification User Guide*. Washington, DC: U. S. Bureau of Labor Statistics.

Wang, M.-T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy–value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/J.DR.2013.08.001

Webb, R. M., Lubinski, D., and Benbow, C. P. (2002). Mathematically facile adolescents with math-science aspirations: new perspectives on their educational and vocational development. *J. Educ. Psychol.* 94, 785–794. doi: 10.1037/0022-0663.94.4.785

Wigfield, A., and Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015

Williams, J. E., and Best, D. L. (1990). *Measuring Sex Stereotypes: A Multination Study (Rev. ed.)*. Thousand Oaks, CA: Sage.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX A

Traditional Gender Role Belief Scale

In general, men are more reliable on the job than women.

In general, men are naturally more competitive than women.

It bothers me to see a man being told what to do by a woman.

Men are naturally better than women at mechanical things.

It is usually better for everyone involved if the man is the achiever outside the home and the woman takes care of the home and family.

frontiers in Education

ORIGINAL RESEARCH
published: 17 May 2019
doi: 10.3389/feduc.2019.00034

Do Teachers' Beliefs About Math Aptitude and Brilliance Explain Gender Differences in Children's Math Ability Self-Concept?

Anke Heyder[1*], *Ricarda Steinmayr*[1] *and Ursula Kessels*[2]

[1] *Department of Psychology, Technical University Dortmund, Dortmund, Germany,* [2] *Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany*

The self-concept of ability in math in elementary school is an early predictor for future math-related choices and careers. Unfortunately, already at this early age girls report lower ability self-concepts in math than boys—despite their comparable performances in objective math competence tests. In the present study we focus on teachers' beliefs as factors explaining these gender differences. Women's underrepresentation in math and science in academia has recently been explained by the belief held by the environment that success in these domains requires an innate ability that cannot be taught ("brilliance"). In addition, teachers' beliefs regarding their students' mathematical aptitude have also been found to influence students' self-concepts of ability. Here, we study if teachers' beliefs regarding their students' mathematical aptitude and brilliance beliefs may account for gender differences in elementary school students' self-concept of ability in math and thus potentially contribute to entering the gendered path into math and science professions. In a sample of 830 fourth graders ($M = 9.14$ years old, 49% female) and 56 elementary school teachers from Germany, we assessed teachers' beliefs regarding their students' mathematical aptitude and their belief that children need brilliance to succeed in math as well as children's mathematical ability self-concept and competencies. In line with prior research, boys reported a statistically significantly more positive math ability self-concept ($d = 0.50$), although boys and girls reached similar scores in a standardized math competence test ($d = 0.07$). However, multilevel regression analyses revealed that teachers' math brilliance beliefs were not related to the gender gap in students' ability self-concept in expense of girls whereas the gender gap was mediated by teachers' beliefs about their students' mathematical aptitude. These findings suggest that math brilliance beliefs held by important socializers such as teachers might not play a role in explaining gender differences in math-related motivation in elementary school whereas teachers' beliefs about students' math aptitude do. Results are discussed against the background of teacher expectancy effects, developmental changes in elementary school, and cultural differences.

Keywords: gender differences, ability self-concept, math, teacher beliefs, implicit theories, elementary school

OPEN ACCESS

Edited by:
Manuela Paechter,
University of Graz, Austria

Reviewed by:
Gerardo Ramirez,
University of California, Los Angeles, United States
Dreisiebner Gernot,
University of Graz, Austria

Correspondence:
Anke Heyder
anke.heyder@tu-dortmund.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *08 January 2019*
Accepted: *08 April 2019*
Published: *17 May 2019*

Citation:
Heyder A, Steinmayr R and Kessels U (2019) Do Teachers' Beliefs About Math Aptitude and Brilliance Explain Gender Differences in Children's Math Ability Self-Concept? Front. Educ. 4:34. doi: 10.3389/feduc.2019.00034

INTRODUCTION

"You need to be a math person to be good at math." Such beliefs are widespread in Western industrialized societies (Rattan et al., 2012; Chestnut et al., 2018). In the context of higher education, research has already shown their potentially damaging implications for the diversity of the learners. Believing in an innate ability required for success in math—that is, math brilliance, which you either have or have not—seems to be particularly detrimental for women and people of color, two social groups that are stereotyped as lacking an innate ability for math (Storage et al., 2016; Chestnut et al., 2018). Thus, believing in brilliance required for success in a particular domain has been identified as one factor contributing to the underrepresentation of women in domains such as science, technology, engineering, and math (i.e., STEM domains, Wang and Degol, 2016).

Gender differences regarding math, however, do not just emerge when it comes to math course selection in high school or college but much earlier in students' educational careers. Therefore, it is important to understand the school experiences and motivational processes that lay the groundwork for (not) pursuing math-related careers (Wang and Degol, 2013, p. 323, 324). In this study, we focus on students in their last year of elementary school, i.e., fourth grade, and their teachers. We test if teachers' belief that math requires innate ability and the ascription of math aptitude to boys also account for gender differences in these students' motivation in math and thus potentially contribute to entering the gendered path into math and science professions.

In what follows, we shortly summarize empirical evidence on gender differences in students' math achievement and motivation in elementary school, two important predictors of students' math-related choices (e.g., Eccles et al., 1983; Wang and Degol, 2013). Subsequently, we synthesize (a) research on the general role of teachers' beliefs for gender differences in school with (b) research on the belief that math requires brilliance, which seems to be widespread in academia and to be a barrier for women's success and participation in STEM fields (Leslie et al., 2015; Rattan et al., 2018). This results in our research question whether teachers' beliefs about (a) students' math aptitude and (b) brilliance in general explain gender differences in children's math ability self-concept in expense of girls already at the end of elementary school.

Gender Differences in Math Achievement and Motivation in Elementary School

Many people believe that an average male person possesses higher math skills and talent than an average female person (e.g., Nosek et al., 2002; Steffens et al., 2010; Ertl et al., 2017). However, contrary to this stereotype, actual mean gender differences in math competence tests are generally small and often negligible (for meta-analyses see e.g., Else-Quest et al., 2010; Reilly et al., 2015). This has also been found for elementary school students. In the recent Trends in Math and Science Study (TIMSS 2015), for instance, no significant differences between male and female fourth graders were found across all participating countries (Mullis et al., 2016). In countries where boys outperformed girls, differences were mostly small. Other studies with elementary school samples yielded similar results (Herbert and Stipek, 2005; Cvencek et al., 2015). However, while mean differences in boys' and girls' math performance in early years tend to be small or non-existent, larger gender differences are found at the highest ability level (Brunner et al., 2008). Both in earlier and more recent cohorts of early childhood studies, the gender gap developed early at the top of the achievement distribution and spread throughout the distribution during the first few years of elementary school (Cimpian et al., 2016). For example, when considering the combined PIRLS and TIMSS test results the gender ratio in the highest performance level was four girls to five boys (Bergold et al., 2017).Gender differences in math also become larger in older samples concerning both the average performance and the gender gap among top performers (Hyde, 2005; Reilly, 2012). In older students, i.e., PISA 2009 participants, twice as many male (2.5%) than female (1.2%) students reached the highest ability levels in math (Reilly, 2012). However, most researchers agree that gender differences in math ability are not the primary explanation for why female students end up working and majoring less often than male students in math-intensive fields (see also Riegle-Crumb et al., 2012; Wang and Degol, 2013, p. 308).

The seminal Eccles et al. expectancy-value model (e.g., Eccles et al., 1983; Eccles, 2011) provides a popular theoretical model of the processes that lead to gender differences in academic achievement and choices. According to Eccles et al., the two most proximal precursors of academic choices are students' belief about how well they will do on upcoming tasks in the respective domain (i.e., their expectation of success in math) and how much they like, value, and enjoy the respective domain (i.e., their math values), with both constructs being influenced by the child's ability self-concept. The child's ability self-concept is defined as the child's beliefs about its competence in the respective domain (here, math) and is assumed to be directly influenced by socializer's beliefs and stable child characteristics such as child's gender or aptitude (Eccles and Wigfield, 2002, p. 119). As math ability self-concept is an empirically established, powerful predictor of students' math-related career choices (Wang and Degol, 2013; Musu-Gillette et al., 2015) we concentrate on ability self-concepts in the present study. And already in elementary school, boys report a more positive ability self-concept than girls (e.g., Tiedemann, 2000b; Fredricks and Eccles, 2002; Herbert and Stipek, 2005). Importantly, these gender differences in math ability self-concept in expense of female students are larger than those found in test performance, and thus, cannot be traced back on actual competence differences. Moreover, gender gaps in ability self-concepts (or confidence) seem to be larger than in intrinsic motivation (or interest) or the perceived usefulness and importance of math (Wigfield et al., 1997; Ganley and Lubienski, 2016) suggesting that children's competence beliefs are more likely to be affected by prior learning experiences and socializers' beliefs in a gender-specific way than children's valuing of the respective domain. In addition, these findings indicate that a domain that is stereotyped as male will still be highly valued by girls and women irrespective of their perceptions of their own abilities, possibly reflecting the high status of those fields or occupations perceived as male or masculine in our society.

The Role of Socializers' Beliefs

Socializers such as teachers or parents are considered as powerful influences on gender gaps in academic motivation, choices, and achievement during school (see e.g., Gunderson et al., 2012). According to the expectancy-value model (e.g., Eccles et al., 1983) socializers' beliefs directly affect students' beliefs such as their stereotypes and ability self-concepts. As teachers' beliefs are thought to influence parents' beliefs concerning their child competencies more profoundly than vice versa (Simpkins et al., 2015), we concentrate on teachers' beliefs. In general, the important role of teachers' beliefs concerning children's ability self-concept formation has been underlined by many empirical studies (e.g., Keller, 2001; Herbert and Stipek, 2005; Wolter et al., 2015; for related research on the role of parents, see e.g., Tiedemann, 2000b; Tomasetto et al., 2015). For instance, teachers' belief that girls have superior reading skills than boys predicted a decline in boys' reading self-concept whereas girls' reading self-concept was unrelated (Retelsdorf et al., 2015). Regarding math, there is evidence that also elementary school teachers share math-male gender stereotypes (for a review see Li, 1999), that is, they ascribed greater talent and more importance of math to male than female persons (e.g., Fennema et al., 1990; Tiedemann, 2000a,b, 2002; McKown and Weinstein, 2002; Hand et al., 2017). Furthermore, they even revealed a shift in evaluation standards when rating the math performance of a girl compared to a boy's math performance (Holder and Kessels, 2017). Elementary teachers underrated the skills of girls throughout the achievement distribution as early as in Grade 1, when past and current math achievement and behavioral ratings where controlled for (Cimpian et al., 2016). In contrast, only a few studies found no gender differences in teachers' ratings of their students' abilities (e.g., Bennett et al., 1993; Dickhäuser and Stiensmeier-Pelster, 2003) and did not reveal any gender-biased expectations regarding math (Lorenz et al., 2016).

Moreover, teachers (like parents, see e.g., Frome and Eccles, 1998; Herbert and Stipek, 2005; Lazarides and Watt, 2017) seem to transmit stereotyped beliefs (Keller, 2001) to their students supporting or maintaining students' stereotyped belief that math is a "male" domain (e.g., Steffens et al., 2010). The "math-male stereotype" has been consistently found in students in various countries when assessing the stereotyping of domains using implicit measures (like the IAT; Greenwald et al., 1998). Interestingly, the implicit math-male stereotype was evident in children as young as four years old (Cvencek et al., 2011), in elementary school children (Passolunghi et al., 2014), secondary school students (Steffens et al., 2010), and college students (Nosek et al., 2002; Smeding, 2012). However, children at elementary school age did not yet consistently stereotype math as a male domain when expressing their views using explicit measures (e.g., Ambady et al., 2001; Passolunghi et al., 2014). Only later, at adolescence, most students endorsed clear-cut math-male stereotypes and ascribed more talent, ability, and interest in math to boys than to girls (Chatard et al., 2007; Steffens et al., 2010; Steffens and Jelenec, 2011; however compare Passolunghi et al., 2014). The perceived fit between one's own gender and the perceived maleness or femaleness of a domain constitutes a crucial factor for students' willingness to get involved in that domain (Kessels et al., 2014). Accordingly, domain gender stereotypes have been proven important predictors of many domain-related emotions, cognitions, and behaviors. The more female students perceived of math as a male domain, the lower their math achievements, math liking and intentions to follow a respective career were (e.g., Nosek and Smyth, 2011; Steffens and Jelenec, 2011; Lane et al., 2012). According to the expectancy-value model (e.g., Eccles et al., 1983) this view is influenced and shaped by, for example, teachers' beliefs and stereotypes increasing gender gaps in STEM-related ability self-concepts, e.g., mathematical ability self-concepts (e.g., Hand et al., 2017). Indeed, studies with German elementary school samples (Tiedemann, 2000a,b, 2002) showed that the more teachers endorsed gender stereotypes concerning math the greater was their gender-stereotyped view on their students' math abilities. Since teachers' ratings of students' math aptitude predict students' ability self-concepts (e.g., Herbert and Stipek, 2005), this study aims to test whether teachers' gender-stereotyped perception of girls' and boys' math aptitude might explain gender differences in math ability self-concepts.

Brilliance Beliefs and Women's Underrepresentation in STEM

Above and beyond the "traditional" math-male gender stereotype, a different belief has recently gained an increasing amount of interest in research aiming at explaining STEM-related gender differences in the context of academia and higher education in the United States. Originally, it was coined domain-specific ability belief and encompasses the extent to which important agents in the learning environment perceive success in a given domain as requiring an innate ability that cannot be taught ("brilliance") (Leslie et al., 2015). That is, such beliefs can be interpreted as essentially a domain-specific version of what has been called entity theory of intelligence or fixed mindset before, i.e., the belief that intelligence is fixed (Dweck, 2006, 2007; Gunderson et al., 2017). If the beliefs of important learning agents instead of student's field-specific ability beliefs are in focus, the construct is also closely related to the so-called meta-lay theories of a domain-specific aptitude (Rattan et al., 2018). In what follows, we use the term "brilliance belief" to describe the extent to which success in a given domain is perceived as requiring innate ability.

A larger misfit between female students' self and STEM subjects has been attributed in the past to the clashing perceptions that while learning in STEM depends mainly on high ability and less on effort, female academic success is seen as mainly due to effort and hard work (Kessels, 2015). Research in the context of academia has shown that the stronger success in a given domain was perceived as requiring innate ability, the lower the portion of women in this field (Leslie et al., 2015; Meyer et al., 2015). In higher education, STEM subjects were among those domains strongly requiring innate ability as well as being dominated by male persons. In more detail, different disciplines within the STEM field differ in the degree that innate ability is seen as required, and while on average Social Science/Humanities are regarded as requiring less brilliance than STEM, some of those are seen as requiring more or less of innate ability and thus resulting in smaller or larger proportion of female PhDs (Leslie et al., 2015). Also experimental research suggests that presenting job opportunities as requiring innate ability reduces women's interest and belonging but increases their anxiety (Bian et al., 2018). Similarly, perceiving STEM faculty as believing in a special STEM aptitude (which only some students have and others do not) predicted less sense of belonging among female PhD candidates in STEM. Male PhD candidates' sense of belonging, however, was unrelated to their meta-lay theories (Rattan et al., 2018). These findings with adults raise the question whether similar beliefs and processes are already present in earlier stages of math careers, that is, in elementary school, where gender differences in math ability self-concept seem to emerge (Herbert and Stipek, 2005).

Given the alienating consequences of statements that innate ability is required for a domain or task on female students' engagement, teachers' brilliance beliefs seem crucial for understanding the underrepresentation of female students in STEM. As school students have the full range of domains as school subjects, the relative importance ascribed to innate ability in the different school subjects should signal students which subjects to avoid and to approach instead, according to their gender. First studies showed that U.S. American teachers held more fixed views of intelligence for math and science performance compared to performance in languages (Patterson et al., 2016). The construct of brilliance beliefs has so far been studied only once outside the United States, that is in an elementary school teacher sample from Germany (Heyder et al., 2019). It was found that also this group of socializers held stronger brilliance beliefs for math than language arts although their mean level was lower than in prior US studies with academics or lay persons from the general public (see Meyer et al., 2015).The finding that in both countries, that is, Germany and the US, there are domain differences in brilliance beliefs is in line with the fact that both countries are considered Western industrialized countries that with—at least in parts—shared beliefs and stereotypes (Wilde and Diekman, 2005).

As there is also experimental evidence suggesting that in the United States girls from the age of six on view themselves as less smart than boys and refrain from tasks that require to be smart (Bian et al., 2017), an environment that believes that math requires to be smart might discourage girls from math not just at later stages of their career but already at the very beginning. More precisely, such beliefs if held by the teacher might increase the detrimental effects of the math-male gender stereotype on girls, and thus lead to a larger gender gap in math ability self-concept than if teachers held less pronounced math brilliance beliefs. In a German sample, the present study thus aims to explore whether already in elementary school math brilliance beliefs held by the teacher relate to girls' lower ability self-concept in math, despite boys and girls showing similar math competence.

STUDY OVERVIEW

Brilliance beliefs are powerful predictors of women's underrepresentation in certain fields in academia such as STEM domains (e.g., Leslie et al., 2015; Storage et al., 2016). A recent study showed that also German elementary school teachers believed that success in math requires more innate ability than success in language (Heyder et al., 2019). Further, elementary school girls have been found to report a less positive ability self-concept than boys despite their comparable competences (e.g., Herbert and Stipek, 2005; Ganley and Lubienski, 2016) and also teachers perceive boys as more talented in math than girls (e.g., Li, 1999; Tiedemann, 2000a). In this study, we bring these strands of research together. First, we want to explore whether teachers' belief that math requires innate ability predicts a less positive ability self-concept in girls than boys in elementary school. Second, we want to test whether teachers' perceptions of students' math aptitude mediate the gender gap in students' ability self-concept. With teachers' math brilliance beliefs, we thus test a novel facet of students' learning environment in elementary school as an additional factor that might contribute to the beginning of the leaky STEM pipeline for girls and women. More precisely, we test the following hypotheses: (1) Independent from the actual performance girls report lower self-concept of ability in math than boys. (2) Independent from the actual performance teachers perceive girls as less talented in math than boys. (3) Teachers' math brilliance beliefs moderate the effect of gender on teachers' ratings of their students' mathematical aptitude and students' ability self-concept in the expense of girls. That is, we expect larger gender differences in teachers' math aptitude ratings and students' ability self-concept if teachers hold strong brilliance beliefs than if they hold weak brilliance beliefs. (4) Teachers' ratings of their students' mathematical aptitude mediate the effect of gender on students' ability self-concept.

METHODS

Sample

For this study, we re-analyzed the elementary school data set from the FA(IR)BULOUS Study, a German research project on social inequality in school transitions (Steinmayr et al., 2017). The FA(IR)BULOUS data set had already been analyzed in a prior study on the role of teachers' math-specific ability beliefs for low-achieving students' intrinsic motivation in math (Heyder et al., 2019). This study exceeds the former study by focusing on *gender differences* in *ability self-concept* in math and how they are related to teachers' math-specific ability beliefs.

The FA(IR)BULOUS elementary school sample was collected in 2016 and consisted of 837 fourth graders nested in 56 classes. Seven refugee children were excluded from the full sample because they could not fill out the questionnaire due to a lack of German language skills. The remaining 830 student were on average 9.14 years old ($SD = 0.54$), 409 of them were female (49.4%, two students did not indicate their gender), and 667 of them reported German to be their first language (80.4%). More than half of the students reported to have 100 books at home or less (59.6%) which serves as an indicator for students' resources for learning at home or their cultural capital. These descriptive characteristics of the sample suggest that the sample can be considered as similar to the representative German PIRLS 2016 elementary school sample with regard to language spoken at home (83.4% always or nearly always German) and number of books (53.5% 100 books or less at home; Hußmann et al., 2017).

In German elementary schools, each class has one (main) teacher. The 56 teachers of the FA(IR)BULOUS elementary school sample were mostly women (94.6%) and 44.56 years old on average ($SD = 11.48$). The teacher sample corresponds to the total population of primary school teachers in the federal state of Germany in which the study was conducted in terms of gender composition and age (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2016).

Measures

Students' Math Ability Self-Concept

Students' ability self-concept in math was assessed with a four-item short version of the validated German Scales for the Assessment of School-Related Competence Beliefs (SESSKO; Schöne et al., 2002). Each item was rated on a 5-point scale from 1 = *totally disagree* to 5 = *totally agree*. An example item is "In math I know a lot." The reliability of the scale was high (Cronbach's alpha = 0.90).

Students' Math Competencies

Students' math competencies were assessed using the arithmetic subscale of the DEMAT 3+ (Deutscher Mathematiktest für dritte Klassen; Roick et al., 2004), a standardized German math competence test for third graders. The test consists of 15 arithmetic math tasks and had a good reliability (Cronbach's alpha = 0.82). In the following analyses, we used the raw sum score of the test. In addition, we provide the corresponding T-score, i.e., the score on a standardized scale with a $M = 50$ and $SD = 10$.

Teachers' Math Aptitude Ratings

Teachers indicated for each student how talented in math the student is compared with the total student population. Response options were 1 = *far below average (ca. 2% of all students)*, 2 = *below average (ca. 14% of all students)*, 3 = *slightly below average (ca. 15.5% of all students)*, 4 = *average (ca. 37% of all students)*, 5 = *slightly above average (ca. 15.5% of all students)*, 6 = *above average (ca. 14% of all students)*, and 7 = *far above average (ca. 2% of all students)*.

Teachers' Math Brilliance Beliefs

For assessing teachers' math brilliance beliefs, we translated the four items of Leslie et al. (2015) from English to German and adopted them with regard to math and elementary school students. That means, that teachers were asked to indicate to which degree they believe that *children* need an innate ability to be successful in math. An example item is "If you want to succeed in math, hard work alone just won't cut it; for that, children need to have an innate gift or talent". Teachers rated each item on a scale from 1 = *strongly disagree* to 7 = *strongly agree*. Cronbach's alpha was 0.76.

Procedure

This study was carried out in accordance with the recommendations of the German Research Foundation. In compliance with the guidelines established by the institutional ethic committee, participation was voluntary and parents' written informed consent was given before participation in accordance with the Declaration of Helsinki. Teachers were informed about the study during a staff meeting and declared their informed consent verbally. A formal ethics approval was not required by the TU Dortmund guidelines or German regulations. Data was collected during regular class hours by trained research assistants at the beginning of fourth grade. All self-report items were read aloud to make sure that all students understood the items and worked at the same speed.

Analytic Approach

We applied multilevel regression analyses in Mplus 8 (Muthén and Muthén, 1998-2017) with standard errors corrected for non-normality of variables (MLR). All continuous variables were grand-mean centered. Students' gender was coded 0 = *female*, 1 = *male*. Students' first language was coded 0 = *German*, 1 = *other language*.

In order to explore, whether teachers' math brilliance beliefs moderate the size of the gender gap in students' and teachers' aptitude ratings, we specified two cross-level interaction models with gender, teachers' math brilliance beliefs and their interaction terms as predictors of students' ability self-concept or teachers' aptitude rating, respectively. In a second step, we added students' math competencies, mother tongue and number of books as control variables to the model which all have been found to be related with students' motivation in prior research (e.g., Evans et al., 2010; Stanat et al., 2010; Ganley and Lubienski, 2016). Hypothesis 4 was tested in a lower level mediation model with teachers' aptitude ratings as a mediator of the effect of gender on students' ability self-concept.

The percentage of missing data was low. Information on gender was missing for 0.2%, number of books for 0.4%, math test scores for 3.4%, and teachers' aptitude ratings for 2.8% of the student sample. The data on class membership, students' ability self-concept, students' mother tongue, and teachers' math brilliance beliefs were complete. Little's MCAR test with all target variables included suggested missing completely at random, $X^2 = 11.47$, $df = 9$, $p = 0.245$. Thus, the full information maximum likelihood algorithm (FIML) was applied to handle cases with missing values.

RESULTS

Descriptive Statistics

On average, students earned $M = 8.36$ ($SD = 3.48$) out of 15 points in the math competence test. This is equivalent to an average T-score of $T = 47.81$ ($SD = 10.48$) suggesting that, on average, students performed at the expected level. They further reported a positive ability self-concept, $M = 3.64$, $SD = 0.84$.

TABLE 1 | Means, standard deviations, and bivariate correlations.

	M (*SD*)	2)	3)	4)	5)	6)
1) Math competence[a]	8.36 (3.48)	0.42***	0.56***	0.03	0.19***	−0.05
2) Ability self-concept in math	3.64 (0.84)		0.54***	0.24***	0.23***	0.00
3) Teachers' math aptitude ratings	4.25 (1.39)			0.23***	0.25***	−0.05
4) Male[b]	0.51 (0.50)				0.04	−0.01
5) Number of books in home[c]	3.30 (1.18)					−0.21***
6) Nonnative speaker[b]	0.20 (0.40)					

[a]*Test scores ranged from 0 to 15.*
[b]*Mean values here represent the portions of nonnative speakers and male students, respectively, in the sample.*
[c]*Number of books in home ranges from 1 = 0-10 books to 5 = more than 200 books.*
****p < 0.001.*

Teachers' aptitude ratings had a mean of $M = 4.25$ ($SD = 1.39$). Bivariate correlations between the variables were medium to high. All means, standard deviations, and correlations are presented in **Table 1**.

Gender Differences in Math

As in prior studies (e.g., Herbert and Stipek, 2005), only negligible and statistically not significant differences between girls' and boys' scores on the math competence test were found; Girls: $M = 8.26$, $SD = 3.42$, boys: $M = 8.47$, $SD = 3.53$, $t_{(798)} = -0.85$, $p = 0.398$, Cohen's $d = 0.06$. Nonetheless, moderate gender differences in students' ability self-concept and teachers' aptitude ratings were found. Boys reported a more positive ability self-concept than girls, corroborating Hypothesis 1; Girls $M = 3.44$, $SD = 0.83$, boys $M = 3.84$, $SD = 0.80$, $t_{(826)} = -7.15$, $p < 0.001$, $d = 0.50$. Furthermore, also teachers rated boys as having more math talent than girls, supporting Hypothesis 2; Girls $M = 3.03$, $SD = 1.39$, boys $M = 4.56$, $SD = 1.33$, $t_{(803)} = -6.55$, $p < 0.001$, $d = 0.46$.

Moderation by Teachers' Math Brilliance Beliefs

As reported in a previous study (Heyder et al., 2019), teachers on average perceived success in math as requiring a moderate amount of innate ability, $M = 4.17$, $SD = 1.12$. In order to explore whether this belief relates to the gender gap in students' ability self-concept and teachers' aptitude ratings, we ran two series of cross-level interaction models including different control variables (see **Tables 2**, **3**).

As presented in **Table 2**, girls reported a less positive ability self-concept than boys. Teachers' math brilliance beliefs, however, did not predict students' ability self-concept or the gender gap in students' ability self-concept. Furthermore, there was no significant variation in the relation between gender and students' ability self-concept between classes. Students' math competencies and number of books at home predicted higher scores in students' ability self-concept. Controlling for these variables as well as students' mother tongue only slightly reduced the gender gap in students' ability self-concept. The non-significance of the cross-level interaction remained unaffected.

TABLE 2 | Regression of students' ability self-concept in math on students' gender and teachers' math brilliance beliefs.

	1	2	3
Intercept	3.43 (0.04)***	3.45 (0.04)***	3.43 (0.05)***
Predictors			
Male (L1)	0.41 (0.05)***	0.38 (0.05)***	0.37 (0.05)***
Teachers' beliefs (L2)	0.04 (0.03)	0.04 (0.03)	0.04 (0.03)
Teachers' beliefs (L2) x male (L1)	−0.04 (0.05)	−0.04 (0.05)	−0.03 (0.04)
Math competence (L1)		0.11 (0.01)***	0.10 (0.01)***
Number of books (L1)			0.12 (0.02)***
Nonnative speaker (L1)			0.13 (0.07) #
Variance components			
Residual variance L1	0.64 (0.04)***	0.51 (0.03)***	0.49 (0.03)***
Residual variance slope male	0.01 (0.05)	0.01 (0.03)	0.01 (0.03)
Residual variance L2	0.02 (0.03)	0.02 (0.02)	0.02 (0.02)
Covariance intercept-slope L2	< 0.01 (0.03)	0.01 (0.02)	0.01 (0.02)

Cross-level interaction models. Unstandardized regression coefficients with standard errors in parentheses. Students' math competence test scores, number of books, and teachers' math brilliance beliefs were grand-mean-centered.
*# p < 0.10 ***p < 0.001.*

TABLE 3 | Regression of teachers' aptitude ratings on students' gender and teachers' math brilliance beliefs.

	1	2	3
Intercept	3.93 (0.07)***	3.97 (0.07)***	3.96 (0.07)***
Predictors			
Male (L1)	0.63 (0.07)***	0.54 (0.08)***	0.54 (0.08)***
Teachers' beliefs (L2)	0.08 (0.05)	0.07 (0.05)	0.08 (0.06)
Teachers' beliefs (L2) x male (L1)	−0.07 (0.05)	−0.07 (0.05)	−0.08 (0.05)
Math competence (L1)		0.24 (0.01)***	0.23 (0.01)***
Number of books (L1)			0.18 (0.04)***
Non-native speaker (L1)			0.12 (0.10)
Variance components			
Residual variance L1	1.74 (0.11)***	1.07 (0.07)***	1.04 (0.07)***
Residual variance slope male	0.02 (0.05)	0.03 (0.08)	0.03 (0.07)
Residual variance L2	0.07 (0.06)	0.12 (0.06)*	0.12 (0.06)*
Covariance intercept-slope L2	0.03 (0.03)	0.05 (0.05)	0.05 (0.05)

Cross-level interaction models. Unstandardized regression coefficients with standard errors in parentheses. Students' math competence test scores, number of books, and teachers' math brilliance beliefs were grand-mean-centered.
**p < 0.05 ***p < 0.001.*

Table 3 shows the results for the regression analysis of teachers' aptitude ratings. Here again, being male predicted higher scores than being female. However, this gender difference in teachers' aptitude ratings was not related to teachers' belief whether math requires innate ability. That is, the cross-level interaction did not reach statistical significance. Moreover, there was no significant variation in the relation between gender and teachers' math aptitude ratings between classes. With regard to the control variables, students with higher math competences and students who reported a higher number of books at home were rated by the teachers as having higher math aptitude than other students. Controlling for students' math competence, number of books at home and mother tongue did not substantially change the results. Hypothesis 3 was thus not supported by the data.

Mediation by Teachers' Math Aptitude Ratings

The lack of significant variation in the relations between gender and (a) teachers' aptitude ratings, and (b) students' ability self-concept at Level 2 as well as the non-significant cross-level interaction effects suggested that teachers' brilliance beliefs do not moderate the relations in focus. Striving for a parsimonious model in order to ensure model convergence, we thus tested Hypothesis 4 in a lower level mediation model assuming all effects to be fixed. We did not control for actual math performance as we did not find any significant gender differences in the mathematical competency test. As presented in **Table 4**, all relations specified at Level 1 were statistically significant. Teachers rated boys' math aptitude as higher than girls, and teachers' aptitude ratings in turn predicted students' ability self-concept.

In support of our hypothesis, teachers' aptitude ratings mediated the effect of gender on students' ability self-concept within classes (*indirect effect within* $= 0.20$, $SE = 0.03$, $p < 0.001$). The results indicate that approximately half of the gender gap in students' ability self-concepts was statistically explained by teachers' aptitude ratings. No significant indirect effect was found at Level 2 (*indirect effect between* $= -0.07$, $SE = 0.55$, $p = 0.905$).

TABLE 4 | Mediation of the effect of gender on students' math ability self-concept via teachers' aptitude ratings.

	Coefficient (standard error)
Intercepts	
Teachers' Ratings	−0.65 (7.14)
Students' Ability Self-Concept	4.67 (2.94)
Level 1	
Male ➔Teachers' Ratings	0.62 (0.08)***
Teachers' Ratings ➔Students' Ability Self-Concept	0.33 (0.02)***
Male ➔Students' Ability Self-Concept	0.21 (0.05)***
Residual variance Teachers' Ratings	1.75 (0.11)***
Residual variance Students' Ability Self-Concept	0.46 (0.03)***
Level 2	
Male ➔Teachers' Ratings	1.28 (14.11)
Teachers' Ratings ➔Students' Ability Self-Concept	−0.05 (0.22)
Male ➔Students' Ability Self-Concept	−2.03 (5.77)
Residual variance Teachers' Ratings	0.10 (0.04)*
Residual variance Students' Ability Self-Concept	0.02 (0.02)

Lower level mediation model with fixed slopes. Unstandardized regression coefficients with standard errors in parentheses. Teachers' aptitude ratings were grand-mean-centered.
*$p < 0.05$ ***$p < 0.001$.

DISCUSSION

Why do women end up pursuing less math-intensive careers than men? In this study, we aimed to explain gender differences in a powerful and early predictor of math-related achievement and career choices, that is elementary school students' ability self-concept (e.g., Eccles, 2011; Musu-Gillette et al., 2015). We focused on the role that teachers' beliefs play in gender differences in math ability self-concepts. More precisely, we were interested in exploring whether teachers' ascription of higher math talent to boys compared to girls might contribute to girls' lower ability self-concept in math. In addition, we tested whether teachers' belief that success in math requires an innate ability might be detrimental for girls' but not boys' self-concept of ability. Our analyses were based on an elementary school sample of teachers and fourth graders from Germany. In line with prior meta-analyses (e.g., Else-Quest et al., 2010; Reilly et al., 2015), only negligible gender differences in a standardized math competence test were found. Nonetheless, boys already reported a more positive ability self-concept in math than girls, supporting Hypothesis 1. This finding is in line with findings from earlier studies with elementary students (Tiedemann, 2000b; Fredricks and Eccles, 2002; Herbert and Stipek, 2005; Ganley and Lubienski, 2016; Gentrup and Rjosk, 2018). While male and female students did not differ in their standardized test performance, teachers described their male students as more talented in math than their female students, corroborating Hypothesis 2 and validating earlier studies on teachers' gender-bias in math (e.g., Li, 1999; Cimpian et al., 2016; Hand et al., 2017; Holder and Kessels, 2017; Gentrup and Rjosk, 2018). As expected in Hypotheses 4, these gender-biased aptitude ratings of the teachers proved to account for half of the gender gap in math ability self-concepts. Students at the end of elementary school seem to have internalized their teachers' gender bias in talent ascription in math, with the result that girls perceive their talent for math to be lower than boys. This is an alarming finding, given the importance of math ability self-concepts for future achievement and choices in STEM subjects (e.g., Wang and Degol, 2013; Musu-Gillette et al., 2015).

Based on the literature suggesting detrimental effects of brilliance beliefs for female students' aspirations in STEM (e.g., Leslie et al., 2015; Rattan et al., 2018), we further tested if teachers' brilliance beliefs had a more negative relation with girls' self-concept of ability in math than with boys'. However, our analysis found both boys' and girls' math ability self-concept to be unrelated to their teachers' brilliance beliefs, thus contradicting Hypothesis 3. Whereas such socially shared brilliance beliefs have been found to be related to female students' underrepresentation in STEM careers in higher education (e.g., Leslie et al., 2015; Rattan et al., 2018), our results suggest that teachers' brilliance beliefs may not play a role in explaining the gender gap in math ability self-concept in elementary school.

Several reasons might account for this. First, it is possible that teachers did not transfer their beliefs to the children, so students were not aware of these beliefs, which in turn could not influence their motivation. For instance, teachers might not communicate their beliefs to their students or students might not be able to decipher any messages conveying the respective beliefs (directly or via teaching practices). However, our finding that the students' math ability self-concept was related to the teacher's aptitude rating supports the validity of the expectancy-value model (e.g., Eccles et al., 1983). Moreover, research on teacher expectations has shown many times (e.g., Jussim et al., 2009) that such positive or negative assumptions of a student's (mathematical) abilities can be conveyed in the classroom by the actions of the teacher. In contrast, any deterrent effect of a teacher's brilliance belief on a specific student's ability self-concept seems to be more complicated or subtle and has to be conveyed via several, more abstract intermediate steps than the message "I think you are (not) good at math."

Another possible explanation for the lacking interaction effect could be found in the understanding of the construct "math." More precisely, elementary teachers seem to assume that innate ability is needed for children to succeed in "math" (in general), but they might not apply this belief to basic math in elementary school, considering basic math still as relatively easy and accessible not only to exceptional children. An earlier study with teachers from a variety of grade levels did ask separately about domain-specific beliefs toward advanced and basic math, but combined these two in their analyses (Patterson et al., 2016). Our study raises the question to be studied in future research if advanced math might elicit stronger brilliance beliefs than basic math.

Generally, the young age of the students in our study might be another factor explaining the independency of their math ability self-concept from their teachers' brilliance beliefs. Most importantly, younger children do not consistently differentiate the concepts of ability and of effort (Nicholls, 1990), what implies that messages of required innate ability might not result in the conclusion that effort would not help in order to succeed. Accordingly, at earlier ages, teachers' brilliance beliefs might not execute their full range of negative implications for girls compared to later ages. In addition, earlier research showed that even though first and second graders believed that success in an adult job requires more fixed ability in math than reading and writing, the children did not think that their own grades in math were depending more on fixed ability than their reading and writing grades (Gunderson et al., 2017), thus not yet applying the stereotypes on their own achievements.

And while in the U.S., children as young as six were found to endorse the stereotype that brilliant children are male (Bian et al., 2017), evidence from European countries is scarce. Most European studies with school students on the stereotyping of genius and "effortless achievement" as something male and masculine (Jackson, 2003; Jackson and Dempster, 2009; Heyder and Kessels, 2017) have so far concentrated on adolescents. State of research regarding the gender stereotyping of math in younger children is inconclusive, as some studies found implicit gender stereotyping at an early age (Cvencek et al., 2011), but others, using explicit measures, could not find these (Ambady et al., 2001). Research on the perception of adult stereotypes further indicates that Italian elementary school children thought that teachers viewed boys and girls similarly in math (Muzzatti and Agnoli, 2007) and that French fourth graders of each gender reported that people view their own gender as better in math (Martinot and Désert, 2007). A study from the U.S. however found that male, but not female, fourth graders perceived that adults believed that boys are better at math and science than girls (Kurtz-Costes et al., 2008). Taken together, the existing research on elementary school students and teachers seems to find more and stronger math brilliance beliefs and math-male stereotypes in the U.S. than in Europe. Thus, the role not only of students' age but also of the cultural background on math-related beliefs should be studied more explicitly in comparative studies in the future.

In this vein, this study is limited as it focused only on a German sample. Another limitation is its cross-sectional nature. Even if earlier research has shown that socializers' ability beliefs longitudinally predict children's self-concept of ability (Frome and Eccles, 1998; Lazarides and Watt, 2017), we cannot rule out that other mechanisms are behind the relationship between teachers' aptitude perceptions and children's ability self-concept found in our data. For instance, it is possible that a more positive ability self-concept signals to the teacher that this student is talented, while a negative ability self-concept might be interpreted by the teacher as a lack of talent. Thus, any implications for practice should be drawn and interpreted very cautiously. This study's finding that the gender gap in teachers' aptitude ratings was larger than in students' actual math competencies corroborates once more the prevalence of math-male stereotypes in teachers (see also e.g., Li, 1999; Tiedemann, 2002; Hand et al., 2017). Increasing teachers' awareness of their own math-male stereotypes and confronting them with the fact that average gender differences in math competencies are actually very small could be one approach to reduce teachers' math-male stereotypes leading to more accurate perceptions of boys' and girls' math aptitude. This could happen both during teacher education in universities as well as in professional development courses for in-service teachers. Reducing a potential gender bias in elementary school teachers' aptitude perceptions seems also a fruitful means for increasing girls' participation in STEM particularly in Germany because in Germany teachers give tracking recommendations at the end of elementary school, and these have also been found to be predicted by teachers' gender stereotypes (Nürnberger et al., 2016).

Above and beyond testing the research questions derived above from this study's results, future research should study in more detail—and longitudinally—how domain specific brilliance beliefs develop in children. It would be interesting to understand how the emergence of brilliance beliefs in children relates to their ability self-concept as well as to both the emergence of a compensatory concept of effort and ability and to the emergence of a general view on intelligence as fixed. Teachers' entity theories can be both comforting and demotivating for their students (Rattan et al., 2012), and from early adolescence on, girls report lower incremental views on intelligence than boys (Diseth et al., 2014). As the belief that success in a given domain requires innate ability goes along with an underrepresentation of female persons, it is most crucial to understand from what age on children actually infer from their socializers' brilliance beliefs that their own striving for success in math may be useless, and if girls might gain this understanding even earlier than boys might.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the German Research Foundation. In compliance with the guidelines established by the institutional ethic committee, participation was voluntary and parents' written informed consent was given before participation in accordance with the Declaration of Helsinki. An institutional approval of the protocol by the university ethics board is not typically required for projects of this type.

AUTHOR CONTRIBUTIONS

AH and RS contributed conception and design of the study. AH provided a first draft of the manuscript and finalized the paper. RS and UK wrote sections of the manuscript and gave valuable feedback on the first draft.

FUNDING

The research reported in this article is based on the project FA(IR)BULOUS – FAIRe BeUrteilung des LeistungspOtenzials Unterprivilegierter Schülerinnen und Schüler; Fair Evaluation of underprivileged students' potential. The project was funded by the Stiftung Mercator (Foundation Mercator). Furthermore, we acknowledge financial support by Deutsche Forschungsgemeinschaft and Technische Universität Dortmund/TU Dortmund Technical University within the funding programme Open Access Publishing.

REFERENCES

Ambady, N., Shih, M., Kim, A., and Pittinsky, T. L. (2001). Stereotype susceptibility in children: effects of identity activation on quantitative performance. *Psychol. Sci.* 12, 385–390. doi: 10.1111/1467-9280.00371

Bennett, R. E., Gottesman, R. L., Rock, D. A., and Cerullo, F. (1993). Influence of behavior perceptions and gender on teachers' judgments of students' academic skill. *J. Educ. Psychol.* 85, 347–356.

Bergold, S., Wendt, H., Kasper, D., and Steinmayr, R. (2017). Academic competencies: their interrelatedness and gender differences at their high end. *J. Educ. Psychol.* 109, 439–449. doi: 10.1037/edu0000140

Bian, L., Leslie, S.-J., and Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science* 355, 389–391. doi: 10.1126/science.aah6524

Bian, L., Leslie, S.-J., Murphy, M. C., and Cimpian, A. (2018). Messages about brilliance undermine women's interest in educational and professional opportunities. *J. Exp. Soc. Psychol.* 18, 404–420. doi: 10.1016/j.jesp.2017.11.006

Brunner, M., Krauss, S., and Kunter, M. (2008). Gender differences in mathematics: does the story need to be rewritten? *Intelligence* 36, 403–421. doi: 10.1016/j.intell.2007.11.002

Chatard, A., Guimond, S., and Selimbegovic, L. (2007). "How good are you in math?" the effect of gender stereotypes on students' recollection of their school marks. *J. Exp. Soc. Psychol.* 43, 1017–1024. doi: 10.1016/j.jesp.2006.10.024

Chestnut, E., Lei, R., Leslie, S.-J., and Cimpian, A. (2018). The myth that only brilliant people are good at math and its implications for diversity. *Educ. Sci.* 8:65. doi: 10.3390/educsci8020065

Cimpian, J. R., Lubienski, S. T., Timmer, J. D., Makowski, M. B., and Miller, E. K. (2016). Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts. *AERA Open* 2, 1–19. doi: 10.1177/2332858416673617

Cvencek, D., Kapur, M., and Meltzoff, A. N. (2015). Math achievement, stereotypes, and math self-concepts among elementary-school students in Singapore. *Learn. Instruct.* 39, 1–10. doi: 10.1016/j.learninstruc.2015.04.002

Cvencek, D., Meltzoff, A. N., and Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x

Dickhäuser, O., and Stiensmeier-Pelster, J. (2003). Wahrgenommene Lehrereinschätzungen und das Fähigkeitsselbstkonzept von Jungen und Mädchen in der Grundschule. [Perceptions of teacher evaluations and the academic self-concept of boys and girls in elementary school]. *Psychologie in Erziehung und Unterricht* 50, 182–190.

Diseth, Å., Meland, E., and Breidablik, H. J. (2014). Self-beliefs among students: Grade level and gender differences in self-esteem, self-efficacy and implicit theories of intelligence. *Learn. Individ. Differ.* 35, 1–8. doi: 10.1016/j.lindif.2014.06.003

Dweck, C. S. (2006). *Mindset: The New Psychology of Success*. New York, NY: Ballantine Books.

Dweck, C. S. (2007). "Is math a gift? beliefs that put females at risk," in *Why aren't More Women in Science?: Top Researchers Debate the Evidence,* eds S. J. Ceci and W. M. Williams (Washington, DC: American Psychological Association), 47–55.

Eccles, J. S. (2011). Gendered educational and occupational choices: applying the Eccles et al. model of achievement-related choices. *Int. J. Behav. Dev.* 35, 195–201. doi: 10.1177/0165025411398185

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). "Expectancies, values, and academic behaviors,". in *A Series of Books in Psychology. Achievement and Achievement Motives. Psychological and Sociological Approaches* eds J. T. Spence (San Francisco, CA: W.H. Freeman), 76–146.

Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Evans, M. D. R., Kelley, J., Sikora, J., and Treiman, D. J. (2010). Family scholarly culture and educational success: books and schooling in 27 nations. *Res. Soc. Stratif. Mobil.* 28, 171–197. doi: 10.1016/j.rssm.2010.01.002

Fennema, E., Peterson, P. L., Carpenter, T. P., and Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educ. Stud. Math.* 21, 55–69. doi: 10.1007/BF00311015

Fredricks, J. A., and Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: growth trajectories in two male-sex-typed domains. *Dev. Psychol.* 38, 519–533. doi: 10.1037/0012-1649.38.4.519

Frome, P. M., and Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *J. Pers. Soc. Psychol.* 74:435. doi: 10.1037/0022-3514.74.2.435

Ganley, C. M., and Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: examining gender patterns and reciprocal relations. *Learn. Individ. Differ.* 47, 182–193. doi: 10.1016/j.lindif.2016.01.002

Gentrup, S., and Rjosk, C. (2018). Pygmalion and the gender gap: Do teacher expectations contribute to differences in achievement between boys and girls at the beginning of schooling? *Edu. Res. Eval.* 18, 1–29. doi: 10.1080/13803611.2018.1550840

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Gunderson, E. A., Hamdan, N., Sorhagen, N. S., and D'Esterre, A. P. (2017). Who needs innate ability to succeed in math and literacy? academic-domain-specific theories of intelligence about peers versus adults. *Dev. Psychol.* 53, 1188–1205. doi: 10.1037/dev0000282

Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles* 66, 153–166. doi: 10.1007/s11199-011-9996-2

Hand, S., Rice, L., and Greenlee, E. (2017). Exploring teachers' and students' gender role bias and students' confidence in STEM fields. *Soc. Psychol. Educ.* 20, 929–945. doi: 10.1007/s11218-017-9408-8

Herbert, J., and Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *J. Appl. Dev. Psychol.* 26, 276–295. doi: 10.1016/j.appdev.2005.02.007

Heyder, A., and Kessels, U. (2017). Boys don't work? On the psychological benefits of showing low effort in high school. *Sex Roles* 77, 72–85. doi: 10.1007/s11199-016-0683-1

Heyder, A., Weidinger, A. F., Cimpian, A., and Steinmayr, R. (2019). Teachers' belief that math requires innate ability predicts lower intrinsic motivation among low-achieving students. Manuscript submitted for publication.

Holder, K., and Kessels, U. (2017). Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective. *Soc. Psychol. Educ.* 20, 471–490. doi: 10.1007/s11218-017-9384-z

Hußmann, A., Wendt, H., Bos, W., Bremerich-Vos, A., Kasper, D., and Lankes, E.-M. (eds.) (2017). *IGLU 2016: Lesekompetenzen von Grundschulkindern in Deutschland im Internationalen Vergleich [Elementary School Students' Reading Competencies in Germany in International Comparision]*. Münster: Waxmann.

Hyde, J. S. (2005). The gender similarities hypothesis. *Am. Psychol.* 60, 581–592. doi: 10.1037/0003-066X.60.6.581

Jackson, C. (2003). Motives for 'laddishness' at school: fear of failure and fear of the 'feminine'. *Br. Educ. Res. J.* 29, 583–598. doi: 10.1080/01411920301847

Jackson, C., and Dempster, S. (2009). 'I sat back on my computer … with a bottle of whisky next to me': constructing 'cool' masculinity through 'effortless' achievement in secondary and higher education. *J. Gender Stud.* 18, 341–356. doi: 10.1080/09589230903260019

Jussim, L., Robustelli, S., and Cain, T. (2009). "Teacher expectations and self-fulfilling prophecies," in *Handbook of motivation at school,* eds A. Wigfield and K. Wentzel (Mahwah, NJ: Erlbaum), 349–380.

Keller, C. (2001). Effect of teachers' stereotyping on students' stereotyping of mathematics as a male domain. *J. Soc. Psychol.* 141, 165–173. doi: 10.1080/00224540109600544

Kessels, U. (2015). Bridging the gap by enhancing the fit: How stereotypes about STEM clash with stereotypes about girls. *Int. J. Gender Sci. Technol.* 7, 280–296.

Kessels, U., Heyder, A., Latsch, M., and Hannover, B. (2014). How gender differences in academic engagement relate to students' gender identity. *Educ. Res.* 56, 219–228. doi: 10.1080/00131881.2014.898916

Kurtz-Costes, B., Rowley, S. J., Harris-Britt, A., and Woods, T. A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *Merrill Palmer Q.* 54, 386–409. doi: 10.1353/mpq.0.0001

Lane, K. A., Goh, J. X., and Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles* 66, 220–234. doi: 10.1007/s11199-011-0036-z

Lazarides, R., and Watt, H. M. (2017). Student-perceived mothers' and fathers' beliefs, mathematics and english motivations, and career choices. *J. Res. Adolesc.* 27, 826–841. doi: 10.1111/jora.12317

Leslie, S.-J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 262–265. doi: 10.1126/science.1261375

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: a review. *Educ. Res.* 41, 63–76. doi: 10.1080/0013188990410106

Lorenz, G., Gentrup, S., Kristen, C., Stanat, P., and Kogan, I. (2016). Stereotype bei Lehrkräften? eine untersuchung systematisch verzerrter lehrererwartungen. [stereotypes among teachers? a study of systematic bias in teacher expectations]. *Kölner Z. Soziol. Sozialpsychol.* 68, 89–111. doi: 10.1007/s11577-015-0352-3

Martinot, D., and Désert, M. (2007). Awareness of a gender stereotype, personal beliefs and self-perceptions regarding math ability: when boys do not surpass girls. *Soc. Psychol. Educ.* 10, 455–471. doi: 10.1007/s11218-007-9028-9

McKown, C., and Weinstein, R. S. (2002). Modeling the role of child ethnicity and gender in children's differential response to teacher expectations. *J. Appl. Soc. Psychol.* 32, 159–184. doi: 10.1111/j.1559-1816.2002.tb01425.x

Meyer, M., Cimpian, A., and Leslie, S. J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Front. Psychol.* 6:235. doi: 10.3389/fpsyg.2015.00235

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2016). *Das Schulwesen in Nordrhein-Westfalen aus Quantitativer Sicht* 2015/16 *[A Quantitative Perspective on Schools in North Rhine-Westphalia 2015/16]*. Available online at: https://www.schulministerium.nrw.de/docs/bp/Ministerium/Service/Schulstatistik/Amtliche-Schuldaten/Quantita_2015.pdf (accessed January 7, 2019).

Mullis, I. V. S., Martin, M. O., Foy, P., and Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. Boston, MA: TIMSS & PIRLS International Study Center.

Musu-Gillette, L., Wigfield, A., Harring, J., and Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educ. Res. Eval.* 21, 343–370. doi: 10.1080/13803611.2015.1057161

Muthén, L. K., and Muthén, B. O. (1998-2017). *Mplus User's Guide*, *8th Edn*. Los Angeles, CA: Muthén & Muthén.

Muzzatti, B., and Agnoli, F. (2007). Gender and mathematics: attitudes and stereotype threat susceptibility in Italian children. *Dev. Psychol.* 43, 747–759. doi: 10.1037/0012-1649.43.3.747

Nicholls, J. G. (1990). "What is ability and why are we mindful of it? A developmental perspective," in *Competence Considered,* eds R. J. Sternberg and J. Kolligian (New Haven, CT: Yale University Press), 11–40.

Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *J. Pers. Soc. Psychol.* 83, 44–59. doi: 10.1037/0022-3514.83.1.44

Nosek, B. A., and Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *Am. Educ. Res. J.* 48, 1125–1156. doi: 10.3102/0002831211410683

Nürnberger, M., Nerb, J., Schmitz, F., Keller, J., and Sütterlin, S. (2016). Implicit gender stereotypes and essentialist beliefs predict preservice teachers' tracking recommendations. *J. Exp. Educ.* 84, 152–174. doi: 10.1080/00220973.2015.1027807

Passolunghi, M. C., Rueda Ferreira, T. I., and Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learn. Individ. Differ.* 34, 70–76. doi: 10.1016/j.lindif.2014.05.005

Patterson, M. M., Kravchenko, N., Chen-Bouck, L., and Kelley, J. A. (2016). General and domain-specific beliefs about intelligence, ability, and effort among preservice and practicing teachers. *Teach. Teach. Educ.* 59, 180–190. doi: 10.1016/j.tate.2016.06.004

Rattan, A., Good, C., and Dweck, C. S. (2012). "It's ok — Not everyone can be good at math": instructors with an entity theory comfort (and demotivate) students. *J. Exp. Soc. Psychol.* 48, 731–737. doi: 10.1016/j.jesp.2011.12.012

Rattan, A., Savani, K., Komarraju, M., Morrison, M. M., Boggs, C., and Ambady, N. (2018). Meta-lay theories of scientific potential drive underrepresented students' sense of belonging to science, technology, engineering, and mathematics (STEM). *J. Pers. Soc. Psychol.* 115, 54–75. doi: 10.1037/pspi0000130

Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. *PLoS ONE* 7:e39904. doi: 10.1371/journal.pone.0039904

Reilly, D., Neumann, D. L., and Andrews, G. (2015). Sex differences in mathematics and science achievement: a meta-analysis of national assessment of educational progress assessments. *J. Educ. Psychol.* 107, 645–662. doi: 10.1037/edu0000012

Retelsdorf, J., Schwartz, K., and Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *J. Educ. Psychol.* 107, 186–194. doi: 10.1037/a0037107

Riegle-Crumb, C., King, B., Grodsky, E., and Muller, C. (2012). The more things change, the more they stay the same? prior achievement fails to explain gender inequality in entry into STEM college majors over time. *Am. Educ. Res. J.* 49, 1048–1073. doi: 10.3102/0002831211435229

Roick, T., Gölitz, D., and Hasselhorn, M. (2004). *Deutscher Mathematiktest für dritte Klassen DEMAT 3+ [German math test for third graders]*. Göttingen: Hogrefe Publishing.

Schöne, C., Dickhäuser, O., Spinath, B., and Stiensmeier-Pelster, J. (2002). *SESSKO Skalen zur Erfassung des schulischen Selbstkonzept [Scales for the Assessment of School-Related Competence Beliefs]*. Göttingen: Hogrefe.

Simpkins, S. D., Fredricks, J. A., and Eccles, J. S. (2015). The role of parents in the ontogeny of achievement-related motivation and behavioral choices: IV. child factors and parent belief models. *Monogr. Soc. Res. Child Dev.* 80, 65–84. doi: 10.1111/mono.12160

Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): an investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles* 67, 617–629. doi: 10.1007/s11199-012-0209-4

Stanat, P., Segeritz, M., and Christensen, G. (2010). "Schulbezogene Motivation und Aspiration von Schülerinnen und Schülern mit Migrationshintergrund [School-related motivation and aspiration of students with migration background]," in *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert,* eds W. Bos, E. Klieme, and O. Köller (Münster: Waxmann.), 31–57.

Steffens, M. C., and Jelenec, P. (2011). Separating implicit gender stereotypes regarding math and language: Implicit ability stereotypes are self-serving for boys and men, but not for girls and women. *Sex Roles* 64, 324–335. doi: 10.1007/s11199-010-9924-x

Steffens, M. C., Jelenec, P., and Noack, P. (2010). On the leaky math pipeline: comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *J. Educ. Psychol.* 102, 947–963. doi: 10.1037/a0019920

Steinmayr, R., Michels, J., and Weidinger, A. (2017). *Fa(ir)bulous: Faire Beurteilungen des Leistungspotenzials von Schülerinnen und Schülern [Fa(ir)bulous- Fair Judgments of Students' Academic Potential]*. Dortmund: Technische Universität Dortmund.

Storage, D., Horne, Z., Cimpian, A., and Leslie, S.-J. (2016). The frequency of "brilliant" and "genius" in teaching evaluations predicts the representation of women and African Americans across fields. *PLoS ONE* 11:e0150194. doi: 10.1371/journal.pone.0150194

Tiedemann, J. (2000a). Gender-related beliefs of teachers in elementary school mathematics. *Educ. Stud. Math.* 41, 191–207. doi: 10.1023/A:1003953801526

Tiedemann, J. (2000b). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *J. Educ. Psychol.* 92, 144–151. doi: 10.1037/0022-0663.92.1.144

Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educ. Stud. Math.* 50, 49–62. doi: 10.1023/A:1020518104346

Tomasetto, C., Mirisola, A., Galdi, S., and Cadinu, M. (2015). Parents' math–gender stereotypes, children's self-perception of ability, and children's appraisal of parents' evaluations in 6-year-olds. *Contemp. Educ. Psychol.* 42, 186–198. doi: 10.1016/j.cedpsych.2015.06.007

Wang, M.-T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM Fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001

Wang, M.-T., and Degol, J. L. (2016). Gender gap in science, technology, engineering, and mathematics (STEM): current knowledge, implications for practice, policy, and future directions. *Educ. Psychol. Rev*. 29, 119–140. doi: 10.1007/s10648-015-9355-x

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., et al. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: a 3-year study. *J. Educ. Psychol.* 89, 451–469. doi: 10.1037/0022-0663.89.3.451

Wilde, A., and Diekman, A. B. (2005). Cross-cultural similarities and differences in dynamic stereotypes: a comparison between Germany and the United States. *Psychol. Women Q.* 29, 188–196. doi: 10.1111/j.1471-6402.2005.00181.x

Wolter, I., Braun, E., and Hannover, B. (2015). Reading is for girls!? the negative impact of preschool teachers' traditional gender role attitudes on boys' reading related motivation and skills. *Front. Psychol.* 6:1267. doi: 10.3389/fpsyg.2015.01267

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 28 May 2019
doi: 10.3389/fpsyg.2019.01114

Beliefs in "Brilliance" and Belonging Uncertainty in Male and Female STEM Students

Anne Deiglmayr[1,2], Elsbeth Stern[1] and Renate Schubert[1]*

[1] *ETH Zürich, Department of Humanities, Social and Political Sciences, Zurich, Switzerland,* [2] *Faculty of Education, University of Leipzig, Leipzig, Germany*

A wide-spread stereotype that influences women's paths into STEM (or non-STEM) fields is the implicit association of science and mathematics with "male" and with requiring high levels of male-associated "brilliance." Recent research on such "field-specific ability beliefs" has shown that a high emphasis on brilliance in a specific field goes along with a low share of female students among its graduates. A possible mediating mechanisms between cultural expectations and stereotypes on the one hand, and women's underrepresentation in math-intensive STEM fields on the other hand, is that women may be more likely than men to feel that they do not belong in these fields. In the present study, we investigated field-specific ability beliefs as well as belonging uncertainty in a sample of n = 1294 male and female university students from five STEM fields (Mathematics, Physics, Computer Science, Electrical Engineering, and Mechanical Engineering) at a prestigious technical university in Switzerland. Field-specific ability beliefs of both men and women emphasized brilliance more in more math-intensive fields (Mathematics, Physics) than in less math-intensive fields (Engineering). Women showed higher beliefs in brilliance than men did, and also reported higher levels of belonging uncertainty. For both genders, there was a small, positive correlation (r = 0.19) of belief in brilliance and belonging uncertainty. A relatively small, but significant portion of the effect of gender on belonging uncertainty was mediated by women's higher belief in brilliance.

Keywords: field-specific ability beliefs, belonging uncertainty, STEM gender gap, gender stereotypes, university students

OPEN ACCESS

Edited by:
Bernhard Ertl,
Universität der Bundeswehr München, Germany

Reviewed by:
Grzegorz Sedek,
SWPS University of Social Sciences and Humanities, Poland
Hanna Gaspard,
University of Tübingen, Germany

***Correspondence:**
Anne Deiglmayr
anne.deiglmayr@uni-leipzig.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *11 December 2018*
Accepted: *29 April 2019*
Published: *28 May 2019*

Citation:
Deiglmayr A, Stern E and Schubert R (2019) Beliefs in "Brilliance" and Belonging Uncertainty in Male and Female STEM Students. Front. Psychol. 10:1114. doi: 10.3389/fpsyg.2019.01114

INTRODUCTION

Although the gender-gap in achievement in STEM fields has narrowed down in recent years, women remain underrepresented in many math-intensive fields (Ceci et al., 2014; Wang and Degol, 2017). The dimension of this gender gap and possible explanations for its sustained existence have been analyzed from many perspectives, and based on large data sets, in recent years (for overviews, see for example: Ceci et al., 2014; Miller et al., 2015; Cheryan et al., 2017; Wang and Degol, 2017; Stoet and Geary, 2018). Analysts generally agree that the underrepresentation of women in math-intensive STEM fields results from the interplay of multiple factors. Biological factors and differences in basic cognitive abilities may contribute to the phenomenon, but cannot explain the substantial cross-cultural and historic variability in gender inequality in entry into STEM (Berkowitz et al., in press; Wang and Degol, 2017; Stoet and Geary, 2018). There are indications that women face implicit negative biases when decision-makers judge their abilities and performance in math-intensive STEM fields, for instance when teachers grade girls (Hofer, 2015) or when faculty members rate applicants (Moss-Racusin et al., 2012). To a large extent, however, the underrepresentation of women in STEM fields seems to reflect choices that girls and women make themselves, e.g., by choosing hobbies, academic specializations, study subjects, or career paths leading them into less math-intensive or non-STEM fields (Ceci et al., 2014). Of course, even though such choices seem to be free at the first glance, they are constrained by cultural expectations and stereotypes that associate science and mathematics with stereotypically male, rather than stereotypically female traits (e.g., Thébaud and Charles, 2018).

A wide-spread stereotype that influences women's paths into STEM (or non-STEM) fields is the implicit association of science and mathematics with "male" traits (Nosek et al., 2002). This stereotype is present even in societies with high levels of gender-equity (e.g., Miller et al., 2015). The association of science as being male is linked to gender-specific attributions of success: across a broad range of fields and age groups, success has been shown to be implicitly attributed to innate talent for males and to hard work for females (Proudfoot et al., 2015; Verniers and Martinot, 2015). Recent research on "field-specific ability beliefs" has shown that a high emphasis on "brilliance" (i.e., raw talent) as a requirement for success goes along with a low share of female students among the graduates of a specific field (Leslie et al., 2015; Meyer et al., 2015). To summarize, across academic fields, success is attributed more to some form of innate "brilliance" for males than females. In combination with the field-specific belief in the importance of "brilliance" that dominates many STEM fields, this may result in negative stereotypes against women.

A possible mediating mechanisms between cultural expectations and stereotypes on the one hand, and women's underrepresentation in math-intensive STEM fields on the other hand, is that women may be more likely than men to feel that they do not belong in these fields. Social belonging, more precisely the feeling of "belonging uncertainty" (Walton and Cohen, 2007), has been linked to students' persistence, well-being, and academic achievement in STEM subjects (Walton and Cohen, 2011; Walton et al., 2015). Belonging uncertainty, to specify the term, is an individual's perception that "people like me do not belong here" (Walton and Cohen, 2007, p. 83). Often, belonging uncertainty is reported by members of underrepresented social groups against whom negative stereotypes exist, like women in math-intensive STEM fields (Walton et al., 2015), or minority students in college (Walton and Cohen, 2011). Students experiencing belonging uncertainty are more negatively affected by difficulties they face during their studies, and are more likely to give up their course of study or study field (Walton and Cohen, 2011; Walton et al., 2015).

We argue that the perception that a specific STEM field requires male-associated "brilliance" may contribute to women's belonging uncertainty with regard to the respective field. Thus, it may contribute to women's reluctance to choose such a field, or to remain in it when facing difficulties. We had the chance to correlate field-specific ability beliefs and belonging-uncertainty in a group of students from five different STEM subjects at a prestigious university in Central Europe. To our knowledge, this is the first study measuring field-specific ability beliefs (Leslie et al., 2015) in a sample of university students enrolled in the respective fields. The field-specific ability beliefs of faculty are of course important, as they may influence the reactions and feedbacks that male vs. female students receive for their efforts (Leslie et al., 2015). However, the field-specific ability beliefs held by the students themselves will likely have a more direct impact on their feeling of belonging, and will thus influence their willingness to choose and to persist in a math-intensive STEM field. We hypothesize that this is particularly true for female students: There are negative stereotypes against women regarding their possession of raw talent ("They might be the harder workers, but compared to men, women have lower levels of raw talent"; compare Proudfoot et al., 2015; Verniers and Martinot, 2015). These may lead to belonging uncertainty, and eventually underrepresentation of women, in fields that they perceive as requiring high levels of raw talent that cannot be compensated for by hard work.

STUDY GOALS

In the present study, we aim to investigate field-specific differences in ability beliefs in university students in a range of STEM subjects, and to demonstrate a correlation between field-specific ability beliefs emphasizing "brilliance" on the one hand and belonging uncertainty on the other hand.

A further goal is to replicate the findings by Leslie et al. (2015) in a student sample. The authors found a negative correlation between faculty's endorsement of brilliance as a prerequisite for success and the percentage of female Ph.D. recipients in the respective field. Here, we explore whether there also is a negative correlation between students' endorsement of brilliance as a prerequisite for success in their chosen field and the percentage of female students in the same field. Our sample allows, to some extent, to disentangle the impact of the degree of "math intensiveness" on the one hand, and the minority status of women on the other hand: We recruited participants from the departments of Mathematics, Physics, Computer Science, Electrical Engineering, and Mechanical Engineering. Traditionally, at least at our university, the demands on competencies in mathematics required by the curriculum are higher in the study programs in Mathematics and Physics than they are in the subjects of Computer Science, Electrical Engineering, and Mechanical Engineering. At the same time, the proportion of female students is lower in these three programs than in Mathematics and Physics. This allows us to explore whether belonging uncertainty and the endorsement of brilliance are, in our sample, lower in the subjects with the highest math intensiveness (Mathematics, Physics), or in the subjects with the lowest number of women (Computer Science, Electrical Engineering, and Mechanical Engineering).

The research questions and hypotheses of this study are:

(1) To what extent do the field-specific ability beliefs of university STEM students emphasize "brilliance"? Are there

differences between the different STEM subjects and between genders? Is the highest endorsement of "brilliance" to be found in the very math-intensive fields, such as Math and Physics, or in study programs with the a very low percentage of female students, such as Engineering?

(2) To what extent can field-specific ability beliefs emphasizing "brilliance" predict belonging uncertainty in male and in female STEM students? Based on the theoretical framework behind the concept of belonging uncertainty, doubt about one's competence (e.g., from experiencing actual failure, or from activated negative stereotypes about one's social group) should be the more detrimental to one's feeling of belonging, the more one believes that success depends on some form of innate talent or "brilliance." Therefore, our hypotheses are that (a) there is a positive correlation between belief of brilliance and belonging uncertainty, and (b) that gender differences in belief in brilliance mediate gender differences in belonging uncertainty.

MATERIALS AND METHODS

Sample

Our participants were first-year students enrolled at a prestigious, rather male-dominated, technical university in Switzerland, i.e., the ETH Zurich. They came from five departments (Mathematics, Physics, Computer Science, Electrical Engineering, and Mechanical Engineering), all representing math-intensive STEM fields, although to a different degree. The curricula in Mathematics and Physics have a much stronger requirement in mathematics than the Computer Science and Engineering curricula (cf. Berkowitz and Stern, 2018). On the other hand, the proportion of female students is typically higher in Mathematics and Physics than in Computer Science and Engineering. All first-year students in the five departments were invited to participate in a short online survey toward the middle of their first term. In two consecutive years (2016 and 2017 cohorts), the survey was sent out to the students by the University administration as part of a larger teaching development project. The data of both cohorts was combined for all analyses. Initially, roughly 3000 students were invited to participate in the survey. A total of $n = 1424$ participated, of which $n = 1294$ gave informed consent for their data to be used for research purposes. Data on the survey items relevant for this study were missing for $n = 3$ students, leaving a total sample of $n = 1291$. The sample included $n = 235$ women (18% of the sample), which is roughly representative for the mean male-to-female ratio in the five surveyed departments.

The data for calculating the percentage of female student's in the five departments were collected from the University administration and encompass the entire cohorts of students starting their studies in the five departments in Fall 2016 or Fall 2017. The percentages were 25% female students in Mathematics, 19% in Physics, 15% in Electrical Engineering, and 12% in Mechanical Engineering and Computer Science, respectively.

Survey

The online survey had several parts, of which only one is relevant for the study at hand. In a first part of the survey, students answered questions that concerned a change in the first-year examination mode (13 items). The second part of the survey assessed students' fields-specific-ability beliefs and belonging uncertainty (7 items). The third part of the survey asked students for alternative plans to studying (4 items), and the fourth part aimed to gauge their general well-being at their new school (4 items). Finally, students could give feedback in an open-answer item. At the end of the survey, students received information on the further handling of their data and were asked to give their consent to use their answers for research purposes.

For the current study, only students' answers in the second part of the survey are of relevance. In this part of the questionnaire, students' answered the items of the field-specific ability belief scale (FSAB; 4 items) originally published by Leslie et al. (2015) and of the belonging uncertainty scale (BU; 3 items) by Walton and Cohen (2007). The items of both scales were presented on one page, intermixed in one block titled "Your studies at [school name]." All items were rated on a 7-point answering scale. Only the endpoints were labeled as "do not agree at all" (German original: "trifft gar nicht zu"; coded as 1) or "completely agree" (German original: "trifft völlig zu," coded as 7).

The FSAB items were translated to German and reworded to assess the perception of non-faculty members, i.e., students in a given field (sample item: "Being successful in my subject of study requires a special aptitude that just can't be learned."). The BU items were translated to German and adapted to be specific for the students' school, i.e., ETH Zurich (sample item: "When things are going badly, I feel that maybe I don't belong at ETH after all."). The complete list of questions (in German and in the English re-translation) is available as **Supplementary Table 1**.

RESULTS

Field-Specific Ability Beliefs

In our sample, the internal consistency of the four FSAB items proved satisfactory (Cronbach's $\alpha = 0.78$). Thus, for all further analyses, the mean of all four items was calculated. **Table 1** gives the mean of the FSAB scale for students of the five departments and for both genders. Higher values indicate greater endorsement of brilliance or innate talent as a prerequisite of success in the chosen field of study (the answering scale runs from 1 to 7, with 4 representing the middle of the scale). All but one of the mean values are in the lower half of the scale, indicating disagreement rather than agreement with items expressing beliefs in brilliance. An ANOVA with department and gender as factors yielded two statistically significant main effects (department: $F_{(4, 1281)} = 4.83$; $p = 0.001$; gender: $F_{(1, 1281)} = 6.97$; $p = 0.01$], but no statistically significant interaction [$F_{(4, 1281)} = 1.46$; $p = 0.21$]. Thus, regardless of their gender, students from different departments differed systematically in their endorsement of brilliance as a necessary precondition for success. Descriptively, Engineering and Computer Science students reported lower beliefs in brilliance than students of Physics or Mathematics. The effect size for the difference between the lowest and highest scoring departments (Mechanical Engineering vs. Physics) is $d = 0.35$. Across departments, women were more likely than men to endorse raw talent as a necessary condition for academic success in their chosen field of study ($d = 0.22$).

Belonging Uncertainty

In our sample, the internal consistency of the three BU items was very low (Cronbach's $\alpha = 0.42$), which was due to one of the items ("When things are going well, I feel that I really belong at ETH") not loading with the other two. This is an issue already discussed by the authors of the original scale, who report that this items loads with the other two in some samples but not in others (Walton, 2018).

Thus, we dropped the item and used only the mean of the two remaining items for our sample (Cronbach's $\alpha = 0.82$). **Table 2** gives the BU score for students of the five department and for both genders. Higher values indicate greater belonging uncertainty (the answering scale runs from 1 to 7, with 4 representing the middle of the scale). Most mean values are in the upper half of the scale, indicating agreement rather than disagreement with the belonging uncertainty items.

An ANOVA with department and gender as factors yielded two statistically significant main effects [department: $F_{(4,1281)} = 3.33$; $p = 0.01$; gender: $F_{(1,1281)} = 33.22$; $p < 0.001$], but no statistically significant interaction [$F_{(4,1281)} = 0.84$; $p = 0.49$]. Across departments, women reported higher levels of belonging uncertainty than men did ($d = 0.45$). Comparing departments, Mechanical Engineering students reported the lowest level of belonging uncertainty, and Physics students the highest, with the other departments somewhere in between. The effect size for the difference between the lowest and highest scoring departments (Mechanical Engineering vs. Physics) is $d = 0.46$.

Correlation and Mediation Analyses

The two scales, FSAB and BU, showed a small, positive correlation ($r = 0.19$; $p < 0.001$; $n = 1291$) that was very similar in size for men ($r = 0.19$; $p < 0.001$; $n = 1057$) and women ($r = 0.18$; $p = 0.01$; $n = 234$).

Further, the results were similar for all departments, with r ranging between $r = 0.16$ (Electrical Engineering) and $r = 0.21$ (Mechanical Engineering).

In a mediation analysis (using the PROCESS macro for SPSS as described by Hayes, 2017), we explored the possibility whether women's higher belief in brilliance might explain their higher belonging uncertainty. The total effect of gender on belonging uncertainty ($B = 0.79$; $p < 0.001$) could be split into a direct effect (i.e., unmediated; $B = 0.72$; $p < 0.001$), and an indirect effect (i.e., mediated by belief in brilliance; $B = 0.07$, bootstrapped 95% CI: [0.02; 0.13]). Thus, a relatively small, but significant portion of the effect of gender on belonging uncertainty could be explained by women's higher belief in brilliance.

Figure 1 shows a scatterplot of students' endorsement of brilliance against the percentage of female students in their field. A negative trend would have been expected in analogy to the negative correlation between faculty's endorsement of brilliance and the percentage of female PhD recipients in their field found by Leslie et al. (2015). We only have five departments to plot, and thus cannot reliably calculate a statistical correlation. However, the direction of the association appears to be positive rather than negative, with the two departments with the highest percentage of female students (Physics and Mathematics) also being those with the highest general endorsement of brilliance as a prerequisite for success. Thus, at least in our sample, it is math-intensiveness, rather than the minority status of women, which is associated with higher belief in brilliance of students in the respective field.

DISCUSSION

This is the first study testing the field-specific ability beliefs (using the instrument developed by Leslie et al., 2015) of university students in the respective fields. In the following, we compare our results to the previous findings by Leslie and colleagues in United States samples of academics (Leslie et al., 2015) and lay people (Meyer et al., 2015). Further, we discuss the correlation between field-specific ability beliefs emphasizing brilliance (belief in brilliance) and belonging uncertainty that we found in our sample.

Comparison With Previous Findings by Leslie et al. (2015)

In line with the previous studies by Leslie and colleagues (Leslie et al., 2015; Meyer et al., 2015), we find that academic fields differ with respect to the amount of "brilliance" that is assumed to be required for success. Within STEM fields, we find the highest endorsement of brilliance-related statements in the fields of Physics and Mathematics, which replicates the findings by Leslie and colleagues (Leslie et al., 2015; Meyer et al., 2015). In our sample, students in nearly all fields and of both genders were more likely to reject, rather than endorse, brilliance-related statements. In contrast, the mean values reported by Leslie et al. (2015) for faculty members were all above the midpoint of their answering scale, indicating agreement rather than rejection.

TABLE 1 | Belief in brilliance (min. 1, max. 7) according to subject and gender.

	Male students			Female students			All students		
	M	SD	*n*	*M*	SD	*n*	*M*	SD	*n*
Mathematics	3.79	1.21	140	4.10	0.97	48	3.87	1.16	188
Physics	3.87	1.20	161	3.95	1.06	54	3.89	1.16	215
Comp. Sci.	3.53	1.26	234	4.07	1.24	37	3.60	1.27	271
Electric. Eng.	3.50	1.11	198	3.71	1.12	39	3.53	1.11	237
Mech. Eng.	3.52	1.04	324	3.48	0.89	56	3.51	1.02	380
Total	3.61	1.16	1057	3.85	1.06	234	3.65	1.14	1291

TABLE 2 | Belonging Uncertainty (min. 1, max. 7) according to subject and gender.

	Male students			Female students			All students		
	M	SD	*n*	*M*	SD	*n*	*M*	SD	*n*
Mathematics	4.13	1.92	140	5.28	1.51	48	4.42	1.89	188
Physics	4.78	1.90	161	5.19	1.76	54	4.89	1.87	215
Comp. Sci.	4.51	1.77	234	5.23	1.71	37	4.61	1.78	271
Electric. Eng.	4.41	1.77	198	5.12	1.97	39	4.53	1.82	237
Mech. Eng.	3.94	1.75	324	4.74	1.63	56	4.05	1.75	380
Total	4.31	1.83	1057	5.10	1.71	234	4.45	1.83	1291

Thus, overall, the belief in brilliance was substantially lower in our student sample than in Leslie et al.'s (2015) faculty sample. This might have been due, first if all, to different operationalizations of success in our student vs. Leslie et al.'s (2015) faculty sample: While we used the term "success" in a rather unspecific way [e.g., "Being successful in my subject of study requires a special aptitude (. . .)"], Leslie et al. (2015) specifically asked their faculty respondents what it would take to be a top scholar in a given field [e.g., "Being a top scholar in my subject of study requires a special aptitude (. . .)"]. It is possible that we would have found higher agreement rates if we had framed success in this way to our students as well, or that Leslie et al. (2015) would have found lower agreement rates if they had asked faculty members about factors influencing success as a student of their field more generally. This is, however, a question that only additional research can answer. On the other hand, the different endorsement rates could also reflect the different levels of seniority and professional status of the respondents: Given the human tendency for self-enhancing attributions of success (Miller and Ross, 1975), the successful academics surveyed by Leslie et al. (2015) should have been more willing to attribute success to talent or "brilliance" than the first-year students surveyed in our study. Further, to the extent that emphasis on "brilliance" is part of the culture of a given field that experts acquire during the course of their studies, belief in brilliance would be expected to be stronger in field experts (i.e., faculty) than novices (i.e., first-year students). Finally, the difference in our findings to those of Leslie et al. (2015) might also be due to more general cultural differences between the United States and Switzerland (or in a broader sense Europe).

Meyer et al. (2015) studied a United States sample of lay people with varying degrees of exposure to the various fields of science. They, too, found that their respondents predominantly endorsed brilliance-related statements (with means for the majority of all surveyed fields, and for all STEM fields, above the scale mid-point). Thus, the belief in the necessity of innate talent for success in academic fields might be a more typical belief in United States than in Swiss samples.

Our data set allowed us to disentangle "math intensiveness" and "minority status of women"; we found that the highest belief in brilliance was found in the most math-intensive fields, rather than in those fields with the fewest numbers of women. In contrast, previous research with faculty members (Leslie et al., 2015) as well as with lay people (Meyer et al., 2015) found a (negative) correlation between an emphasis of brilliance for success in a given field and the percentage of women among

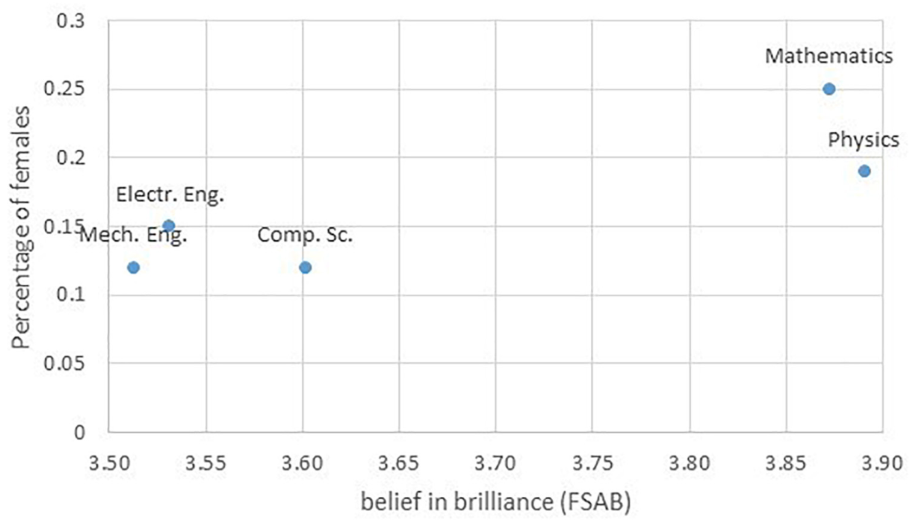


FIGURE 1 | Scatterplot of belief in brilliance and the percentage of female students in the five departments.

its successful graduates (at Ph.D. level). Our study was not a direct replication of the Leslie et al. (2015) study, as it used a different respondent sample (first-year students instead of faculty) and a different criterion (percentage of females among enrolled students, and not among successful graduates). Further, we only studied a very restricted sub-set of five STEM fields. Nevertheless, our findings demonstrate that the correlation found by Leslie et al. (2015) is only observable when considering the full range of academic subjects. Within STEM, math-intensiveness may be a better predictor of belief in brilliance.

Finally, in contrast to previous studies, we found systematic gender differences in field-specific ability beliefs, with women being less reluctant than men to endorse brilliance-related statements. We also found higher levels of belonging uncertainty in women than in men, as well as a positive correlation between belief in brilliance and belonging uncertainty for both genders. These findings will be discussed in detail in the next section.

Belief in Brilliance and Belonging Uncertainty

To the best of our knowledge, our study is the first showing that the more students believe that innate talent is a prerequisite for success in their field of study, the more likely they are to experience belonging uncertainty. This means that they are less likely to think that they actually do belong in their chosen field. The correlation was small (around $r = 0.19$ for both genders), which would be expected given the many possible factors that could influence students' perception of "belonging" to their chosen field of study. Nevertheless, we consider our result relevant, as belonging uncertainty is a plausible mechanism by which belief in brilliance could influence the paths that men and women chose for their future careers inside or outside of STEM fields.

Correlation does not imply causation. On the one hand it may be that the more students believe that success in their field depends on raw talent, the more anxious they feel about the amount of talent that they actually possess and the less certain they are that it will suffice to succeed in their studies, resulting in increased belonging uncertainty. The belief that academic success depends on talent, which cannot be increased and thus is largely out of one's control, corresponds to a "fixed," or entity theory of talent (cf. Dweck, 2007; Yeager and Dweck, 2012). Thus, the correlation between belief in brilliance and belonging uncertainty may have been mediated by students' belief in an entity theory of talent.

On the other hand, students experiencing failures (e.g., trouble keeping up with coursework), or negative stereotypes (e.g., "women are not smart enough to succeed in this field") early in their course of study may develop the hypotheses that (a) their chosen field requires prerequisite talent out of their reach, and therefore (b) is not be the field in which they actually belong. Thus, students searching explanations for their failures and struggles may have developed both stronger beliefs in brilliance and stronger belonging uncertainty.

While the magnitude of the correlation between students' beliefs in brilliance and their belonging uncertainty was similar for both genders, women reported higher levels of both variables than men did. As the score assessing belief in brilliance included several reverse coded items, this finding is unlikely to merely reflect a tendency of women choosing more affirmative answers than men. Also, the gender difference was found in 4 out of 5 of the surveyed departments (see **Table 1**). Thus, it cannot result from the overrepresentation of students from fields with a high emphasis on brilliance in our female sample. According to the two explanations developed in the last paragraph, we cannot exclude the possibility that the women in our sample had a more "fixed" theory of talent *a priori*. This may have made them more anxious about their own talent, and thus they were less certain to belong to their chosen field.

On the other hand, women may also have experienced more failures during their first weeks of study than men did, leading them to report higher levels of belonging uncertainty, and making them more likely to assume that their chosen field requires levels of innate talent of which they do not dispose. Finally, experiencing failure and/or bias may have led women to activate negative stereotypes about their gender. Consequently, they may have developed the belief that they lack essential talents for being successful in their chosen field because of their gender. Therefore, they may have been less likely to reject brilliance-related statements than men are, and more likely to experience belonging uncertainty (cf. Walton and Cohen, 2007). However, no conclusive inferences concerning the reason for women's higher belonging uncertainty and higher beliefs in brilliance can be drawn based on the obtained data. In future studies, in addition to assessing field-specific ability beliefs and belonging uncertainty, it would therefore be interesting to assess students' implicit stereotypes about science and gender, their goal orientations and attribution patterns, and their experiences of successes, failures, and obstacles (including negative stereotypes) during their studies.

Field-Specific Ability Beliefs and Gendered Paths Into STEM

In line with the findings obtained by Leslie and colleagues (Leslie et al., 2015; Meyer et al., 2015), our results show that STEM fields differ in the amount of "brilliance" that people assume to be required for success. From our surveys we cannot conclude to what extent these beliefs reflect true affordances of specific fields. However, a study run with an earlier cohort from the same university revealed that general intelligence could better explain achievement differences in Mathematics exams among students from Physics and Mathematics than among students from Mechanical Engineering (Berkowitz and Stern, 2018). These data suggest that field-specific ability profiles are reflected in field-specific ability beliefs, which themselves may shape processes of evaluation and selection information on who becomes a successful scholar in a given field. Field-specific ability beliefs emphasizing the necessity of brilliance, combined with the cultural stereotype of associating brilliance with men rather than women, will lead to practices and processes in a field that eventually exclude women (Leslie et al., 2015) and may undermine women's interest in specific fields (Bian et al., 2018). Our results show that, even among young women who have chosen to study a STEM subject, biases linking science to "brilliance" are prevalent, and can partly explain their higher belonging uncertainty in these fields.

ETHICS STATEMENT

This study was carried out in accordance with the Guidelines for Research Integrity and Good Scientific Practice by the Ethics Commission at the ETH Zurich. All subjects gave written informed consent. The protocol was approved by the Ethics Commission of the ETH Zurich.

AUTHOR CONTRIBUTIONS

All authors made substantial contributions to the conception and design of the study. ES and RS recruited the support of the participating departments, made substantial contributions to interpreting the results, and to revising several drafts of the manuscript. AD devised the materials, supervised data collection, analyzed the data, and prepared the manuscript, tables, and figure.

ACKNOWLEDGMENTS

The authors would like to thank the Center for Educational Development and Technology (LET) at ETH for its support. Specifically, the authors would like to acknowledge the scientific contributions of Dr. Nora Dittmann Domenichini and Dr. Bruno Rütsche to the design and data collection for this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01114/full#supplementary-material

REFERENCES

Berkowitz, M., and Stern, E. (2018). Which cognitive abilities make the difference? Predicting academic achievements in advanced STEM studies. *J. Intell.* 6, 1–24. doi: 10.3390/jintelligence6040048

Berkowitz, M., Stern, E., Hofer, S., and Deiglmayr, A. (in press). "Girls, boys and schools: On gender (in)equalities in education," in *The Cambridge International Handbook on Psychology of Women*, eds F. M. Cheung and D. F. Halpern (Cambridge: Cambridge University Press).

Bian, L., Leslie, S. J., Murphy, M., and Cimpian, A. (2018). Messages about brilliance undermine women's interest in educational and professional opportunities. *J. Exp. Soc. Psychol.* 76, 404–420. doi: 10.1016/j.jesp.2017.11.006

Ceci, S. J., Ginther, D. K., Kahn, S., and Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychol. Sci. Public Int.* 15, 75–141. doi: 10.1177/1529100614541236

Cheryan, S., Ziegler, S. A., Montoya, A. K., and Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychol. Bull.* 143, 1–35. doi: 10.1037/bul0000052

Dweck, C. (2007). "Is math a gift? Beliefs that put females at risk," in *Why aren't more Women in 1460 Science? Top Researchers Debate the Evidence*, eds S. J. Ceci and W. M. Williams (Washington, WA: APA Press), doi: 10.1037/11546-004

Hayes, A. (2017). *Introduction to Mediation, Moderation, And Conditional Process Analysis*, 2nd Edn. New York, NY: Guilford.

Hofer, S. I. (2015). Studying gender bias in physics grading: the role of teaching experience and country. *Int. J Sci. Edu.* 37, 2879–2905. doi: 10.1080/09500693.2015.1114190

Leslie, S.-J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 262–265. doi: 10.1126/science.1261375

Meyer, M., Cimpian, A., and Leslie, S. J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Front. Psychol.* 6:235. doi: 10.3389/fpsyg.2015.00235

Miller, D. I., Eagly, A. H., and Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: evidence from 66 nations. *J. Edu. Psychol.* 107, 631–644. doi: 10.1037/edu0000005

Miller, D., and Ross, M. (1975). Self-serving biases in the attribution of causality: fact or fiction? *Psychol. Bull.* 82, 213–225. doi: 10.1037/h0076486

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci.* 109, 16474–16479. doi: 10.1073/pnas.1211286109

Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *J. Pers. Soc. Psychol.* 83, 44–59. doi: 10.1037/0022-3514.83.1.44

Proudfoot, D., Kay, A. C., and Koval, C. Z. (2015). A gender bias in the attribution of creativity: archival and experimental evidence for the perceived association between masculinity and creative thinking. *Psychol. Sci.* 96, 1751–1761. doi: 10.1177/0956797615598739

Stoet, G., and Geary, D. C. (2018). The gender-equality paradox in Science. Technology, Engineering, and Mathematics education. *Psychol. Sci.* 29, 581–593. doi: 10.1177/0956797617741719

Thébaud, S., and Charles, M. (2018). Segregation, stereotypes, and STEM. *Soc. Sci.* 7, 1–18.

Verniers, C., and Martinot, D. (2015). Perception of students' intelligence malleability and potential for future success: unfavourable beliefs towards girls. *Br. J Educ. Psychol.* 85, 289–299. doi: 10.1111/bjep.12073

Walton, G. M. (2018). *Belonging and Belonging Uncertainty. Online Ressource.* http://gregorywalton-stanford.weebly.com/uploads/4/9/4/4/49448111/belonging_belonginguncertainty.pdf (accessed October 30, 2018).

Walton, G. M., and Cohen, G. L. (2007). A question of belonging: race, social fit, and achievement. *J. Pers. Soc. Psychol.* 92, 82–96. doi: 10.1037/0022-3514.92.1.82

Walton, G. M., and Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* 331, 1447–1451. doi: 10.1126/science.1198364

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., and Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *J. Edu. Psychol.* 107, 468–485. doi: 10.1037/a0037461

Wang, M. T., and Degol, J. L. (2017). Gender gap in Science, Technology, Engineering, and Mathematics (STEM): current knowledge, implications for practice, policy, and future directions. *Edu. Psychol. Rev.* 29, 119–140. doi: 10.1007/s10648-015-9355-x

Yeager, D. S., and Dweck, C. S. (2012). Mindsets that promote resilience: when students believe that personal characteristics can be developed. *Educ. Psychol.* 47, 302–314. doi: 10.1080/00461520.2012.722805

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 June 2019
doi: 10.3389/fpsyg.2019.01243

Gendered Paths Into STEM-Related and Language-Related Careers: Girls' and Boys' Motivational Beliefs and Career Plans in Math and Language Arts

Rebecca Lazarides[1]* *and Fani Lauermann*[2]

[1] Department of Education, Universität Potsdam, Potsdam, Germany, [2] Center for Research on Education and School Development (IFS), TU Dortmund University, Dortmund, Germany

Women are often underrepresented in math-intensive fields like the physical sciences, technology, engineering and mathematics. By comparison, boys relative to girls are less likely to strive for jobs in social and human-services domains. Relatively few studies have considered that intra-individual comparisons across domains may contribute to gendered occupational choices. This study examines whether girls' and boys' motivational beliefs in mathematics and language arts are predictive of their career plans in these fields. The study focusses on same domain and cross-domain effects and investigates bidirectional relations between motivational beliefs and career plans. Data for this study stem from 1,117 ninth and tenth graders (53.2% girls) from secondary schools in Berlin, Germany. Findings show systematic gender differences in same-domain effects in mathematics: girls' comparatively lower mathematics self-concept and intrinsic value predicted a lower likelihood of striving for a math-related career. Cross-domain effects were not related to gender-specific career plans, with only one exception. Girls' lower levels of intrinsic value in mathematics corresponded to a higher likelihood of striving for a career in language-related fields, which subsequently predicted lower levels of intrinsic value in mathematics. This finding points to a need to address both gender-specific motivational beliefs and gender-specific career plans in school when aiming to enhance more gender equality in girls' and boys' occupational choices.

Keywords: gendered motivational beliefs, career plans, mathematics, language arts, dimensional comparison

OPEN ACCESS

Edited by:
Carl Senior,
Aston University, United Kingdom

Reviewed by:
Uner Tan,
Çukurova University, Turkey
Malte Jansen,
Institute for Educational Quality Improvement (IQB), Germany

Correspondence:
Rebecca Lazarides
rebecca.lazarides@uni-potsdam.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *03 December 2018*
Accepted: *10 May 2019*
Published: *06 June 2019*

Citation:
Lazarides R and Lauermann F (2019) Gendered Paths Into STEM-Related and Language-Related Careers: Girls' and Boys' Motivational Beliefs and Career Plans in Math and Language Arts. Front. Psychol. 10:1243. doi: 10.3389/fpsyg.2019.01243

INTRODUCTION

A substantial amount of research has focused on social and individual factors contributing to persistent gender disparities in the selection and pursuit of particular career paths (for an overview, see for example Watt, 2016; Wang and Degol, 2017). This research shows that women are often underrepresented in math-intensive fields like science, technology, engineering and mathematics (STEM) (Watt et al., 2012; Watt, 2016). By comparison, boys relative to girls are less likely to strive for jobs in social and human-services domains (Su and Rounds, 2015; Wolter et al., 2015), which often require higher levels of verbal than math skills (see National Center for O*Net Development, 2014; Lauermann et al., 2015). Research based in expectancy-value theory (EVT; Eccles et al., 1983) and the dimensional comparison theory (DCT; Möller and Marsh, 2013) suggests that systematic differences in students' domain-specific motivational beliefs (i.e., academic self-concepts and task values) can contribute to such gender-specific career paths. Girls often report lower levels of intrinsic and utility value of mathematics than boys (Gaspard et al., 2015) as well as lower self-concept of ability in this domain (Marsh and Yeung, 1998). Boys in turn report lower levels of intrinsic value and self-concept in language arts (Jacobs et al., 2002; Watt, 2004). Differences in academic beliefs about mathematics and language arts can thus potentially shape subsequent career preferences for occupations that are perceived as either math-intensive (e.g., STEM) or verbal-intensive (e.g., communication, teaching). Relatively few studies, however, have considered that intraindividual comparisons across such domains as math and language arts may also contribute to gendered educational and occupational choices (Nagy et al., 2006; Lauermann et al., 2015). A choice against a math-intensive career, for instance, may be linked to a comparatively higher interest in the verbal domain rather than a low interest in math.

In the present study, we build upon this research and examine whether adolescent girls' and boys' motivational beliefs in mathematics and language arts are predictive of their career plans in these fields. In line with previous research (Lauermann et al., 2017; Lazarides et al., 2017), we understand career plans both as an outcome and as a precursor of students' motivational beliefs. Academic motivations in math and language arts may lead students to choose careers for which these domains are important; at the same time, choosing a career that requires relatively high levels of math or verbal skills may increase students' motivations to engage in these academic domains as a means of accomplishing their career goals (for the math domain, see e.g., Lauermann et al., 2017; Lazarides et al., 2017). Therefore, we investigate potential bidirectional relations between students' motivational beliefs and career plans.

Gendered Motivational Beliefs and Career Plans

Eccles and colleagues' expectancy-value theory (EVT; Eccles et al., 1983, 1998) proposes that individuals' motivational beliefs – defined as their subjective valuing of and expected success in a given task – are important predictors of students' task-related activities, achievements, career plans, and career attainment. Task values are defined as "the quality of the task that contributes to the increasing or decreasing probability that an individual will select it" (Eccles, 2005, p. 109) and are described in terms of four components: students' task-related enjoyment (intrinsic value), the perceived usefulness of activities and tasks for own short- and long-term goals (utility value), the personal importance of doing well on a given task (attainment value), and the subjective cost related to engaging in given activities and tasks (cost value). In this study, we focus on students' intrinsic and utility values because these components have been shown to be important antecedents of students' educational and occupational choices (Nagy et al., 2006; Watt et al., 2012; Lauermann et al., 2017; Lazarides et al., 2019). Success expectancies are defined as individuals' beliefs about how well they will do on upcoming tasks, either in the immediate or long-term future (Eccles and Wigfield, 2002). The key conceptual difference between students' success expectancies and academic self-concept of ability is that success expectancies refer to future achievements (Wigfield and Eccles, 2000), whereas academic self-concept refers to past accomplishments that inform students' self-evaluations (Marsh et al., 2018). However, these two constructs are often not empirically distinguishable (Wigfield and Eccles, 2000), possibly because students use their past experiences as an important reference point to estimate the subjective likelihood of succeeding in a given academic domain in the future (e.g., Marsh et al., 2005). When both constructs reference the same domain (e.g., math or reading), they typically form one factor (e.g., Eccles and Wigfield, 1995; Wigfield and Eccles, 2000). Similar to previous research in EVT, we focus on students' academic self-concept of ability as an important antecedent of students' expected success in a given domain. Studies have shown that students' academic self-concept is highly related to their achievement (Marsh et al., 2005), whereas their task values are comparatively more strongly related to career choices and aspirations (e.g., Meece et al., 1990; Watt et al., 2012; Lazarides and Watt, 2015; Lauermann et al., 2017).

A number of studies demonstrate persistent gender differences in adolescents' domain-specific task values (e.g., Watt et al., 2012; Gaspard et al., 2015) and academic self-concept of ability (Marsh and Yeung, 1998). Girls, compared to boys, tend to report lower levels of intrinsic value (Frenzel et al., 2010; Watt et al., 2012; Gaspard et al., 2015) and lower academic self-concepts in mathematics (Marsh and Yeung, 1998). Girls also report lower levels of perceived utility of mathematics for their future life and for their job prospects (Gaspard et al., 2015). By comparison, boys report lower self-concept of ability (Marsh and Yeung, 1998; Arens and Jansen, 2016) and lower levels of interest in language-related domains (Yeung et al., 2011). Such gender-specific motivational beliefs are associated with gender differences in students' educational and career paths (Watt et al., 2012; Lauermann et al., 2017). In the math domain, girls tend to report comparatively lower levels of motivation and lower levels of interest in math-intensive careers. In an Australian sample of adolescents, Watt et al. (2012) found that girls participated less often in math courses than did boys and less often aspired to math-related careers. In a longitudinal U.S. sample, Lauermann et al. (2017) found a weak positive association between gender and grade 12 self-concept of ability in mathematics favoring male students, and male students were more likely to strive for and attain math-related careers as adults. In a longitudinal sample in Germany, Lazarides et al. (2017) found that boys reported higher levels of interest and utility value in math and were more likely than girls to strive for math-related careers. Regarding gender differences in domains in which women are typically overrepresented, Nagy et al. (2006) found that boys were less likely than girls to choose an advanced biology course in grade 12, and findings reported in Lauermann et al. (2015) suggest that girls were more likely than boys to consider human services occupations, which tend to be verbal-intensive. Building on this previous evidence, we examine whether gender differences in students' academic motivations, namely self-concepts of ability and task values, are linked to corresponding differences in adolescents' career choices. We focused on the domains of mathematics and language arts due to their critical role for a variety of occupational fields and due to persistent gender differences in these domains.

Dimensional Comparisons and Gendered Career Plans

Individuals' motivational beliefs are influenced by internal and external comparison processes (Eccles, 2009; Möller and Marsh, 2013). Individuals tend to assess their own skills by comparing their performance in a given domain with the performance of relevant peers (*external comparisons*) and by comparing their levels of performance across different domains (*internal comparisons*). Such cross-domain comparison processes play a central role in the development of students' academic self-concept of ability, as described in the internal/external frame of reference model (I/E model; Marsh, 1986). According to the I/E model, a continuum of core academic self-concepts exists, which include students' self-concept in the verbal domain and their self-concept in the math domain (Marsh et al., 2015). Students evaluate their abilities by comparing their performance in a given domain to their own past performances in this domain, to the observed performance of relevant peers, or to their own performance across domains. Consistent with the theoretical assumptions of the I/E model, a number of studies have documented negative contrast effects across the math and verbal domains (e.g., Brunner et al., 2008; Möller et al., 2009, 2011; Niepel et al., 2014). Whereas students' verbal achievement positively predicts their verbal self-concept of ability ("*same-domain effect*"), it has a negative effect on students' self-concept of ability in math ("*cross-domain effect*"). High performance in the verbal domain sets a high standard against which students' math performance is being compared, which then negatively affects their self-evaluated competence in math. Analogous contrast effects have been documented with regard to students' math performance and verbal self-concept of ability.

The dimensional comparison theory (DCT; Möller and Marsh, 2013) was developed as an extension of the I/E model (Marsh et al., 2015). A central contribution of DCT (Möller and Marsh, 2013) is that it incorporates contrast effects, assimilation effects, and same-domain effects across a wide range of academic subjects that are relatively similar ("near comparisons") or dissimilar ("far comparisons"). Negative contrast effects, or cross-domain effects, of students' achievement on their self-concept of ability are likely to apply across dissimilar domains like math and language arts (e.g., a negative effect of math achievement on verbal self-concept of ability and vice versa); positive assimilation effects are likely to apply across subjects that are similar to each other (e.g., a positive effect of math achievement on physics self-concept of ability); and same-domain effects apply within the same domain (e.g., a positive effect of math achievement on math self-concept of ability).

Furthermore, DCT expands upon the IE-framework by focusing on the "why," "with what" and "with what effect" questions of dimensional comparisons (Möller and Marsh, 2013). Notably, Möller and Marsh (2013, p. 553) point out that the vast majority of available evidence on the effects of dimensional comparisons (i.e., the "with what" question) has focused on students' domain-specific academic self-concepts, even though dimensional comparisons can also affect other outcomes such as mood, course selection, or career choices. Dickhäuser et al. (2005), for example, focused on academic self-concept and course selection in biology and chemistry, and showed significant negative paths from students' self-concepts on the selection of non-corresponding subjects. Lauermann et al. (2015) examined the relations between adolescents' motivational beliefs across two academic domains, English and math, on their math/science-related and human services-related career plans and identified significant negative paths from students' English self-concept and English task values on their career plans in math.

In the present study, we focus on dimensional comparison effects among motivational beliefs (academic self-concept and task values) and career plans in math and language-related domains and examine whether these dimensional comparison effects may contribute to gender disparities in adolescents' domain-specific motivations and career plans.

A few recent studies have examined gender differences in educational (Nagy et al., 2006; Wang et al., 2013; Guo et al., 2017) and occupational choices (Parker et al., 2012; Lauermann et al., 2015) based on the theoretical assumptions of EVT and DCT. These studies showed that dimensional comparison effects might partially explain gender-specific educational and occupational choices. For instance, in a study in the United States, girls reported significantly higher valuing of English as a subject domain than did boys, which not only positively predicted their preference for human-services careers but also negatively predicted their interest in pursuing careers in math and science (Lauermann et al., 2015). Wang et al. (2013) showed similar effects for STEM careers in a U.S. sample; girls were more likely than boys to have high math and high verbal ability, which corresponded to a lower likelihood of pursuing STEM careers. Another study with German adolescents (Nagy et al., 2006) found that having high levels of math achievement and math self-concept of ability negatively predicted boys' enrolment in advanced biology courses, but did not affect girls' enrolment in such courses. These studies thus suggest that negative cross-domain effects may differentially affect girls' and boys' educational and career choices.

Taken together, this evidence suggests that dimensional comparison processes can contribute to gendered educational and occupational choices. However, these studies have focused on ability (Wang et al., 2013), single task value components such as intrinsic value (Nagy et al., 2006), or on a composite score of all task values (Lauermann et al., 2015). Thus, the role of different motivational components like students' intrinsic, utility, and attainment value has not been systematically examined. Furthermore, the reciprocal longitudinal associations between students' academic motivations and career plans remain understudied (e.g., Lauermann et al., 2017). Finally, most of the cited research has focused on the math domain, and only a handful of studies have focused on career plans in verbal domains (e.g., Durik et al., 2006). Thus, the present study examines the longitudinal relations between girls' and boys' task value components (intrinsic, utility, and attainment value), academic self-concepts, and career plans in mathematics and language-related domains.

The Present Study

Informed by both EVT and DCT, the primary objective of this longitudinal study is to examine the predictive effects of student gender on their motivational beliefs and career plans in mathematics and language arts. We examine same-domain and cross-domain effects and consider the potential reciprocity of the relations between motivational beliefs and career plans. Based on our review of literature and theoretical considerations, we derived a set of five hypotheses focusing on gender differences, same-domain associations, and cross-domain effects in the math and language arts domains. First, we hypothesize that girls will report lower motivational beliefs (academic self-concept and task values) in mathematics than boys, and that girls will be less likely than boys to strive for careers in math-intensive fields (*Hypothesis 1*). We also hypothesize that boys will report lower motivational beliefs (academic self-concept and task values) in language arts than girls, and that boys will be less likely than girls to strive for careers in language-related fields (*Hypothesis 2*). Third, we expect to find positive same-domain associations between motivational beliefs and career plans, such that mathematics (vs. language-related) task values and self-concepts will positively predict math-related (vs. language-related) career plans (*Hypothesis 3*). We also expect to find negative cross-domain effects between math- and language arts-related motivational beliefs and career plans; we expect that mathematics (vs. language-related) task values and self-concepts will negatively predict language-related (vs. math-related) career plans and vice versa (*Hypothesis 4*). Additionally, in line with the I/E model (Marsh, 1986), we expect to find positive same-domain effects, and negative cross-domain effects among students' grades and their motivational beliefs (self-concept of ability and task values) (Gaspard et al., 2018) (*Hypothesis 5*). Lastly, we expect to identify gender-specific (same-domain and cross-domain) motivational processes. We assume that the predictive effects of students' gender on career plans in math- and language-related domains are at least partly attributable to gender differences in motivational beliefs in math and language arts (*Hypothesis 6*).

The following control variables were included in all analyses: whether German was a native language and school type (academic track vs. comprehensive school). The schematic model of the tested relations is depicted in **Figure 1**.

MATERIALS AND METHODS

Sample

Data was drawn from the German Move Study (Motivation and Valuing in Mathematics; Lazarides and Rubach, unpublished), which examines the relations between students' perceptions of their mathematics teachers' beliefs, perceptions of teachers' instructional behaviors, and students' motivations. In the longitudinal study Move, data was obtained from parents, students, and their mathematics teachers concerning perceived teaching quality, learning support and motivation for mathematics at three measurement points, two of which were included in the present study· The participating schools were randomly selected from a list of all secondary schools in Berlin, and data were collected by trained research assistants at the end of a compulsory class, approximately 2 months after the beginning of the 2015 school year (Time 1), as well as after the mid- year mark in the spring of 2016 (Time 2). The survey administration took approximately 30 min. In this study, we used the data from 1,117 students (age: M = 14.59 years, SD = 0.88) who participated at the first time point. A total of 746 9th (54.0%) and 10th graders (46%) (age: M = 14.50 years, SD = 0.86) participated at the first two time points included in this study. Written informed consent was obtained from the parents of the participants. The Berlin Senate for Education, Youth, and Research approved the study. An ethics approval was not required at the time the study was conducted as per the then applicable institutional and national guidelines and regulations. The students (53.2% girls) came from 58 classrooms across 13 secondary schools in Berlin, Germany. The sample consisted of ninth (48%) and tenth (52%) graders.

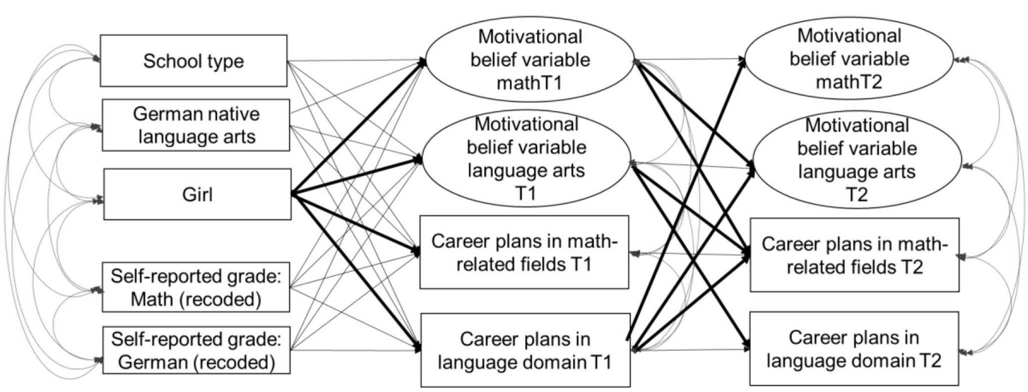


FIGURE 1 | Schematic model of the tested relations.

Most students (69.8%) reported that they were native speakers of German. Approximately half of the students attended a gymnasium school (the highest academic track in Germany; 51.8%), whereas the remaining students attended comprehensive schools (a type of secondary school that provides courses for different ability levels; 48.2%). Students' participation was voluntary.

Measures

The following sections provide an overview of all scales used in this study (the items are reported in **Appendix A**).

Ability Self-Concept

Students' self-concepts in mathematics and in language arts were assessed with an 8-item scale, with answer options ranging from 1 to 5 (see Steinmayr and Spinath, 2010). Four parallel domain-specific items were used to assess student's self-concept in mathematics (e.g., "I think I am . . . in mathematics" from "1 [not talented] to 5 [very talented]") and language arts (e.g., "I think I am . . . in German" from "1 [not talented] to 5 [very talented]"). The scales had very good internal consistency for math ($\alpha = 0.87$ at Time 1 and $\alpha = 0.88$ at Time 2) and language arts ($\alpha = 0.86$ at Time 1 and $\alpha = 0.87$ at Time 2).

Utility Value

Students' utility values in mathematics and language arts were assessed with a six-item scale based on Steinmayr and Spinath (2010), with answer choices ranging from 1 (*does not apply at all*) to 5 (*fully applies*). Three parallel items were used to assess utility value in mathematics (e.g., "Mathematics is useful for my future.") and language arts (e.g., "German is useful for my future"). The internal consistencies of these scales were very good in math ($\alpha = 0.88$ at Time 1 and $\alpha = 0.89$ at Time 2) and language arts ($\alpha = 0.91$ at Time 1 and $\alpha = 0.91$ at Time 2).

Intrinsic Value

Students' intrinsic values in mathematics and language arts were assessed with a six-item scale based on Steinmayr and Spinath (2010), with answer choices ranging from 1 (*does not apply at all*) to 5 (*fully applies*). Similar to utility value, three parallel items were used to assess intrinsic value in mathematics (e.g., "I like mathematics") and language arts (e.g., "I like German). The internal consistencies were very good in math ($\alpha = 0.92$ at Time 1 and $\alpha = 0.92$ at Time 2) and language arts ($\alpha = 0.93$ at Time 1 and $\alpha = 0.92$ at Time 2).

Career Plans

Students' mathematics-related career plans were assessed with the item "What occupation do you think are you going to have when you are 30 years old?" Two independent coders coded the math-relatedness of students' open-ended answers for relatedness to mathematics and language domains per nominated career using the Occupational Information Network (O*NET; National Center for O*Net Development, 2014) to quantify the importance of "knowledge of arithmetic, algebra, geometry, calculus, statistics, and their applications" (for level of importance of mathematics for the job) and of "the structure and content of the English language including the meaning and spelling of words, rules of composition, and grammar." (for level of importance of language arts) for each occupation named by the students, on a scale ranging from 0 (*not at all math-/language-related*) to 100 (*highly math-/language-related*). The interrater reliability was good, $\kappa = 0.82$.

Self-Reported Achievement

Achievement in mathematics and in the verbal domain was assessed by students' self-reported school grades at the end of the last semester in the school year. In Germany, school grades range from 1 (*very good*) to 6 (*unsatisfactory*), with lower values indicating better performance. To facilitate the interpretation of the findings, we reverse-coded the grades so that higher values reflect better achievement.

Statistical Analyses

A longitudinal structural equation modeling approach with a cross-lagged panel design was used, and the same variables were measured across time points (Kenny, 1975). This design enabled us to test the stability of constructs and the bidirectionality of effects between constructs. Three separate models were tested for students' self-concepts of ability and task values because these constructs tend to be highly correlated: Model 1 included students' academic self-concept, Model 2 included utility value, and Model 3 included intrinsic value. Each model included the motivational belief variable at Times 1 and 2 (autoregressive path) and career plans at Times 1 and 2 (autoregressive path) in both mathematics and language arts. Students' gender, immigrant background, school type and self-reported achievement in math and language arts were included as predictors of the Time 2 outcomes in all tested models. Reciprocal associations across Time 1 and Time 2 were tested between the motivational belief variables (academic self-concept, utility value, intrinsic value) and students' career plans in mathematics and language arts.

Before testing the structural equation models, scalar measurement invariance was tested for the latent variables in the full sample (Byrne, 2004). Scalar measurement invariance (intrinsic value) or partial scalar invariance (self-concept, utility value) was established indicating that the same latent constructs were assessed across time (for more detailed information, see **Appendix B**). Measurement invariance restrictions were kept when testing the hypothesized effects with longitudinal structural equation modeling. Measurement invariance was also tested across gender (see **Appendix B**). Mplus 8.0 was used for all analyses (Muthén and Muthén, 1998–2019). The TYPE IS COMPLEX function of Mplus was used to account for the nested structure of the data (students nested within classrooms), and maximum likelihood estimation with robust standard errors (MLR) was applied in all models. Missing data were handled by using full-information maximum likelihood (FIML) estimation. Information about participation rates per school, attrition rate across waves, and missing values on the study variables for each wave are reported in **Appendix C**. The following criteria were used to evaluate the goodness of fit of the models (Tanaka, 1993): Yuan-Bentler scaled χ^2 (mean-adjusted test-statistic robust to non-normality), Tucker and Lewis index (TLI), comparative fit index (CFI), and root mean square of approximation (RMSEA) with associated confidence intervals (CIs). Additionally, standardized root mean residual values (SRMR) were reported. TLI and CFI values greater than 0.95, RMSEA values lower than 0.06, and SRMR lower than 0.08 indicate satisfactory model fit (Hu and Bentler, 1999). Indirect effects were tested with the MODEL INDIRECT command and the CINTERVAL option. Bootstrapped standard errors and confidence intervals were obtained to evaluate the estimated indirect effects (Muthén and Muthén, 1998–2019). Indirect effects were estimated based on the product of coefficients method (MacKinnon et al., 2007; Williams and MacKinnon, 2008).

RESULTS

Descriptive Statistics and Bivariate Associations

Observed means and standard deviations for all variables included in subsequent analyses are reported in **Table 1**, and manifest bivariate correlations are reported in **Tables 2**, **3**. These correlational patterns suggest that – both at Time 1 and Time 2 – girls were less likely than boys to report career plans in mathematics (consistent with *Hypothesis 1*) and more likely to report career plans related to the language arts domain (consistent with *Hypothesis 2*). Furthermore, students with comparatively higher self-reported math achievement at Time 1 also reported comparatively higher career plans in mathematics at both Time 1 and Time 2. However, students' math achievement was also significantly positively related to their career plans in language arts at Time 2. Students' self-reported achievement in language arts at Time 1 was positively related to their career plans in the math and language arts domains at Time 1 and Time 2. These correlational patterns support positive same-domain associations for achievement and career plans, but no negative cross-domain associations emerged. Achievement is thus positively related to career aspirations across domains.

However, positive same-domain and negative cross-domain effects were corroborated for the associations between math-related career plans and math- and language arts-related motivations. Specifically, students' math-related career plans (at Time 1 and Time 2) were significantly and positively correlated with students' self-concept, utility value, and intrinsic value in math at both Time 1 and Time 2 (consistent with *Hypothesis 3* in the math domain) and were significantly and negatively correlated with students' self-concept, utility value, and intrinsic value in language arts at both Time 1 and Time 2 (consistent with *Hypothesis 4* in the math domain). Analogous same-domain associations were confirmed for the language arts domain (consistent with *Hypothesis 3* in the verbal domain). Specifically, language arts-related career plans at Time 1 and Time 2 were significantly and positively correlated with students' self-concept, utility value and intrinsic value in language arts at Time 1 and Time 2.[1] However, significant negative cross-domain associations were corroborated only for career plans in language arts at Time 1 and Time 2 and utility value in mathematics at Time 1 and Time 2 (only partly consistent with *Hypothesis 4* in the verbal domain). Thus, our expectations were fully supported in the math domain but were only partially supported in the language arts domain. In the following sections, these associations are further examined in the context of cross-lagged structural equations models.

Students' Self-Concept, Task Values, and Career Plans in Math and Language Arts

Model 1: Self-Concept and Career Plans Model

The model had good fit to the data, $\chi^2(211) = 363.13$, CFI = 0.98, TLI = 0.98, RMSEA = 0.03, SRMR = 0.02. Standardized and significant coefficients for this model are reported in **Figure 2**. The standardized coefficients of this model are reported in **Tables 4**, **5**. In line with our expectations (*Hypotheses 1*), girls, relative to boys, reported lower levels of self-concept in

[1]The correlation between Time 1 intrinsic value and Time 2 career plans was positive but failed to reach significance.

TABLE 1 | Descriptives of the study variables at time 1 (data for time 2 in parentheses) for boys (n = 506) and girls (n = 594).

	Girls		Boys				
Variable	***M***	***SD***	***M***	***SD***	**Wald χ^2, *df* = 1**	***d***	**Range**
Self-reported mathematics grade	4.03	1.02	3.96	1.23	0.68 n.s.	0.06	1–6
Self-reported grade in language arts	4.45	0.87	4.10	0.85	26.96***	0.41	1–6
Self-concept in mathematics	3.02 (3.09)	0.87 (0.86)	3.37 (3.45)	0.89 (0.93)	33.70***(25.26***)	0.40 (0.40)	1–5
Utility value in mathematics	2.90 (2.92)	0.93 (0.92)	3.16 (3.17)	0.93 (0.96)	11.80*** (7.48**)	0.30 (0.27)	1–5
Intrinsic value in mathematics	2.82 (2.93)	1.09 (1.12)	3.22 (3.19)	1.13 (1.17)	27.30*** (7.25***)	0.36 (0.23)	1–5
Self-concept in language arts	3.50 (3.55)	0.79 (0.80)	3.44 (3.38)	0.77 (0.78)	1.68 n.s. (9.08**)	0.07 (0.21)	1–5
Utility value in language arts	3.73 (3.74)	0.94 (0.94)	3.60 (3.55)	1.06 (0.99)	2.76 n.s. (4.61*)	0.13 (0.20)	1–5
Intrinsic value in language arts	3.52 (3.55)	1.05 (1.00)	3.25 (3.25)	1.09 (1.04)	12.14*** (14.43***)	0.25 (0.29)	1–5
Career plans related to mathematics	42.54 (41.83)	8.91 (9.04)	44.61 (45.41)	9.92 (10.06)	8.38** (17.91***)	0.22 (0.37)	0–100
Career plans related to language arts	57.65 (57.82)	6.39 (6.19)	55.70 (55.45)	6.67 (6.44)	16.92***(12.81***)	0.30 (0.38)	0–100

Latent means are reported for the latent variables including measurement invariance across time and gender groups. Grades were recoded. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TABLE 2 | Intercorrelations between the study variables.

	1	2	3	4	5	6	7	8	9	10	11	12
1) Girl												
2) German native	−0.04											
3) Lang achiev	0.20***	0.11*										
4) Math achiev	0.03	0.13***	0.46***									
5) Comp. school	−0.13***	−0.06	−0.26***	−0.23***								
6) Int math T1	−0.18***	0.01	0.10**	0.47***	−0.06							
7) Int math T2	−0.12**	0.04	0.12**	0.45***	−0.07	0.71***						
8) Int lang T1	0.12***	−0.01	0.30***	−0.16***	0.01	−0.09*	−0.09*					
9) Int lang T2	0.14***	−0.08*	0.15***	−0.13**	0.01	−0.13***	−0.11***	0.70***				
10) Self-concept math T1	−0.20***	0.01	0.22***	0.65***	−0.11***	0.81***	0.66***	−0.16***	−0.20***			
11) Self-concept math T2	−0.20***	0.06	0.20***	0.56***	−0.09*	0.66***	0.81***	−0.17***	−0.21***	0.80***		
12) Self-concept lang T1	0.04	0.03	0.46***	−0.09*	−0.04	−0.13***	−0.15***	0.71***	0.55***	−0.09**	−0.10**	
13) Self-concept lang T2	0.10**	0.04	0.36***	−0.07	−0.05	−0.13***	−0.19***	0.54***	0.71***	−0.10**	−0.11**	0.67***
14) Utility math T1	−0.14**	−0.10**	0.01	0.17***	0.13***	0.52***	0.40***	0.04	0.02	0.44***	0.38***	−0.01
15) Utility math T2	−0.13**	0.07	0.04	0.22***	0.12**	0.40***	0.54***	−0.06	−0.04	0.36***	0.49***	−0.10*
16) Utility lang T1	0.07	−0.08*	0.15***	−0.04	0.10*	0.03	0.04	0.56***	0.38***	−0.03	−0.05	0.40***
17) Utility lang T 2	0.10*	−0.06	0.17***	−0.05	0.09	−0.02	0.02	0.42***	0.54***	−0.09	−0.04	0.34***
18) Career math T1	−0.08*	−0.02	0.09*	0.19***	−0.25***	0.29***	0.24***	−0.08*	−0.10*	0.28***	0.23***	−0.09*
19) Career math T2	−0.12***	−0.03	0.09	0.16**	−0.21***	0.26***	0.23***	−0.10*	−0.07	0.28***	0.27***	−0.08
20) Career lang T1	0.15**	−0.05	0.17***	0.04	−0.25***	−0.04	−0.08	0.11**	0.13***	−0.01	−0.05	0.17***
21) Career lang T2	0.17**	0.01	0.14*	0.08	−0.28***	−0.01	−0.03	0.06	0.07	0.03	−0.02	0.10**

p* < 0.05, ** *p* < 0.01,**p* < 0.001; *Lang achiev, self-reported grade German; Comp. school, comprehensive school ("Integrierte Sekundarschule"); Math achiev, self-reported grade Mathematics; Int math T1/T1, Intrinsic value in mathematics at Time 1/Time 2; Intl ang T1/T2, Intrinsic value in language arts at Time 1/Time 2; Self-concept maths T1/T2, Self-concept in mathematics at Time 1/Time 2; Self- concept lang T1/T2, Self-concept in language arts at Time 1/Time 2; Utility maths T1/T2, Utility value in mathematics at Time 1/Time 2; Utility lang T1/T2, Utility value in language arts at Time 1/Time 2; Career math T1/T2, Career plans in math-related fields at Time 1/ Time 2; Career lang T1T/2, Career plans in language domain at Time 1/Time 2.*

TABLE 3 | Intercorrelations between the study variables – continuation of **Table 2**.

	13	14	15	16	17	18	19	20	21
14) Utility maths T1	−0.01								
15) Utility maths T2	−0.07	0.59***							
16) Utility lang T1	0.31***	0.18***	0.07						
17) Utility lang T2	0.40***	0.09*	0.09*	0.57***					
18) Career math T1	−0.10*	0.20***	0.12*	−0.11*	−0.10*				
19) Career math T2	−0.13***	0.17***	0.17***	−0.12**	−0.08*	0.63***			
20) Career lang T1	0.15***	−0.11**	−0.12**	0.11***	0.16***	−0.08*	−0.10*		
21) Career lang T2	0.13**	−0.04	−0.10*	0.13***	0.16***	−0.03	−0.06	0.63***	

**p* < 0.05, ***p* < 0.01, ****p* < 0.001; *Utility lang T1/T2, Utility value in language arts at Time 1/Time 2; Career math T1/T2, career plans in math-related fields at Time 1/Time 2; Career lang T1T/2, career plans in language domain at Time 1/ Time 2.*

mathematics at Time 1 (β = −0.22, *SE* = 0.03, $p < 0.001$) and aspired to occupations that required lower levels of math knowledge at Time 1 (β = −0.12, *SE* = 0.04, $p = 0.001$). Controlling for achievement differences in school grades, girls reported lower levels of self-concept in language arts at Time 1 compared to boys (β = −0.07, *SE* = 0.03, $p = 0.005$), but aspired to occupations that required higher levels of knowledge in language arts than boys (β = 0.10, *SE* = 0.04, $p = 0.007$), in partial support of *Hypothesis 2*. Notably, correlational patters in **Tables 2**, **3** are fully consistent with *Hypothesis 2*, so that the negative predictive effect of gender on self-concept suggests a larger discrepancy between achievement and self-evaluated abilities for girls than for boys.

Model 1 reveals *positive same-domain effects* but *Hypothesis 3* was supported only in the math domain and not in the language arts domain. These positive same-domain effects were unidirectional from self-concept at Time 1 to career plans at Time 2: Although students' self-concept in mathematics at Time 1 (ψ = 0.17, *SE* = 0.04, $p < 0.001$) and in language arts at Time 1 (ψ = 0.12, *SE* = 0.04, $p = 0.001$) were significantly and positively correlated with career plans in the respective domain within time,

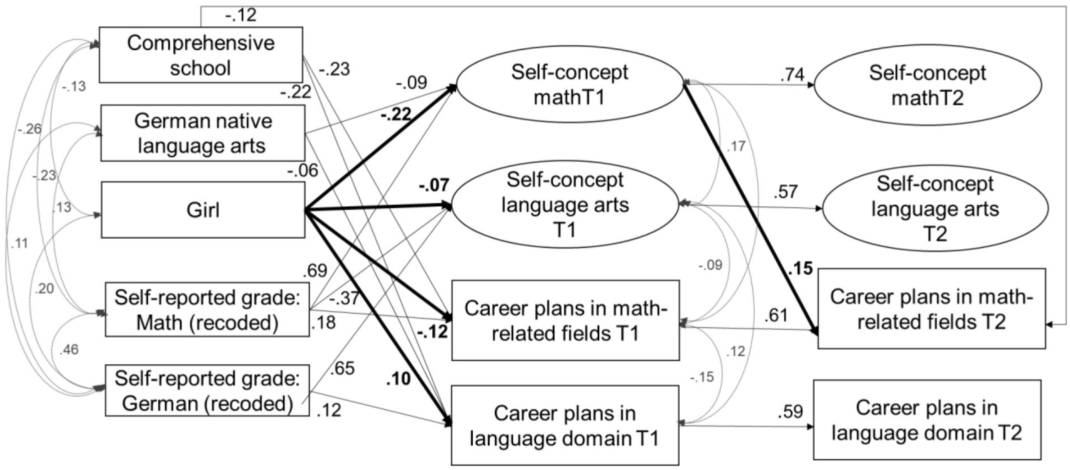


FIGURE 2 | Model 1 – Relations among academic self-concept and career plans in math and language. Standardized and significant ($p < 0.05$) coefficients are depicted.

TABLE 4 | Model 1, Part I: Relations between career plans and academic self-concept.

Variable	Self-concept math T1			Self-concept lang T1			Self-concept math T2			Self-concept lang T2		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Girls	**−0.22**	**0.03**	**<0.001**	**−0.07**	**0.03**	**0.005**	−0.05	0.03	0.076	0.04	0.03	0.121
German native	**−0.09**	**0.03**	**0.001**	0.01	0.03	0.815	0.05	0.03	0.092	0.01	0.03	0.690
Math achiev	**0.69**	**0.03**	**<0.001**	**−0.37**	**0.04**	**<0.001**	0.06	0.04	0.119	−0.02	0.05	0.700
Lang achiev	−0.04	0.04	0.254	**0.65**	**0.04**	**<0.001**	0.03	0.04	0.368	0.09	0.05	0.064
Comp. school	−0.01	0.03	0.964	0.04	0.04	0.360	0.03	0.03	0.310	−0.01	0.03	0.906
Career math T1							0.04	0.03	0.175	−0.03	0.03	0.295
Career lang T1							−0.01	0.02	0.121	0.04	0.02	0.121
Self-concept math T1							**0.74**	**0.03**	**<0.001**	−0.05	0.05	0.379
Self-concept lang T1							−0.04	0.04	0.310	**0.61**	**0.06**	**<0.001**

N = 1117; German native, German native language; Math achiev, Self-reported grade Mathematics (recoded); Lang achiev, Self-reported grade German (recoded); Comp school, Comprehensive school ("Integrierte Sekundarschule"); Career math T1, career plans in math-related fields at Time 1; Career lang 1, career plans in language domain at Time 1; Self-concept math T1/T2, Self-concept in mathematics at Time 1/Time 2; self-concept lang T1/T2, self-concept in language arts at Time 1/Time 2. Coefficients which are significant at least at p < 0.05 are depicted in bold.

TABLE 5 | Model 1, Part II: Relations between academic self-concept and career plans.

	Career math T1			Career lang T1			Career math T2			Career lang T2		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Girls	**−0.12**	**0.04**	**<0.001**	**0.10**	**0.04**	**0.007**	−0.05	0.03	0.121	0.07	0.05	0.138
German native	−0.06	0.04	0.090	**−0.06**	**0.03**	**0.039**	−0.02	0.04	0.682	0.04	0.04	0.365
Math achiev	**0.18**	**0.05**	**<0.001**	−0.05	0.04	0.196	−0.09	0.05	0.090	0.04	0.06	0.578
Lang achiev	−0.02	0.03	0.567	**0.12**	**0.05**	**0.010**	0.06	0.05	0.225	−0.04	0.05	0.444
Comp school	**−0.23**	**0.03**	**<0.001**	**−0.22**	**0.03**	**<0.001**	−0.09	0.05	0.058	**−0.12**	**0.04**	**0.005**
Self math T1							**0.15**	**0.04**	**0.001**	0.03	0.06	0.574
Self lang T1							−0.04	0.04	0.357	0.03	0.04	0.500
Career math T1							**0.57**	**0.05**	**<0.001**	−0.02	0.04	0.634
Career lang T1							−0.06	0.05	0.198	**0.59**	**0.06**	**<0.001**

N = 1117; German native, German native language; Math achiev, Self-reported grade Mathematics (recoded); Lang achiev, Self-reported grade German (recoded); Comp school, Comprehensive school ("Integrierte Sekundarschule"); Career math T1/T2, career plans in math-related fields at Time 1/ Time 2; Career lang T1/T2, career plans in language domain at Time 1/ Time 2; Self-concept math T1, Self-concept in mathematics at Time 1; self-concept lang T1, self-concept in language arts at Time 1. Coefficients which are significant at least at p < 0.05 are depicted in bold.

we did not identify significant predictive effects of self-concept in language arts on career plans in language arts across time. Only in mathematics, self-concept at Time1 significantly and positively predicted career plans at Time 2 (β = 0.15, *SE* = 0.04, $p = 0.001$).

Partially confirming our expectations (*Hypothesis 4*), our results also show some *negative cross-domain effects,* but only within time: Students' self-concept in language arts at Time 1 was significantly and negatively correlated with career plans in math-related fields at Time 1 (ψ = –0.09, *SE* = 0.04, $p = 0.02$). We did not identify significant cross-domain effects between academic self-concept and career plans across time.

Within-time relations suggested positive same-domain effects of achievement on self-concept in mathematics and language arts, and negative cross-domain effects of mathematics achievement on students' self-concept in language arts, however, we did not find such cross-domain effects across time (*Hypothesis 5*). Students' beliefs were relatively stable, which may explain the lack of significant longitudinal associations. Although we did not find direct cross-domain effects across time, we were able to identify indirect cross-domain effects of students' school grade in mathematics at Time 1 on their ability self-concept in language arts at Time 2, mediated via self-concept in language arts at Time 1 (β = –0.20, *SE* = 0.04, $p < 0.001$; 95% CI [–0.27 –0.13]).

In accordance with our expectations (*Hypothesis 6*), there was a significant and indirect effect from student gender to student career plans in math-related fields at Time 2 through student mathematics self-concept – girls reported lower mathematics self-concepts than boys at Time 1, which in turn partially explained their low math-related career plans, β = –0.032, *SE* = 0.01, $p = 0.02$; 95% CI [–0.05 – 0.001].

The following pattern of results emerged for included control variables. Compared to students whose mother tongue was not German, native speakers of German reported lower levels in mathematics self-concept at Time 1 (β = –0.09, *SE* = 0.03, $p = 0.001$). Students in comprehensive schools reported lower career plans in language-related domains at both time points (Time 1: β = −0.12, *SE* = 0.04, $p = 0.005$; Time 2: β = –0.22, *SE* = 0.03, $p < 0.001$) and lower career plans in math-related domains at Time 1 (Time 1: β = −0.23, *SE* = 0.03, $p < 0.001$; Time 2: β = –0.09, *SE* = 0.05, $p = 0.058$) than students in academic track schools. Students' mathematics achievement at Time 1 positively predicted their mathematics self-concept at Time 1 (β = 0.69, *SE* = 0.03, $p < 0.001$) and their math-related career plans at Time 1 (β = 0.18, *SE* = 0.05, $p < 0.001$). Students' achievement in language arts at Time 1 positively predicted their self-concept of ability in language arts at Time1 (β = 0.65, *SE* = 0.04, $p < 0.001$) as well as their career plans in the language arts domain at Time 1 (β = 0.12, *SE* = 0.05, $p = 0.010$). The stability of students' academic self-concept in both mathematics (β = 0.74, *SE* = 0.03, $p < 0.001$) and language arts (β = 0.61, *SE* = 0.06, $p < 0.001$) was relatively high.

The model explained significant amounts of variance in career plans in math-related fields (T1: 10.4%; T2: 42.3%), in language arts-related career plans (T1: 9.1%; T2: 41.3%), as well as in students' mathematics self-concept (T1: 48.1%; T2: 65.3%) and language arts self-concept (T1: 32.1%; T2: 46.7%).

Model 2: Utility Value and Career Plans Model

The model showed a good fit to the data, $\chi^2(123) = 152.89$, CFI = 0.99, TLI = 0.99, RMSEA = 0.02, SRMR = 0.02. Standardized and significant coefficients for this model are reported in **Figure 3**. The standardized coefficients of this model are reported in **Tables 6**, **7**. In line with our expectations (*Hypotheses 1*), girls reported comparatively lower levels of utility value in mathematics at Time 1 (β = –0.13, *SE* = 0.03, $p < 0.001$) as well as lower career plans in math-related fields at both time points (Time 1: β = –0.11, *SE* = 0.04, $p = 0.001$; Time 2: β = –0.08, *SE* = 0.04, $p = 0.038$). Also in line with our

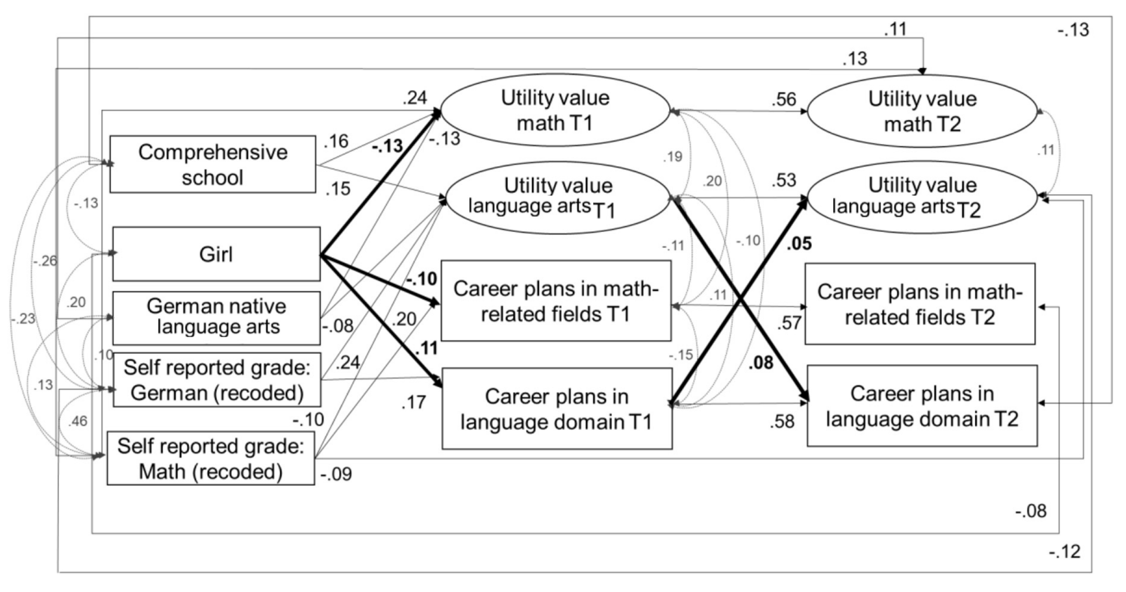


FIGURE 3 | Model 2 – Relations among utility value and career plans in math and language. Standardized and significant ($p < 0.05$) coefficients are depicted.

TABLE 6 | Model 2, Part I: Relations between career plans and utility value.

Variable	Utility math T1			Utility lang T1			Utility math T2			Utility lang T2		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*P*
Girls	**−0.13**	**0.03**	**<0.001**	0.06	0.04	0.168	−0.03	0.04	0.441	0.03	0.03	0.437
German native	**−0.13**	**0.04**	**<0.001**	**−0.08**	**0.04**	**0.041**	**0.11**	**0.03**	**0.001**	−0.02	0.04	0.507
Math achiev	**0.24**	**0.04**	**<0.001**	**−0.11**	**0.04**	**0.007**	**0.13**	**0.04**	**0.002**	**−0.09**	**0.03**	**0.004**
Lang achiev	−0.02	0.04	0.629	**0.24**	**0.04**	**<0.001**	−0.01	0.04	0.838	**0.12**	**0.04**	**0.005**
Compr school	**0.16**	**0.05**	**0.001**	**0.15**	**0.04**	**0.001**	0.07	0.04	0.066	0.03	0.03	0.317
Career math T1							−0.01	0.04	0.704	−0.05	0.03	0.130
Career lang T1							−0.04	0.03	0.114	**0.05**	**0.03**	**0.037**
Utility math T1							**0.56**	**0.04**	**<0.001**	0.02	0.04	0.605
Utility lang T1							−0.02	0.04	0.598	**0.53**	**0.04**	**<0.001**

N = 1117; German native, German native language; Math achiev, Self-reported grade Mathematics (recoded); Lang achiev, Self-reported grade German (recoded); Comp school, Comprehensive school ("Integrierte Sekundarschule"); Career math T1, career plans in math-related fields at Time 1; Career lang T1, career plans in language domain at Time 1; Utility math T1/T2, Utility value in mathematics at Time 1/Time 2 Utility lang T1/T2, Utility value in language arts at Time 1/Time 2. Coefficients which are significant at least at p < 0.05 are depicted in bold.

TABLE 7 | Model 2, Part II: Relations between utility value and career plans.

	Career math T1			Career lang T1			Career math T2			Career lang T2		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*P*	β	*SE*	*P*
Girls	**−0.11**	**0.04**	**0.001**	**0.10**	**0.04**	**0.011**	**−0.08**	**0.04**	**0.038**	0.06	0.05	0.173
German native	−0.06	0.04	0.083	**−0.06**	**0.03**	**0.033**	0.05	0.04	0.220	−0.03	0.04	0.446
Math achiev	**0.18**	**0.05**	**<0.001**	−0.05	0.04	0.176	0.01	0.05	0.909	0.05	0.04	0.216
Lang achiev	−0.02	0.03	0.491	**0.13**	**0.05**	**0.007**	0.04	0.05	0.429	−0.05	0.05	0.264
Comp school	**−0.23**	**0.03**	**<0.001**	**−0.22**	**0.03**	**<0.001**	−0.08	0.05	0.232	**−0.13**	**0.04**	**0.001**
Utility math T1							0.06	0.04	0.163	0.03	0.05	0.486
Utility lang T1							−0.06	0.03	0.054	**0.08**	**0.04**	**0.047**
Career math T1							**0.57**	**0.05**	**<0.001**	−0.02	0.04	0.729
Career lang T1							−0.05	0.05	0.291	**0.58**	**0.04**	**<0.001**

N = 1117; Career math T1/T2, career plans in math-related fields at Time 1/ Time 2; Career lang T1/T2, career plans in language domain at Time 1/ Time 2; German native, German native language; Math achiev, Self-reported grade Mathematics (recoded); Lang achiev, Self-reported grade German (recoded); Comp school, Comprehensive school ("Integrierte Sekundarschule"); Utility math T1, Utility value in mathematics at Time 1; Utility lang T1 = Utility value in language arts at Time 1. Coefficients which are significant at least at p < 0.05 are depicted in bold.

assumptions (*Hypothesis 2*), girls reported higher career plans in the language domain at Time 1 ($\beta = 0.10$, $SE = 0.04$, $p = 0.011$) compared to boys.

Consistent with *Hypothesis 3*, we identified *positive same domain effects*, but only for language arts and not for mathematics: We identified positive same-domain effects between utility value and career plans across time for language arts; utility value in language arts at Time 1 positively predicted career plans in language-related domains at Time 2 ($\beta = 0.08$, $SE = 0.04$, $p = 0.047$). Our assumptions about *cross-domain effects* between motivational beliefs and career plans (*Hypothesis 4*) were not confirmed for utility value longitudinally: neither utility value in language arts at Time 1 predicted career plans in mathematics at Time 2 ($\beta = -0.02$, $SE = 0.03$, $p = 0.054$) nor did utility value in mathematics at Time 1 predict career plans in language-related domains at Time 2 ($\beta = 0.03$, $SE = 0.05$, $p = 0.486$).

In line with our expectations (*Hypothesis* 5), we found positive same-domain effects of achievement at Time 1 on utility value at Time 2 for both mathematics and language arts, and a negative cross-domain effect of mathematics achievement at Time 1 on students' utility value in language arts at Time 2. Although we did not find direct negative cross-domain effects for the relation between self-reported achievement and utility value across time, we were able to identify indirect cross-domain effects of students' self-reported grade in mathematics at Time 1 on utility value in language arts at Time 2 via utility value in language arts at Time 1 ($\beta = -0.05$, SE = 0.02, $p = 0.007$; 95% CI [−0.08 −0.01]).

Contrary to expectations about gender-specific (same-domain and cross-domain) motivational processes (*Hypothesis 6*), we did not find any significant indirect effects from gender on career plans via utility value (language arts: $\beta = 0.01$, $SE = 0.01$, $p = 0.28$; 95% CI [−0.05 0.17]; mathematics: $\beta = -0.01$, $SE = 0.01$, $p = 0.18$; 95% CI [−0.38 0.07]).

With regard to our control variables, we found that, compared to students whose mother tongue was not German, students whose mother tongue was German reported lower utility value in mathematics ($\beta = -0.13$, $SE = 0.04$, $p < 0.000$) and language arts ($\beta = -0.08$, $SE = 0.04$, $p = 0.041$) at Time 1, but higher utility value in mathematics at Time 2 (β = 0.11, *SE* = 0.03, $p < 0.001$). Students in comprehensive schools reported higher utility value of language-related domains (β = 0.15, *SE* = 0.04, $p < 0.001$) and math (β = 0.16, *SE* = 0.05, $p < 0.001$) at Time 1. Students' mathematics achievement at Time 1 positively predicted their utility value in mathematics at Time 1 (β = 0.24, *SE* = 0.04, $p < 0.001$) and Time 2 (β = 0.13, *SE* = 0.04, $p = 0.002$), as well as their math-related career plans at Time 1 (β = 0.18, *SE* = 0.05, $p < 0.001$). Students' achievement in language arts at Time 1 positively predicted their utility value in language arts at Time 1 (β = 0.24, *SE* = 0.04, $p < 0.001$) and Time 2 (β = 0.12, *SE* = 0.04, $p = 0.005$) as well as their career plans in the language arts domain at Time 1 (β = 0.13, *SE* = 0.05, $p = 0.007$). The stability of reported utility value in both mathematics (β = 0.56, *SE* = 0.04, $p < 0.001$) and language arts (β = 0.53, *SE* = 0.04, $p < 0.001$) was relatively high across both time points.

The model explained significant amounts of variance in career plans in math-related fields (T1: 10.3%; T2: 40.7%), career plans in language arts-related fields (T1: 9.1%; T2: 41.9%), as well as in mathematics utility value (T1: 9.7%; T2: 38.3%) and language arts utility value (T1: 6.7%; T2: 35.0%).

Model 3: Intrinsic Value and Career Plans Model

The model showed a good fit to the data, $\chi^2(124) = 192.19$, CFI = 0.99, TLI = 0.98, RMSEA = 0.02, SRMR = 0.02. Standardized and significant coefficients for this model are reported in **Figure 4**. The standardized coefficients of this model are reported in **Tables 8**, **9**. In line with our assumptions (*Hypothesis 1*), girls relative to boys reported lower levels of intrinsic value in mathematics at Time 1 (β = −0.18, *SE* = 0.03, $p < 0.001$) and were less likely to report career plans in math-related fields at Time 1 (β = −0.11, *SE* = 0.04, $p = 0.001$). Also in line with our expectations (*Hypothesis 2*), girls were more likely than boys to report career plans in language arts domains at Time 1 (β = 0.10, *SE* = 0.04, $p = 0.009$). The stability of students' intrinsic value in both mathematics (β = 0.63, *SE* = 0.04, $p < 0.001$) and language arts (β = 0.65, *SE* = 0.04, $p < 0.001$) was high across the two time points.

Our assumptions about *positive same-domain effects*: (*Hypothesis 3*) were –, similarly, to our results for self-concept – only confirmed for the domain of mathematics, but not for language arts: Intrinsic value in mathematics at Time 1 significantly and positively predicted career plans in math-related fields at Time 2 (β = 0.08, *SE* = 0.03, $p = 0.020$). The effect was unidirectional as intrinsic value positively predicted career plans (and not vice versa).

In accordance with our expectations (*Hypothesis 4*), we also identified *cross-domain effects*: Intrinsic value in mathematics at Time 2 significantly and negatively predicted by career plans in language arts at Time 1 (β = −0.05, *SE* = 0.02, $p = 0.047$). This effect was unidirectional: interestingly, career plans predicted subsequent intrinsic value (and not vice versa).

Partially confirming our expectations (*Hypothesis 5*), we found positive same-domain effects of mathematics achievement at Time 1 on mathematics intrinsic value at Time 2, and of achievement in language-arts at Time 1 on intrinsic value in language-arts at Time 2. Although we did not find direct negative cross-domain effects for the relation between self-reported achievement and intrinsic value across time, we were able to identify indirect cross-domain effects from self-reported grade in mathematics at Time 1 on intrinsic value in language arts at Time 2 via intrinsic value in language arts at Time 1 (β = −0.22, *SE* = 0.02, $p < 0.001$; 95% CI

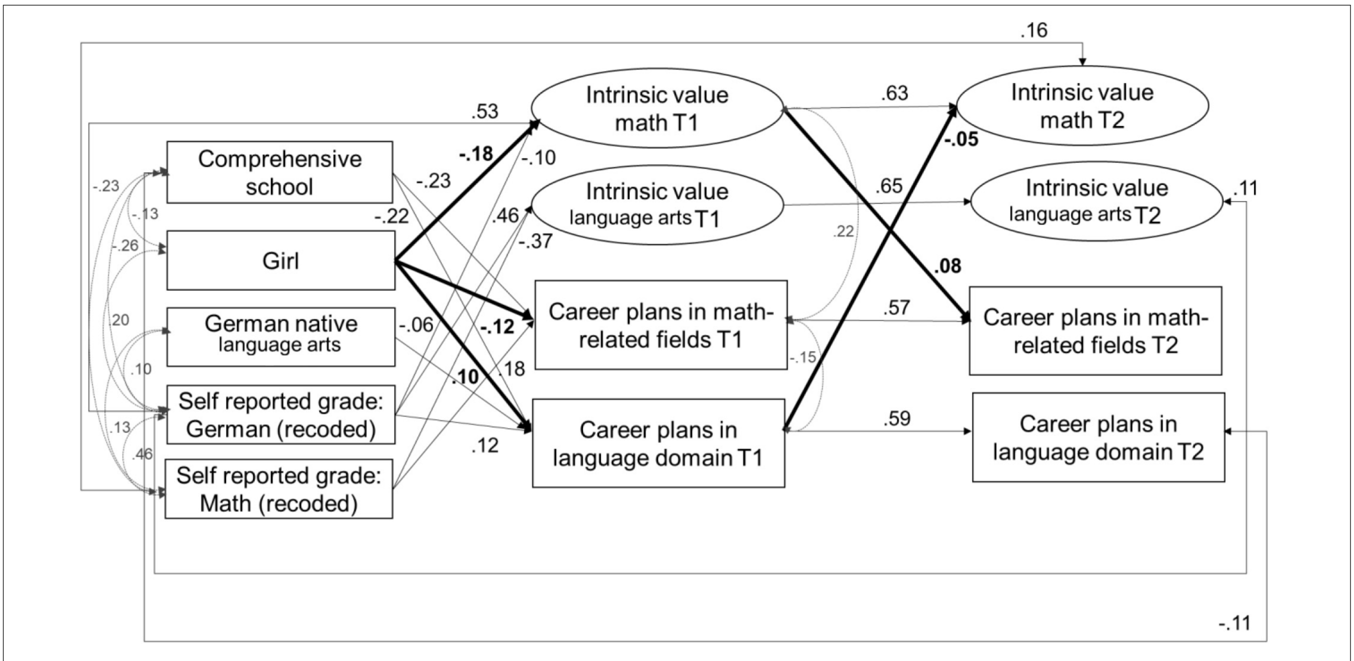


FIGURE 4 | Model 3 – Relations among intrinsic value and career plans in math and language. Standardized and significant ($p < 0.05$) coefficients are depicted.

TABLE 8 | Model 3, Part I: Relations between career plans and intrinsic value.

Variable	Intrinsic math T1			Intrinsic lang T1			Intrinsic math T2			Intrinsic lang T2		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*P*
Girls	**−0.18**	**0.03**	**<0.001**	0.05	0.03	0.125	0.01	0.03	0.871	0.02	0.03	0.491
German native	−0.06	0.03	0.093	−0.01	0.03	0.945	0.02	0.03	0.405	0.02	0.04	0.549
Math achiev	**0.53**	**0.04**	**<0.001**	**−0.36**	**0.02**	**<0.001**	**0.16**	**0.05**	**0.002**	−0.07	0.04	0.091
Lang achiev	**−0.10**	**0.04**	**0.017**	**0.47**	**0.03**	**<0.001**	−0.01	0.05	0.862	**0.11**	**0.05**	**0.025**
Compr school	0.01	0.04	0.971	0.06	0.05	0.228	0.01	0.03	0.769	0.01	0.03	0.709
Career math T1							0.03	0.03	0.414	−0.03	0.03	0.391
Career lang T1							**−0.05**	**0.02**	**0.047**	0.03	0.03	0.213
Intrinsic math T1							**0.63**	**0.04**	**<0.001**	−0.04	0.04	0.298
Intrinsic lang T1							−0.01	0.04	0.917	**0.65**	**0.03**	**<0.001**

N = 1117; German native, German native language; Math achiev, Self-reported grade Mathematics (recoded); Lang achiev, Self-reported grade German (recoded); Comp school, Comprehensive school ("Integrierte Sekundarschule"); Career math T1, career plans in math-related fields at Time 1; Career lang 1, career plans in language domain at Time 1; Intrinsic math T1/ T2, Intrinsic value in mathematics at Time 1/Time 2; Intrinsic lang T1/T2, Intrinsic value in language arts at Time 1/Time 2. Coefficients which are significant at least at $p < 0.05$ are depicted in bold.

TABLE 9 | Model 3, Part II: Relations between intrinsic value and career plans.

	Career math T1			Career lang T1			Career math T2			Career lang T2		
	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Girls	**−0.11**	**0.04**	**0.001**	**0.10**	**0.04**	**0.011**	−0.06	0.04	0.076	0.06	0.05	0.185
German native	−0.06	0.04	0.083	**−0.06**	**0.03**	**0.033**	0.03	0.04	0.477	0.04	0.04	0.344
Math achiev	**0.18**	**0.05**	**<0.001**	−0.05	0.04	0.176	−0.03	0.05	0.470	0.05	0.05	0.351
Lang achiev	−0.02	0.03	0.491	**0.13**	**0.05**	**0.007**	0.06	0.05	0.222	−0.03	0.05	0.546
Comp school	**−0.23**	**0.03**	**<0.001**	**−0.22**	**0.03**	**<0.001**	−0.09	0.05	0.058	**−0.11**	**0.04**	**0.007**
Intrinsic math T1							**0.08**	**0.03**	**0.020**	0.03	0.05	0.486
Intrinsic lang T1							−0.06	0.04	0.173	0.01	0.05	0.844
Career math T1							**0.57**	**0.05**	**<0.001**	−0.02	0.04	0.640
Career lang T1							−0.05	0.05	0.263	**0.59**	**0.05**	**<0.001**

N = 1117; Career math T1/T2, career plans in math-related fields at Time 1/ Time 2; Career lang T1/T2, career plans in language domain at Time 1/Time 2; German native, German native language; Math achiev, Self-reported grade Mathematics (recoded); Lang achiev, Self-reported grade German (recoded); Comp school, Comprehensive school ("Integrierte Sekundarschule"); Intrinsic math T1, Intrinsic value in mathematics at Time 1; Intrinsic lang T1, Intrinsic value in language arts at Time 1. Coefficients which are significant at least at $p < 0.05$ are depicted in bold.

[− 0.27 − 0.17]) and from self-reported grade in German at Time 1 on mathematics intrinsic value at Time 2 via mathematics intrinsic value at Time 1 (β = -0.05, *SE* = 0.02, *p* = 0.03; 95% CI [−0.09 −0.01]).

Gender-specific motivational processes (see *Hypothesis 6*) were identified only for mathematics: There was a significant and indirect effect from student gender to student career plans in math-related fields at Time 2, which was mediated via students' intrinsic valuing of mathematics – girls reported lower mathematics intrinsic value than boys at Time 1, which in turn corresponded to a lower probability of pursuing math-related careers, β = −0.014, *SE* = 0.01, *p* = 0.03; 95% CI [− 0.03 − 0.01]. This effect size, however, was very small.

With regard to our control variables, we found that compared to students whose mother tongue was not German, students whose mother tongue was German had a lower likelihood of striving for careers related to language arts (β = −0.06, *SE* = 0.03, *p* = 0.033) at Time 1. Students in comprehensive schools were less likely to strive for careers in math at Time 1 (Time 1: β = −0.23, *SE* = 0.03, $p < 0.001$) or for careers related to language arts at Time 1 and Time 2 (Time 1: β = −0.22, *SE* = 0.03, $p < 0.001$; Time 2: β = −0.11, *SE* = 0.04, *p* = 0.007). Students' mathematics achievements at Time 1 positively predicted their intrinsic value in mathematics at Time 1 (β = 0.53, *SE* = 0.04, $p < 0.001$) and Time 2 (β = 0.16, *SE* = 0.05, *p* = 0.002), as well as math-related career plans at Time 1 (β = 0.18, *SE* = 0.05, $p < 0.001$). Students' achievement in language arts at Time 1 positively predicted their intrinsic value in language arts at Time 1 (β = 0.47, *SE* = 0.03, $p < 0.001$) and Time 2 (β = 0.11, *SE* = 0.05, *p* = 0.003), as well as their language arts-related career plans at Time 1 (β = 0.12, *SE* = 0.05, *p* = 0.007). Students' mathematics achievements at Time 1 negatively predicted their intrinsic valuing of language arts at Time1 (β = −0.36, *SE* = 0.03, $p < 0.001$). Students' achievement in language arts at Time 1 negatively predicted their intrinsic valuing of mathematics at Time 1 (β = − 0.10, *SE* = 0.04, *p* = 0.012). Career plans in mathematics at Time 1 were significantly and positively correlated with intrinsic value in mathematics at Time 1 (ψ = 0.22, *SE* = 0.04, $p < 0.001$) and were significantly and negatively correlated with career plans in language arts at Time 1 (ψ = –0.15, *SE* = 0.03, $p < 0.001$). The model explained significant amounts of variance in career plans in math-related fields (T1: 10.2%; T2: 41.5%), in language arts-related fields (T1: 9.1%; T2: 41.1%), as well as in intrinsic value in math (T1: 27.3%; T2: 53.5%) and intrinsic value in language arts (T1: 20.5%; T2: 51.1%).

In a set of supplemental analyses reported in **Appendix D**, we included math- and language arts-related academic self-concept, intrinsic and utility values, and career plans in one model. The results of this model show negative cross-domain effects of self-concept on task values. Self-concept in mathematics at Time 1 negatively predicts intrinsic value in language-arts at Time 2 (β = -0.18, *SE* = 0.07, p = 0.010). Self-concept in language-arts at Time 1 negatively predicts intrinsic value in mathematics (β = -0.13, *SE* = 0.04, p = 0.003) and utility value in mathematics (β = -0.10, *SE* = 0.04, p = 0.019) both at Time 2. Thus, our findings confirm the key role of academic self-concept in dimensional comparison effects and show that these effects apply to students' task values as well (see Gaspard et al., 2018).

Our additional results also show positive same-domain effects of self-concept and intrinsic value. Self-concept in mathematics at Time 1 positively predicts intrinsic value in mathematics at Time 2. Intrinsic value in language-arts at Time 1 positively predicts self-concept in language-arts (β = 0.16, *SE* = 0.06, p = 0.010) and utility value in language-arts (β = 0.10, *SE* = 0.05, p = 0.036) both at Time 2. Thus, the same-domain effects that we show confirm reciprocal links between academic self-concept and intrinsic value within-domains.

DISCUSSION

This study investigated whether dimensional comparison processes regarding girls' and boys' motivational beliefs might contribute to gendered career plans in mathematics and language arts. Furthermore, we investigated whether motivational beliefs and career plans were reciprocally related across two academic domains. Our findings revealed systematic gender differences in *same domain effects* in mathematics: girls' comparatively lower mathematics self-concept and intrinsic value predicted a lower likelihood of striving for a math-related career. Furthermore, and contrary to expectations, *cross-domain effects* were not related to gender-specific career plans, with only one exception. Girls' lower levels of intrinsic value in mathematics corresponded to a higher likelihood of striving for a career in language-related fields, which subsequently predicted lower levels of intrinsic value in mathematics. This finding points to a need to address both gender-specific motivational beliefs and gender-specific career plans in school when aiming to enhance more gender equality in girls' and boys' occupational choices.

Gendered Motivational Beliefs and Career Plans in Math and Language-Related Domains

Our hypotheses regarding gender differences in motivational beliefs and career plans of students in math and language-related domains were mostly confirmed. Consistent with prior evidence (Marsh and Yeung, 1998; Watt, 2004; Watt et al., 2012; Gaspard et al., 2015), girls reported lower academic self-concept, intrinsic and utility values in mathematics than boys, and were less likely than boys to strive for careers in math-intensive fields (*Hypothesis 1*). This is notable, given that girls and boys did not differ substantially in terms of self-reported mathematics achievement. Thus, despite gender equality in grades in mathematics, girls felt less competent than did boys in math. Notably, grades are quite important in this context, because they are one of the main factors determining access to higher education, including in the fields of math and science. Yet, negative self-beliefs reduce the likelihood of pursuing math-intensive careers, even when access is possible. Stereotype threat effects might contribute to this discrepancy between achievement feedback and self-perceptions (Steele, 1997). If teachers or parents communicate, for example, through their achievement-related expectations and feedback behaviors that mathematics is a subject that is "typically male" (Tiedemann, 2002; Tenenbaum and Leaper, 2003), girls can feel less competent in the subject despite their high achievement. Our findings point to the need to foster girls' self-concept in mathematics, for example, by providing them with positive feedback about their intellectual performance in math classes (Dweck, 1978).

Our hypotheses about gender differences in language-related fields were only partly confirmed. Without taking into account gender differences in self-reported grades, boys reported lower self-concept and lower utility value than girls in language arts at the middle of the school year, but not at the beginning of the school year. Furthermore, boys reported comparatively lower levels of intrinsic value in language arts. However, when differences in achievement were controlled (girls had higher grades in language arts than boys), girls reported lower self-concept in language arts than boys already at the beginning of the school year. Controlling for grades, there was no longer a statistically significant effect of gender on utility value or intrinsic value in language arts. This finding extends previous research, which has shown that boys report lower levels of interest and competence beliefs in language arts (Jacobs et al., 2002; Watt, 2004). Gender differences in the language arts domain in the present study appeared to be explained to a large degree by differences in teacher-graded achievement, with the exception of differences in students' self-concept. Analyses of motivational differences between girls and boys need to take into account achievement differences as well.

Same-Domain and Cross-Domain Effects

In line with our expectations (*Hypothesis 3*) and based on EVT (Eccles et al., 1983), we found positive same-domain associations between motivational beliefs and career plans. Our expectations were only partially confirmed as we found positive same-domain effects mainly for mathematics. Unidirectional effects were identified showing that academic self-concept and intrinsic value in mathematics predicted subsequent career plans in mathematics-related fields but not vice versa. This result deviates from the reciprocal effects reported by Lauermann et al. (2017). However, whereas Lauermann et al. (2017) asked students about their subjective probability of pursuing careers in math and science, in our study, the importance of math was inferred from open-ended reports of desired careers. If students are not fully aware of the degree to which mathematics is important for their career choice, the predictive power of such choices for subsequent motivations may be reduced. Accordingly, the degree to which students are aware of academic requirements that are relevant for attaining particular careers may play an important moderating role in these reciprocal links.

Most studies that have examined same-domain effects among motivational beliefs and career plans have focused on the domain of mathematics (Lauermann et al., 2017; Lazarides et al., 2017; Wang, 2012; Watt et al., 2012) or science (Guo et al., 2017). Only few studies have examined same-domain effects among motivational beliefs and career plans in domains that are stereotyped as typically female (Lauermann et al., 2015). Building on such previous findings, our study showed that when focusing on the single components of the task value construct, we only found such positive same domain effects for utility value in the domain of language arts. Thus, we find differential effects among the task value components depending on the domain at hand. In mathematics, intrinsic value, and academic self-concept were important predictors of subsequent career plans, whereas in language arts, utility value emerged as a significant predictor of subsequent language-related career plans. Mathematics is often stereotyped as being difficult and not interesting (for mathematics: Watson et al., 1994), which may explain why students' ability beliefs and interest emerged as significant predictors in this domain. It may be that only students who are highly interested and who feel highly competent in math might tend to strive for careers in math-intensive fields. By comparison, in language-related fields, in which students in our sample were more likely to feel competent and interested, the utility of the domain was more important for their career plans.

In line with the dimensional comparison theory (Möller and Köller, 2001), we identified a set of negative cross-domain effects (*Hypothesis 4*). Specifically, we found one negative cross-domain effect for language-related career plans that negatively predicted subsequent mathematics intrinsic value. If students strived for a career in language-related domains (e.g., writer, journalist) at the beginning of the school year, they reported lower intrinsic value for math at the middle of the school year. Thus, students' career plans in language-related domains seemed to initiate specialization processes and led to a reduction of interest in domains that would not help students to achieve their goals.

Interestingly, we identified the expected positive same-domain effects across time for the relation between achievement and task values in mathematics and language arts, but not for the relation between achievement and academic self-concept. We also identified a direct negative cross-domain effect of mathematics achievement (Time 1) on utility value in language arts (Time 2), but not for academic self-concept. However, academic self-concept in our study was highly stable from the beginning of the school year to mid-year. Focusing on a longer time span might be necessary to adequately capture changes in students' beliefs over time, as proposed in the I/E model (Marsh, 1986).

Gendered Career Plans and Dimensional Comparison Processes

One central contribution of our study to previous work is that we examined whether and how both same- and cross-domain effects contributed to gendered career plans in mathematics- and language-related fields. Few studies have examined the potential interrelations between student gender and dimensional comparison effects in predicting students' educational (Nagy et al., 2006; Parker et al., 2012; Guo et al., 2017) and career plans and choices (Wang et al., 2013; Lauermann et al., 2015). Their findings showed that dimensional comparison processes can explain gendered educational and career plans in certain domains. Focusing on math vs. language-related career plans, our study only partially confirmed our expectations about gender-specific (same-domain and cross-domain) motivational processes (*Hypothesis 6*). Only for same-domain effects we found a significant indirect effect of gender on career plans through student motivational beliefs. Girls reported lower mathematics self-concept and intrinsic value at time 1 than boys and were subsequently less likely to strive for careers in math-related fields at time 2. We did not find such effects for language-related motivational beliefs mainly because gender differences in motivational beliefs in language-related domains were not as pronounced as they were in mathematics. This might also be an explanation for the different findings in our study compared to previous studies that found such negative cross-domain effects, for example, for gendered task value and career plans in the field of human-services occupations (Lauermann et al., 2015). Another explanation might be the longitudinal design of the current study. Negative cross-domain effects might have emerged if we had not controlled for prior beliefs because both beliefs and career plans were relatively stable. However, the longitudinal design is an important strength of the present study, as it allows us to examine the effects of motivational beliefs on potential changes in students' career plans within the school year (and vice versa), and thus, applies a developmental perspective. A longer period of time, however, may need to be considered to examine how and why these beliefs may influence each other over time.

Our findings of differential effects for math and language-related domains point to a need to consider the interrelations between learning contexts and student characteristics. Thus, especially in mathematics, girls' lower self-concept and intrinsic value seem to be an explanation for gendered career plans in corresponding fields, whereas boys' lower intrinsic value in language arts did not explain their lower likelihood of striving for careers in language-related domains. More studies are needed that investigate the factors that contribute to boys' lower likelihood of striving for careers in fields that are stereotyped as "typically female." Such factors could be related to the matching of the image of these domains and boys' self-identity (Kessels et al., 2006).

Limitations

Our study has several limitations that need to be discussed when interpreting its findings. First, the operationalization of career plans is referring to concrete ideas more than to aspirations as students were asked "What job do you think are you going to have when you are 30 years old?" This needs to be considered when comparing the results of this study to previous studies that asked students, for example, for the likelihood of pursuing a career in a certain field (Lauermann et al., 2017). Furthermore, it is an important question whether language-related careers are specific enough as an outcome variable as many careers require verbal and communication-related skills. There are careers that require substantially higher language skills and knowledge than mathematical skills and knowledge (e.g., media consultant and journalist), so that the category "language-relatedness" can be meaningful for a specific group of occupations. However, other careers that are typically considered to be math-intensive (e.g., astrophysicist or mathematician) require very high levels of math and verbal skills and knowledge. Our data preclude us from examining potential discrepancies between required levels of domains-specific skills and knowledge across occupations and students' subjective beliefs about these occupations. It is also important to note that we were unable to examine potential gaps between students' career aspirations and educational goals. Schneider and Stevenson (1999) point out, for example, that adolescents with clear understandings of the amount of education needed for their aspired careers are more likely to achieve their aspirations. Thus, future studies need to assess not only adolescents' occupational aspirations but also corresponding educational goals. Lastly, cross-lagged panel studies have recently been criticized because of their inability to differentiate between relatively stable between-person differences and within-person developmental processes (Hamaker et al., 2015). A larger number of time points, and potentially a larger longitudinal sample than the one available for this study would be necessary for systematic analyses of longitudinal state- and trait-level differences.

CONCLUSION

Taken together, our findings corroborate previous research (Watt et al., 2012; Lazarides and Watt, 2015; Lauermann et al., 2017) by showing that gender differences in academic motivations contribute to gendered career plans in mathematics (*same-domain effects*), but, we did not find analogous effects for language arts. In addition, negative *cross-domain effects* did not significantly explain gendered career plans in language-related domains. However, language-related career plans negatively predicted students' intrinsic valuing of mathematics, which in turn predicted a decrease in language-related career plans. For educational practice, our findings suggest that it is likely important for teachers to enhance interest and self-concept of girls in mathematics, but also to directly speak with boys and girls about their career plans in specific fields.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Berlin Senate for Education, Science and Research (senbwf) with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the responsible committee at senbwf.

AUTHOR CONTRIBUTIONS

RL conducted the analyses and collected the data, and drafted the manuscript. FL contributed decisively to the development of the manuscript.

FUNDING

The data of this study stemmed from a pilot study for the research project "Teach!" which is supported by the German Research Foundation from 2019–2021 under grant number LA 3522/5-1 to RL.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01243/full#supplementary-material

REFERENCES

Arens, A. K., and Jansen, M. (2016). Self-concepts in reading, writing, listening, and speaking: a multidimensional and hierarchical structure and its generalizability across native and foreign languages. *J. Educ. Psychol.* 108, 646–664. doi: 10.1037/edu0000081

Brunner, M., Lüdtke, O., and Trautwein, U. (2008). The internal/external frame of reference model revisited: incorporating general cognitive ability and general academic self-concept. *Multivar. Behav. Res.* 43, 137–172. doi: 10.1080/00273170701836737

Byrne, B. M. (2004). Testing for multigroup invariance using AMOS graphics: a road less traveled. *Struct. Equa. Model.* 11, 272–300. doi: 10.1207/s15328007sem1102_8

Dickhäuser, O., Reuter, M., and Hilling, C. (2005). Coursework selection: a frame of reference approach using structural equation modelling. *Br. J. Educ. Psychol.* 75, 673–688. doi: 10.1348/000709905x37181

Durik, A. M., Vida, M., and Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: a developmental analysis. *J. Educ. Psychol.* 98, 382–393. doi: 10.1037/0022-0663.98.2.382

Dweck, C. S. (1978). Sex differences in learned helplessness: II. the contingencies of evaluative feedback in the classroom and III. An experimental analysis. *Dev. Psychol.* 14, 258–278. doi: 10.1037//0012-1649.14.3.268

Eccles, J. S. (2005). "Subjective task value and the Eccles et al. model of achievement-related choices," in *Handbook of Competence and Motivation*, eds A. J. Elliot and C. S. Dweck (New York, NY: Guilford), 105–131.

Eccles, J. S. (2009). Who am I and what am I going to do with my life? personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., et al. (1983). "Expectancies, values and academic behaviors," in *Achievement and Achievement Motives: Psychological and Sociological Approaches*, ed. J. T. Spence (San Francisco, CA: Freeman), 75–146.

Eccles, J. S., and Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Pers. Soc. Psychol. Bull.* 21, 215–225. doi: 10.1177/0146167295213003

Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153

Eccles, J. S., Wigfield, A., and Schiefele, U. (1998). "Motivation to succeed," in *Handbook of Child Psychology*, 5 Edn, Vol. 3, ed. N. Eisenberg (New York, NY: Wiley), 1017–1095.

Frenzel, A. C., Goetz, T., Pekrun, R., and Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *J. Res. Adolesc.* 20, 507–537. doi: 10.1111/j.1532-7795.2010.00645.x

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., et al. (2015). More value through greater differentiation: gender differences in value beliefs about math. *J. Educ. Psychol.* 107, 663–677. doi: 10.1037/edu0000003

Gaspard, H., Wigfield, A., Jiang, Y., Nagengast, B., Trautwein, U., and Marsh, H. W. (2018). Dimensional comparisons: how academic track students' achievements are related to their expectancy and value beliefs across multiple domains. *Contemp. Educ. Psychol.* 52, 1–14. doi: 10.1016/j.cedpsych.2017.10.003

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J., and Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: dimensional comparison processes and expectancy-by-value interactions. *Learn. Instruct.* 49, 81–91. doi: 10.1016/j.learninstruc.2016.12.007

Hamaker, E. L., Kuiper, R. M., and Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychol. Methods* 20, 102–116. doi: 10.1037/a0038889

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., and Wigfield, A. (2002). Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Dev.* 73, 509–527. doi: 10.1111/1467-8624.00421

Kenny, D. A. (1975). Cross-lagged panel correlation: a test for spuriousness. *Psychol. Bull.* 82, 887–903. doi: 10.1037/0033-2909.82.6.887

Kessels, U., Rau, M., and Hannover, B. (2006). What goes well with physics? measuring and altering the image of science. *Br. J. Educ. Psychol.* 76, 761–780. doi: 10.1348/000709905x59961

Lauermann, F., Chow, A., and Eccles, J. S. (2015). Differential effects of adolescents' expectancy and value beliefs about math and english on math/science-related and human services-related career plans. *Int. J. Gender Sci. Technol.* 7, 205–228. doi: 10.1111/jora.12218

Lauermann, F., Tsai, Y.-M., and Eccles, J. (2017). Math-related career aspirations and choices within eccles et al.'s expectancy–value theory of achievement-related behaviors. *Dev. Psychol.* 53, 1540–1559. doi: 10.1037/dev0000367

Lazarides, R., Dicke, A.-L., Rubach, C., and Eccles, J. S. (2019). Profiles of motivational beliefs in math: exploring their development, relations to student-perceived classroom characteristics and impact on future career aspirations and choices. *J. Educ. Psychol.*

Lazarides, R., Rubach, C., and Ittel, A. (2017). Adolescents' perceptions of socializers' beliefs, career-related conversations, and motivation in mathematics. *Dev. Psychol.* 53, 525–539. doi: 10.1037/dev0000270

Lazarides, R., and Watt, H. M. G. (2015). Student-perceived mathematics teacher beliefs, math classroom learning environments and gendered math career intentions. *Contemp. Educ. Psychol.* 41, 51–61. doi: 10.1016/j.cedpsych.2014.11.005

MacKinnon, D. P., Fritz, M. S., Williams, J., and Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: program PRODCLIN. *Behav. Res. Methods* 39, 384–389. doi: 10.3758/bf03193007

Marsh, H. W. (1986). Verbal and math self-concepts: an internal/external frame of reference model. *Am. Educ. Res. J.* 23, 129–149. doi: 10.3102/00028312023001129

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Abduljabbar, A. S., Abdelfattah, F., et al. (2015). Dimensional comparison theory: paradoxical relations between self-beliefs and achievements in multiple domains. *Learn. Instruct.* 35, 16–32. doi: 10.1016/j.learninstruc.2014.08.005

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., et al. (2018). The murky distinction between self-concept and self-efficacy: beware of lurking jingle-jangle fallacies. *J. Educ. Psychol.* 111, 331–353. doi: 10.1037/edu0000281

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., and Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Dev.* 76, 397–416. doi: 10.1111/j.1467-8624.2005.00853.x

Marsh, H. W., and Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: gender differences in the development of math and English constructs. *Am. Educ. Res. J.* 35, 705–738. doi: 10.3102/00028312035004705

Meece, J. L., Wigfield, A., and Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *J. Educ. Psychol.* 82, 60–70. doi: 10.1037/0022-0663.82.1.60

Möller, J., and Köller, O. (2001). Dimensional comparisons: an experimental approach to the internal/external frame of reference model. *J. Educ. Psychol.* 93:826. doi: 10.1037//0022-0663.93.4.826

Möller, J., and Marsh, H. W. (2013). Dimensional comparison theory. *Psychol. Rev.* 120, 544–560. doi: 10.1037/a0032459

Möller, J., Pohlmann, B., Köller, O., and Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Rev. Educ. Res.* 79, 1129–1167. doi: 10.3102/0034654309337522

Möller, J., Retelsdorf, J., Köller, O., and Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: an integration of models of relations between academic achievement and self-concept. *Am. Educ. Res. J.* 48, 1315–1346. doi: 10.1037/dev0000393

Muthén, L., and Muthén, B. (1998–2019). *Mplus User's Guide.* Los Angeles, CA: Muthén & Muthén.

Nagy, G., Trautwein, U., Baumert, J., Köller, O., and Garrett, J. (2006). Gender and course selection in upper secondary education: effects of academic self-concept and intrinsic value. *Educ. Res. Eval.* 12, 323–345. doi: 10.1080/13803610600765687

National Center for O*Net Development (2014). *O*NET OnLine*. Available at: http://www.onetonline.org/ (accessed July 17, 2014).

Niepel, C., Brunner, M., and Preckel, F. (2014). The longitudinal interplay of students' academic self-concepts and achievements within and across domains: replicating and extending the reciprocal internal/external frame of reference model. *J. Educ. Psychol.* 106, 1170–1191. doi: 10.1037/a0036307

Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., and Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multicontext study. *Dev. Psychol.* 48, 1629–1642. doi: 10.1037/a0029167

Schneider, B., and Stevenson, D. (1999). The ambitious generation. *Educ. Leadersh.* 57, 22–25.

Steele, C. M. (1997). A threat in the air: how stereotypes shape intellectual identity and performance. *Am. Psychol.* 52, 613–629. doi: 10.1037//0003-066x.52.6.613

Steinmayr, R., and Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW) [construction and validation of a scale for the assessment of subjective task values (SESSW)]. *Diagnostica* 56, 195–211. doi: 10.1026/0012-1924/a000023

Su, R., and Rounds, J. (2015). All STEM fields are not created equal: people and things interests explain gender disparities across STEM fields. *Front. Psychol.* 6:189. doi: 10.3389/fpsyg.2015.00189

Tanaka, J. S. (1993). "Multifaceted conceptions of fit in structural equation models," in *Testing Structural Equation Models*, eds K. A. Bollen and J. S. Long (Newbury Park, CA: Sage), 10–39.

Tenenbaum, H. R., and Leaper, C. (2003). Parent-child conversations about science: the socialization of gender inequities? *Dev. Psychol.* 39, 34–47. doi: 10.1037/0012-1649.39.1.34

Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educ. Stud. Math.* 50, 49–62.

Wang, M.-T. (2012). Educational and career interests in math: a longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Dev. Psychol.* 48, 1643–1657. doi: 10.1037/a0027247

Wang, M.-T., and Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): current knowledge, implications for practice, policy, and future directions. *Educ. Psychol. Rev.* 29, 119–140. doi: 10.1007/s10648-015-9355-x

Wang, M.-T., Eccles, J., and Kenny, S. (2013). Not lack of ability but more choice individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937

Watson, J., McEwen, A., and Dawson, S. (1994). Sixth form A level students' perceptions of the difficulty, intellectual freedom, social benefit and interest of science and arts subjects. *Res. Sci. Technol. Educ.* 12, 43–52. doi: 10.1080/0263514940120106

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th-through 11th-grade Australian students. *Child Dev.* 75, 1556–1574. doi: 10.1111/j.1467-8624.2004.00757.x

Watt, H. M. G. (2016). "Gender and motivation," in *Handbook of Motivation at School*, eds K. Wentzel and D. Miele (New York, NY: Routledge), 320–339.

Watt, H. M. G., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., and Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: a comparison of samples from Australia, Canada, and the United States. *Dev. Psychol.* 48, 1594–1611. doi: 10.1037/a0027838

Wigfield, A., and Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps1999.1015

Williams, J., and MacKinnon, D. P. (2008). Resampling and distribution of the product methods for testing indirect effects in complex models. *Struct. Equ. Model.* 15, 23–51. doi: 10.1080/10705510701758166

Wolter, I., Braun, E., and Hannover, B. (2015). Reading is for girls!? the negative impact of preschool teachers' traditional gender role attitudes on boys' reading related motivation and skills. *Front. Psychol.* 6:1267. doi: 10.3389/fpsyg.2015.01267

Yeung, A. S., Lau, S., and Nie, Y. (2011). Primary and secondary students' motivation in learning english: grade and gender differences. *Contemp. Educ. Psychol.* 36, 246–256. doi: 10.1016/j.cedpsych.2011.03.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 26 June 2019
doi: 10.3389/fpsyg.2019.01242

Check for updates

Self-Concept and Support Experienced in School as Key Variables for the Motivation of Women Enrolled in STEM Subjects With a Low and Moderate Proportion of Females

Silke Luttenberger[1], Manuela Paechter[2] and Bernhard Ertl[3]*

[1]*Institute in Early Childhood and Primary Teacher Education, University College of Teacher Education Styria, Graz, Austria,* [2]*Educational Psychology, Institute of Psychology, University of Graz, Graz, Austria,* [3]*Learning and Teaching with Media, Department of Education, Universität der Bundeswehr München, Neubiberg, Germany*

OPEN ACCESS

Edited by:
Carl Senior,
Aston University, United Kingdom

Reviewed by:
Calah Ford,
University of Kentucky, United States
Nicole Farris,
Texas A&M University Commerce, United States

***Correspondence:**
Silke Luttenberger
silke.luttenberger@phst.at

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: 08 December 2018
Accepted: 10 May 2019
Published: 26 June 2019

Citation:
Luttenberger S, Paechter M and Ertl B (2019) Self-Concept and Support Experienced in School as Key Variables for the Motivation of Women Enrolled in STEM Subjects With a Low and Moderate Proportion of Females. Front. Psychol. 10:1242. doi: 10.3389/fpsyg.2019.01242

The proportion of women enrolled in STEM courses at university level has remained consistently low for decades. Differences, however, exist between various STEM domains: While engineering and technology appear especially unattractive, subjects such as mathematics, biology, or chemistry have better chances at attracting women. Research has mostly neglected these differences, treating STEM as an overall category. In the light of the differences in the proportions of women enrolled in and dropping out of various STEM subjects, the present study takes a more differentiated look to separately investigate the STEM subjects that have a low or moderate proportion of females. The following study focuses on female university students' intrinsic and extrinsic motivations in these two groups of STEM subjects, asking to what degree the academic STEM self-concept and support experienced in both school and the family contribute to the motivation to study a STEM topic. Four hundred sixty-nine female students took part in the investigation. Two hundred eighty-four of them were enrolled in STEM subjects with a low proportion of females (STEM-LPF) and 185 in STEM subjects with a moderate proportion of females (STEM-MPF). A comparison of the two samples shows that women in STEM-LPF exceed women in STEM-MPF with regard to their academic STEM self-concept and intrinsic and extrinsic motivations. Different variables contribute to motivation in the two samples. For STEM-LPF, a latent regression analysis found positive relationships between the academic STEM self-concept and both intrinsic and extrinsic motivations, while support experienced in school and from the family was not related to motivation. By contrast, in the STEM-MPF sample, the academic self-concept was not related to motivation. Previous interest in STEM subjects in school contributed positively to present intrinsic and extrinsic motivations. An unexpected result, however, was found concerning activities in school that were designed to promote interest in STEM. Memories of these kinds of activities were negatively related to both intrinsic and extrinsic motivations. These measures might be experienced as intrusive support: attempts to promote STEM sometimes might backfire and achieve the opposite of what was intended.

Keywords: gender, STEM, motivation, academic self-concept, school factors, university students, latent regression analysis

INTRODUCTION

Not many women pursue a career in the STEM fields (**S**cience, **T**echnology, **E**ngineering, and **M**athematics). This is particularly the case for engineering and technology. The proportion of females in these fields has remained consistently low over the last decades (e.g., Blickenstaff, 2005; Ihsen et al., 2013) – in the EU, it was only 25.7%, with Germany and its 18.5% scoring even lower (Center of Excellence Women and Science, 2014). There is also an underrepresentation of women in mathematics and sciences, although the difference is less severe, with a proportion of 37% in the EU and 35.6% in Germany (Center of Excellence Women and Science, 2014). Altogether, there are differences in the amounts of women in STEM between different EU member states, as well as between particular STEM fields. One reason for this might be that attitudes toward and attributions to subjects and professions vary between different cultural and subject-specific backgrounds (Else-Quest et al., 2010; Nosek and Smyth, 2011). Nosek and Smyth (2011) point out that women in the Western world still hold strong beliefs about STEM being a "male" domain.

And it is not just that women are less likely to enroll in STEM. When enrolled, they are more likely to drop out of their studies. Two variables in particular are seen as decisive for enrollment and dropout rates: STEM self-concept and motivation (Ellis et al., 2016). These will also serve as key variables in the present study, which investigate to what degree personal variables like the academic self-concept, as well as socializing school- and family-related factors contribute to women's intrinsic and extrinsic motivations in STEM.

In light of the variations in enrollment as well as dropout rates in different STEM subjects (e.g., higher enrollment of women in biology or chemistry than physics or engineering, Watt et al., 2017), it appears inadvisable to treat STEM as a general category. Watt and colleagues argue that it is "imperative to disaggregate discussions of different fields of sciences rather than use an aggregated concept of STEM" (Watt et al., 2017, p. 254). However, in light of an apparently infinite number of STEM majors, it is neither practical nor efficient to separately investigate subjects, which is why this study chooses another approach, classifying STEM majors according to the proportion of women enrolled in them.

The proportion of women in a field is a critical variable. It appears that personal attitudes, assessments, and characteristics are related to enrollment in subjects with a lower or higher proportion of women (Ihsen et al., 2013; Ertl et al., 2014). Gender disparities create a social context in STEM fields that signals to females that they are "minorities" who may not belong there (Murphy et al., 2007) and/or that women will encounter specific obstacles in these subjects. Therefore, the present study investigates the contribution of personal and social variables to motivation in two STEM groups with different proportions of women enrolled in the respective STEM fields.

MOTIVATION, SELF-CONCEPT, AND SOCIALIZING FAMILY- AND SCHOOL-RELATED FACTORS IN STEM

Considering impacts on career choice as well as persistence in a study or a course, one group of theoretical models and research studies focuses on motivation as a driving force. From a view point of motivation, intrinsic motivation, interest, enjoyment, the experience of self-determination as well as persistence, and the wish to achieve certain goals are important reasons to pursue a certain career path (e.g., Dickhäuser and Meyer, 2006; Ihsen et al., 2013; Van Soom and Donche, 2014). By contrast, models and research studies within a socio-cultural background focus more strongly on socializing factors and the cultural environment of a person (e.g., Dick and Rallis, 1991; Adya and Kaiser, 2005).

Motivation

Motivation plays a crucial role when it comes to learning behaviors, career choice, as well as persistence (e.g., Dickhäuser and Meyer, 2006; Ihsen et al., 2013). It describes the combination of a trait-like preference and a positively experienced, situation-specific state when working on a task (Macher et al., 2013). Motivation explains to which degree an individual makes an effort to achieve a particular goal, e.g., a good test score or a degree.

According to self-determination theory (Deci and Ryan, 1991), an individual can be motivated by internal rewards, e.g., enjoyment in doing, exploring, and learning things (intrinsic motivation) or by external rewards, e.g., money or prestige (extrinsic motivation). Intrinsic and extrinsic motivations can be ordered along a continuum. Intrinsic motivation can be described as an experience of competency or autonomy that manifests itself in sustainable efforts over a longer range of time. It is autonomous in the sense that it is experienced as self-determined (Van Soom and Donche, 2014). With regard to extrinsic motivation, self-determination theory distinguishes different forms. It comes from external sources (like external rewards) and is more goal driven and less sustainable. However, extrinsic motivation varies to which degree it is externally triggered; it may have internal sources when personal importance is placed on rewards (e.g., a high salary; Van Soom and Donche, 2014). Extrinsic motivation can serve as an impetus to put effort into learning, e.g., when passing a test is instrumental toward obtaining rewards.

The motivation to enroll in a STEM major and stay on a chosen STEM career path usually results from a combination of both high intrinsic and extrinsic motivations (Aeschlimann et al., 2016), with male STEM students' motivation mostly exceeding that of female students (Ihsen et al., 2013). With regard to motivation in STEM, mathematics has been identified as a critical filter, which excludes girls and women from joining and/or remaining in a STEM field of study (Else-Quest et al., 2010). Mathematics is regarded as a typically male domain, even though achievements have not differed across genders in many studies (Stevenson and Newman, 1986). Nevertheless, boys and men feel more confident overall when solving mathematical problems and have more positive attitudes toward and a higher motivation in mathematics (Vermeer et al., 2000; Else-Quest et al., 2010; Lazarides and Ittel, 2012).

Self-Concept

Self-concept is defined as the trait-like knowledge and perception about oneself. It is multidimensional, having both a nonacademic and an academic self-concept. These two overarching self-concepts are made up of more specific ones such as self-concepts in different academic domains (e.g., math, languages; Marsh and Scalas, 2011).

The academic self-concept is formed already in childhood by experiences, the feedback of others, as well as by attributions to a person's behavior and achievements, and it becomes more stable over time (Luttenberger et al., 2018). Self-concepts in different academic domains comprise an individual's self-assessments and perceptions of competence ("I am good in science"). These kinds of self-assessments may rely on different frames of reference. An external frame of reference uses comparisons of own achievements to the achievements of peers or to a predetermined performance standard, while an internal frame of reference compares former achievements to present ones (Altermatt et al., 2002; Schunk et al., 2014). Assessments can additionally rely on a general perception of individual abilities without the use of a certain frame of reference (Hoferichter et al., 2018).

Self-concept and achievements are related to each other *via* learning behaviors and effort (Valentine et al., 2004). A higher self-concept results in higher effort and persistence and thus implies higher achievements, while high achievements conversely strengthen the self-concept (Marsh and Scalas, 2011). The overall academic self-concept (how quickly I learn and how well I generally do at school) and the domain-specific self-concept (e.g., how good I am in mathematics) may influence preferences and choices of courses, school types, and even professions.

In STEM subjects, women often have a more negative self-concept than males, even if they in fact have the same grades and achievements (Jacobs et al., 2002; Watt, 2004; Frenzel et al., 2010; Nagy et al., 2010). Girls are more likely to attribute success to external factors and failure to internal factors such as a lack of mathematical ability (Parsons et al., 1984). Because mathematics is a crucial filter for enrollment as well as for remaining in STEM education, a lower math self-concept can be detrimental and may also lower learning motivation (Dresel et al., 2007). Generally, an overly critical math and STEM self-concept is a significant factor impacting why females are less motivated in STEM subjects and why they seldom consider a career in STEM (OECD, 2015). These differences can be partially seen as the results of socialization at home and school because the gender-specific math and STEM self-concepts become increasingly significant following primary school (Senler and Sungur, 2009).

School and Family as Socializing Factors

Social cognitive theory emphasizes the role of observing and interacting with others for an individual's personal development. In an ongoing process, individuals learn from significant others (as models), but individuals also vary in what they adopt from significant others and they are of course influenced by the consequences of their behavior and the interactions with others (Bussek and Bandura, 1999). As such significant others, teachers and parents or the closer social context are socializers (Watt et al., 2017), which influence the development of students' academic self-concept as well as their motivation in a field.

Teachers influence their students' self-concept and motivation in different ways. Cognitive activation, style of teaching, feedback, as well as teachers' attributions to achievement are important for the development of students' motivation in a subject (Lazarides and Ittel, 2012). Teachers can help students overcome gender-specific attribution patterns by, e.g., encouraging girls to attribute success to their own abilities and not external reasons, achieving motivation as a result (Dresel et al., 2007).

In a similar fashion, parents and the closer social context shape both the self-concept and motivation. Parents' beliefs about their children's abilities, as well as their feedback and support influence their children's self-concept and motivation (Tiedemann, 2000; Dresel et al., 2007; Gunderson et al., 2012; Watt et al., 2017). Parents also shape their children's career decisions by feedback, role modeling (Kessels, 2015), or content-specific support. In this context, students' interpretations of the social environment are crucial. According to Watt et al. (2017), career choice is more strongly influenced by how children and adolescents perceive parental positions than by what parents themselves report about their support. In a survey on careers in science, female researchers emphasized more than males the impact of their parents, particularly of the father, on their career choice (Sonnert, 2009). Besides parents, the support of peers and friends plays an important role for STEM career decisions (Robnett and Leaper, 2012). Dick and Rallis (1991) postulate in their model on career choice that aptitudes, experiences, socialization factors, and the cultural milieu have an impact on career-specific values, motivation, the self-concept, and career choice as a result.

RESEARCH QUESTIONS

As indicated, motivation is a crucial factor for enrollment and persistence in a STEM field. The present study investigates the contributing factors to the academic STEM self-concept and socializing factors (perceived family and school support) on intrinsic and extrinsic motivations in STEM. Taking into consideration the variations between different STEM fields, this study investigates two samples: Women studying STEM subjects with a low (STEM-LPF) and a moderate proportion of females (STEM-MPF). The following research questions will be investigated:

1. To what degree do STEM-LPF students and STEM-MPF students differ with regard to motivation, STEM self-concept, and socializing factors?
2. How do STEM self-concept and socializing factors contribute to motivation of (1) STEM-LPF students and (2) STEM-MPF students?

MATERIALS AND METHODS

Participants

The sample of this study is female German university students. They were surveyed in the context of a larger study in six European countries by the EU research project SESTEM, a project on equality of job opportunities. Students were contacted through university mailing lists and invited to participate in the study. The sample comprises primarily students in undergraduate and Master's level courses for STEM and STEM teaching.

Following the recommendation by Buchmann et al. (2002) to use a proportion of 30% as a critical threshold value for identifying typically male and more integrated professions, two samples of women were identified: 284 women in STEM-LPF (STEM fields with a low proportion of females, equal to or lower than 30%) and 185 women in STEM-MPF (STEM fields with a moderate proportion of females, higher than 30% but lower than 70%). All participants were enrolled in German universities. Women in the STEM-LPF sample studied the following subjects (ordered according to the number of participants enrolled in the respective fields): mechanical engineering, computer sciences, physics, metal engineering, electrical engineering, civil construction, or other kinds of engineering subjects (including subject combinations). STEM-MPF students studied mathematics, biology, geography, chemistry, STEM teacher education, biotechnology, architecture, or other subjects (ordered according to the number of participants enrolled in the respective fields). Some STEM-MPF students studied a STEM subject plus a non-STEM subject (e.g., languages, history), which are not listed here. In the present study, only STEM subjects within physical/natural sciences were investigated (e.g., medical subjects were excluded).

Measures

A questionnaire was developed by the six partners of the SESTEM consortium. They contributed according to their respective field of expertise and negotiated the specific constructs of the questionnaire, measurement approaches, and scoring systems. This kind of expert negotiation was chosen to ensure the validity of the questionnaire as well as to meet the different goals of the project. This negotiation paid attention to checks and balances to gather as much information as necessary and to keep the questionnaire as short as possible. During this process, an English version of the questionnaire was developed and then translated into five other languages including German. All questionnaire versions were combined to a multi-language questionnaire in LimeSurvey. Thus, students were able to choose their preferred language at the start of the online questionnaire. This online-survey included questions about:

1. Students' *majors* or study subjects. Students could name up to three study subjects that were part of their degree. These were classified according to their respective proportion of women based on the German first-year students' statistics [Destatis (Statistisches Bundesamt), 2013].
2. *Parents' professions*. The professions were entered as text and later classified as a STEM or a non-STEM field.
3. *Intrinsic* (five items) and *extrinsic motivations* (two items, see examples of items in **Table 1**). Higher values, measured on a five-point Likert scale, indicate a higher level of motivation.
4. A*cademic self-concept in STEM* was measured by four items that applied a five-point Likert scale (see **Table 1**). Higher values indicate a more positive self-concept.
5. *School factors*. These had two aspects: the first related to students' favorite subjects in STEM. For that, students were asked about their three most favorite subjects in school as free text. The answers were coded as STEM/non-STEM, and the STEM subjects were summed up to a score. For the analyses in this manuscript, this score only includes STEM subjects that are considered as a "male" domain (excluding, e.g., biology). Higher values indicate a higher number of favorite STEM subjects. The second aspect of the school factors is related to school support for STEM. Students were asked three questions concerning specific school teachers or activities that encouraged students' interest in STEM (e.g., "Were there specific school activities based around STEM, such as school visits or special projects?"). Positive answers were summed up to a score and by multiplication adjusted to a range between 0 (no activities) and 5 (all activities).

TABLE 1 | Overview of the scales used for the study with the number of items, an example, and the internal consistency.

Scale	Items	Example	Cronbach's α
Academic self-concept STEM	4	"I am not skilled enough in mathematics for choosing a career in STEM"	0.82
Intrinsic motivation	5	"I want to work in STEM to contribute to scientific and technical developments"	0.71
Extrinsic motivation	2	"The high salaries make a career in STEM attractive to me"	0.73

Note: The Cronbach's α calculated for the whole German sample (567 students). This larger sample included students with majors, which did not qualify for inclusion in the present study. According to Paechter et al. (2013), Cronbach's α with 0.70 and more can be considered a satisfying indicator of the internal consistency of a scale. If necessary, values for academic self-concepts were recoded, so that higher values indicate a higher self-concept.

6. *Family factors.* They describe support students received by family and peers (e.g., parents, siblings, friends). Students were asked whether they received support for homework or for career decisions in different subjects (e.g., math, science) and from whom they received support (e.g., "Who helped you with your science homework? [Brother or sister]"). Positive answers were summed up to a score for the subject and for the person giving support and by multiplication adjusted to a range between 0 (no support) and 1 (full support). This manuscript applies three variables for the analyses: support for math, support for science, and support by parents.

Table 1 shows the characteristics of the Likert scales including the number of items, an exemplary item, and the internal consistency of the scale. Missing items of singles scales were imputed; missing scales were treated as missing. We used the values of the skewness and kurtosis to analyze the distribution of the data. West et al. (1995) set the criteria for indicators used in structural equation models at a value of >2 for skewness and >7 for kurtosis for deviation from normal distribution. All scales meet the criteria of normal distribution.

RESULTS

Descriptive Statistics for the STEM-LPF Sample

Of the 284 students in STEM-LPF, 50.4% of the students (134) had a father and 11.3% (30) had a mother working in a STEM profession. Most students showed a very positive STEM self-concept (M = 4.58; the means described in the following related to a scale of 1–5, with 1 as the lowest value and 5 as the highest value). Intrinsic (M = 3.99) and extrinsic motivations (M = 3.81) were positive. With respect to school factors, 50 students had three favorite STEM subjects at school, 141 students had two, 84 had just one, while 9 had favorite non-STEM subjects (M = 1.82). They received a moderate amount of STEM support in school (M = 2.33 of a maximum of 5). Considering family factors, the amount of parents' support in math (M = 0.14 of a maximum of 1) and STEM (M = 0.14) was low. General support by the parents was low to medium (M = 0.36). **Table 2** provides an overview of all scales, including their value range, their means, and their standard deviations.

Table 3 provides an overview of the bivariate correlations between the variables.

Descriptive Statistics for the STEM-MPF Sample

Of the 185 students in STEM-MPF, 53.1% of the students (93) had a father and 8.6% (15) had a mother in a STEM profession. Most students showed a very positive STEM self-concept (M = 4.20; the means described in the following related to a scale of 1–5, with 1 as the lowest value and 5 as the highest value). Intrinsic (M = 3.79) and extrinsic motivations (M = 3.40) were positive. With respect to school factors, 22 students had three favorite STEM subjects at school, 88 students had two, 66 had just one, while 9 had favorite non-STEM subjects (M = 1.66). They received a moderate amount of STEM support in school (M = 2.38 of a maximum of 5). Considering family factors, the amount of parents' support in math (M = 0.20 of a maximum of 1) and STEM (M = 0.15) was low. General support by the parents was low to medium (M = 0.38). **Table 4** provides an overview of all scales, including their value range, their means, and their standard deviations.

Table 5 provides an overview of the bivariate correlations between the variables.

TABLE 2 | Ranges, means, standard deviations, and *n* for the reported scales, students in the STEM-LPF sample.

	Range	*M*	*SD*	*n*
Motivation				
Intrinsic	1 … 5	3.99	0.54	284
Extrinsic	1 … 5	3.81	0.74	284
Academic self-concept STEM	1 … 5	4.58	0.55	284
School factors				
STEM favorites	0 … 3	1.82	0.75	277
School support	0 … 5	2.33	2.07	277
Family factors				
Mathematics support	0 … 1	0.14	0.19	284
STEM support	0 … 1	0.14	0.19	284
Parent general support	0 … 1	0.36	0.19	284

TABLE 3 | Bivariate correlations between variables in the STEM-LPF sample.

	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intrinsic motivation (1)	0.327**	0.325**	0.103	−0.057	0.079	−0.029	0.077
Extrinsic motivation (2)	1.000	0.159**	0.110	−0.015	0.008	0.031	0.067
Academic self-concept STEM (3)		1.000	0.094	−0.091	−0.087	−0.061	−0.003
STEM favorites (4)			1.000	−0.210**	−0.034	−0.020	0.007
School support (5)				1.000	0.028	0.068	0.004
Mathematics support (6)					1.000	0.604**	0.618**
STEM support (7)						1.000	0.621**
Parent general support (8)							1.000

*Note: Significant correlations are marked with asterisks (**$p < 0.01$).*

Research Question 1: Differences Between the STEM-LPF and the STEM-MPF Students

A multivariate analysis of variance (MANOVA) with the two groups (STEM-LPF and STEM-MPF) as independent variables and intrinsic and extrinsic motivations as dependent variables was carried out. MANOVA showed an overall significant result, $F(2,463) = 15.96$, $p < 0.01$, $\eta^2 = 0.06$. Women in STEM-LPF scored significantly higher than women in STEM-MPF on intrinsic motivation, $F(1,464) = 29.09$, $p < 0.01$, $\eta^2 = 0.06$, as well as on extrinsic motivation, $F(1,464) = 12.71$, $p < 0.01$, $\eta^2 = 0.03$.

A *t*-test between both groups with the academic STEM self-concept as dependent variable also found higher values for women in STEM-LPF, $t(109) = 6.25$, $p < 0.01$, Cohen's $d = 0.7$.

MANOVA with the two groups (STEM-LPF and STEM-MPF) as independent variables and the school factors as dependent variables was not significant, $F(2,448) = 2.04$, $p > 0.05$, $\eta^2 = 0.01$.

MANOVA with the two groups (STEM-LPF and STEM-MPF) as independent variables and family variables as dependent variables was significant, $F(3,465) = 4.68$, $p < 0.05$, $\eta^2 = 0.03$. Women in STEM-MPF scored significantly higher than women in STEM-LPF on perceived parental support in math, $F(1,467) = 11.78$, $p < 0.01$, $\eta^2 = 0.02$. There were significant differences neither for perceived parental support in STEM, $F(1,467) = 0.34$, $p > 0.05$, $\eta^2 = 0.00$, nor for general parental support, $F(1,467) = 1.43$, $p > 0.05$, $\eta^2 = 0.00$.

TABLE 4 | Ranges, means, standard deviations, and *n* for the reported scales, students in the STEM-MPF sample.

	Range	*M*	*SD*	*n*
Motivation				
Intrinsic	1 … 5	3.79	0.67	185
Extrinsic	1 … 5	3.40	0.86	185
Academic self-concept STEM	1 … 5	4.20	0.70	185
School factors				
STEM favorites	0 … 3	1.66	0.75	174
School support	0 … 5	2.38	2.12	174
Family factors				
Mathematics support	0 … 1	0.20	0.21	185
STEM support	0 … 1	0.15	0.19	185
Parent general support	0 … 1	0.38	0.21	185

Research Question 2a: Latent Regression Analysis for STEM-LPF Students

Latent regression analysis was used to test the relationships between the variables in a multivariate, multiple regression context. Structural relationships between multiple dependent variables and multiple independent variables can be analyzed simultaneously. Regression analyses are specified at the latent level and are corrected for measurement error at the level of the independent and dependent variables. Latent regression analyses have the advantage that the relationships between variables in the regression model can be estimated more accurately (Geiser, 2013). The data were analyzed with Mplus 6 using a maximum likelihood estimator. The goodness of fit of the data to the hypothesized model was assessed using the following indices: χ^2/df, comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). In general, values of $\chi^2/\text{df} < 2$, $\text{CFI} > 0.95$, $\text{RMSEA} < 0.05$, and $\text{SRMR} < 0.05$ are considered indicators of good model fit. The model fit indices suggest a good fit of the latent regression analysis model for STEM-LPF: $\chi^2/\text{df} = 1.039$; CFI = 0.999; RMSEA = 0.012; SRMR = 0.022.

Table 6 displays the standardized solutions for the latent regression analysis with the school and family factors. Each factor comprises different variables. The model shows that the two indicators STEM favorites at school ($\beta = 0.593$) and school support ($\beta = -0.355$) are related to the latent school factor. The three indicators support in mathematics ($\beta = 0.778$), support in STEM ($\beta = 0.777$), and support by parents ($\beta = 0.797$) are related to the latent factor family.

The regression coefficients between the school and family factors and the self-concept as predictor variables and intrinsic and extrinsic motivations as criterion variables show the following results: students in STEM-LPF with higher levels of academic self-concept in STEM report higher intrinsic ($\beta = 0.308$) and extrinsic motivations ($\beta = 0.139$). There were no significant correlations between the school and family factors and intrinsic and extrinsic motivations. The total variance of intrinsic and extrinsic motivations that can be explained is

TABLE 5 | Bivariate correlations between variables in the STEM-MPF sample.

	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Intrinsic motivation (1)	0.410**	0.380**	0.259**	−0.201**	0.022	−0.011	0.077
Extrinsic motivation (2)	1.000	0.299**	0.230**	−0.284**	−0.117	−0.140	−0.050
Academic self-concept STEM (3)		1.000	0.395**	−0.195*	−0.150*	−0.092	0.023
STEM favorites (4)			1.000	−0.251**	−0.146*	−0.162*	−0.089
School support (5)				1.000	0.039	0.173*	0.080
Mathematics support (6)					1.000	0.486**	0.645**
STEM support (7)						1.000	0.648**
Parent general support (8)							1.000

*Note: Significant correlations are marked with asterisks (*p < 0.05, **p < 0.01).*

R^2 (intrinsic) = 0.125 and R^2 (extrinsic) = 0.046. **Figure 1** gives an overview of the indicators and factors of the latent regression analysis model.

Research Question 2b: Latent Regression Analysis for Students in Subjects With a Moderate Proportion of Females

The model fit indices suggest a good fit of the latent regression analysis model for STEM-MPF: $\chi^2/df = 1.759$; CFI = 0.970; RMSEA = 0.064; SRMR = 0.040. **Table 7** displays the standardized solutions for the latent regression analysis with the school and family factors. The model shows that the two indicators STEM favorites at school ($\beta = 0.591$) and school support ($\beta = -0.431$) are related to the latent school factor. The three indicators support in mathematics ($\beta = 0.705$), support in STEM ($\beta = 0.710$), and support by parents ($\beta = 0.910$) are high, positively related to the latent family factor.

The regression coefficients between the school and family factors and the STEM self-concept as predictor variables and intrinsic and extrinsic motivations as criterion variables show the following results: the model shows a moderate relationship between the latent school factor and students' intrinsic ($\beta = 0.403$) and extrinsic ($\beta = 0.461$) motivations. Students in STEM-MPF who reported a higher number of favorite STEM subjects in school have higher intrinsic and extrinsic motivations in STEM. By contrast, higher levels of school support indicate lower intrinsic and extrinsic motivations in STEM. There were no significant correlations between academic self-concept in STEM and support by the family. The total variance of intrinsic and extrinsic motivations that can be explained by the factors is R^2 (intrinsic) = 0.244 and R^2 (extrinsic) = 0.299. **Figure 2** gives an overview of the indicators and factors of the latent regression analysis model.

DISCUSSION

Consistent with the suggestions to differentiate between STEM domains (Watt et al., 2017), this study classified STEM subjects into two groups according to their proportion of female students. Distinct differences and patterns could be found for the two groups.

STEM Self-Concept as a Key Predictor for Motivation in STEM-LPF Subjects

Women in STEM subjects with a low proportion of females showed a significantly higher STEM self-concept and significantly higher intrinsic and extrinsic motivations than women in STEM-MPF. The higher STEM self-concept might be a result of a selection process in which only the most confident and able females decide on a career in strongly male-dominated fields.

Only the STEM self-concept contributed significantly to intrinsic and extrinsic motivations in the latent regression analysis. Neither family factors nor school factors received significant β-weights. Generally, research points at a strong influence of the STEM self-concept on the motivation to achieve

TABLE 6 | Standardized coefficients for the latent regression analysis for the STEM-LPF sample.

Observed variable	Latent factor	*β*	*SE*	*p*
STEM favorites	*School*	0.593	0.250	0.018
Support in school		–0.355	0.157	0.024
Mathematics support	*Family*	0.778	0.034	<0.001
Science support		0.777	0.034	<0.001
Parent support		0.797	0.034	<0.001

TABLE 7 | Standardized coefficients for the latent regression analysis for the students in subjects with a moderate proportion of females.

Observed variable	Latent factor	*β*	*SE*	*p*
STEM favorites	*School*	0.591	0.100	<0.001
Support in school		–0.431	0.090	<0.001
Mathematics support	*Family*	0.705	0.047	<0.001
Science support		0.710	0.048	<0.001
Parent support		0.910	0.041	<0.001

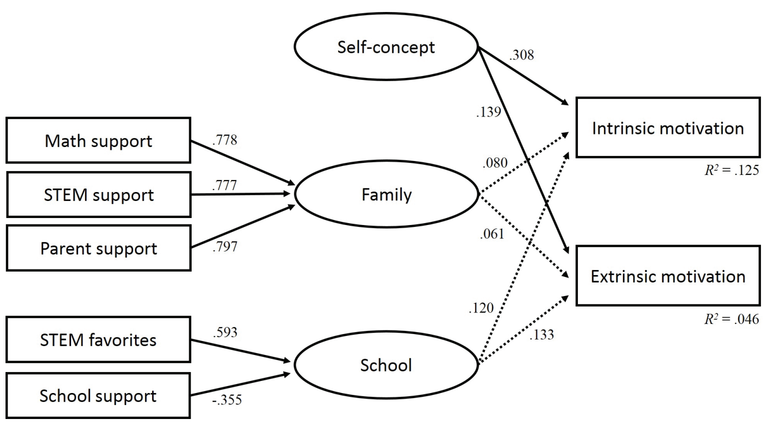


FIGURE 1 | Latent regression analysis for the STEM-LPF sample.

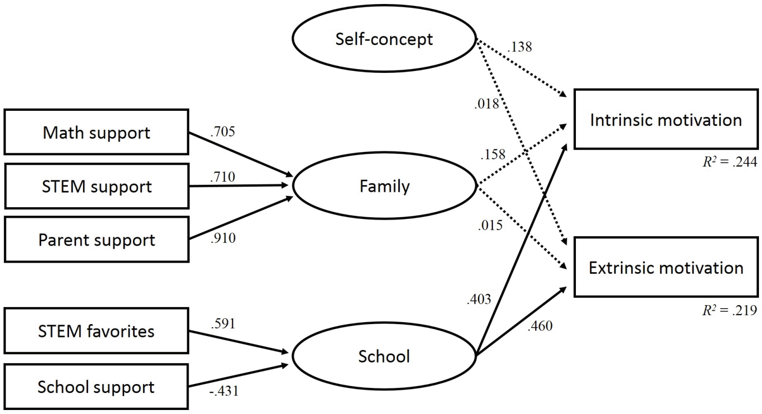


FIGURE 2 | Latent regression analysis for the students in subjects with a moderate proportion of females.

in a subject and invest effort into it (Guay et al., 2010). Both self-determination theory and research on the academic self-concept (Marsh and Craven, 2005; Guay et al., 2010) assume that perceiving oneself as able and competent increases intrinsic motivation and achievement. In the present study, this kind of high self-concept in STEM was related to facets of intrinsic motivation, for example, the wish to contribute positively to scientific and technical developments in the respective field of study (see item example in **Table 1**).

STEM self-concept was also significantly related to extrinsic motivation. However, in comparison to intrinsic motivation, the β-weight was much smaller. This result might be explained by the women's previous experiences. A high self-concept and intrinsic motivation are optimal requirements for long-lasting, persistent involvement in a domain. By contrast, involvement based on external motivation is much more liable to break down in case of failure, disappointment, stereotypes, etc. (Guay et al., 2010) – all obstacles which women in extremely male-dominated subjects are likely to encounter. As a result, an especially strong link between intrinsic motivation and self-concept can be expected.

School Factors as Key Predictors for Motivation in STEM-MPF Subjects

Only school factors contributed to intrinsic and extrinsic motivations in the latent regression analysis. Neither family factors nor the academic self-concept received significant β-weights.

Former experiences and preferences expressed as STEM subjects being favorite subjects in school contributed positively to intrinsic and extrinsic motivations. Favorite subjects can be seen as indicators of intrinsic motivation. So in this case, former appraisal of STEM (to be sure, assessed in hindsight) contributes to present intrinsic motivation. STEM subjects as favorite subjects also contribute to extrinsic motivation, e.g., expectations of a good salary, etc.

By contrast (and at a first glance, somewhat counterintuitively), the variable of school support contributed negatively to intrinsic and extrinsic motivations. It expresses assessments of the support provided by teachers and the school as such, e.g., special activities that encouraged interest in STEM. There are different explanations for this result. It might be that, in hindsight, these activities in school have little to do with present academic experiences and demands. This variable would therefore tend to impede current motivation in STEM. Another reason might be that these kinds of school activities were perceived as intrusive, perhaps carrying the message that STEM subjects are not interesting *per se*, but need special encouragement and emphasis in order to be regarded as attractive, especially for girls. It speaks for this assumption that the variables STEM favorites and school support correlate negatively. Bhanot and Jovanovic (2005) showed for mathematics that parental support often carries these kinds of "hidden" messages. When this is the case, special activities and encouragement provided by teachers might backfire and discourage rather than encourage motivation and interest.

Differences Between Women in STEM-LPF and STEM-MPF Subjects

Differences between the two groups of women concern mainly the academic STEM self-concept and intrinsic and extrinsic motivations. Women in STEM subjects with a low proportion of females excelled by a higher STEM self-concept. One may assume that only women with a strong STEM self-concept that "inoculates" against the multiple barriers (stereotypes, lack of support systems in the family or by peers; Dasgupta and Stout, 2014; Wang et al., 2015) will lastly decide for and stick to a career in a STEM-LPF field. As self-concept is formed in early years of childhood and becomes more stable over time (Nagy et al., 2010; Ertl et al., 2017), this result points at the importance to build up a strong STEM self-concept already in early school years.

However, the decision to take up a study in a STEM-LPF field needs not only a strong self-concept but also a high degree of persistence and stamina. Therefore, it is not surprising that the difference between women in STEM-LPF and STEM-MPF concerning their STEM self-concept is accompanied by a higher degree of intrinsic and extrinsic motivations. Students with higher motivation in a subject invest more time and effort in learning and performance and thus meet important requirements for academic success. In such a long-term context, intrinsic motivation is especially important (Van Soom and Donche, 2014; Luttenberger et al., 2018).

Of the family and school factors, only one variable distinguished between the two groups. Women in STEM-LPF perceived lower math support from the family than women in STEM-MPF. In the light of differences in self-concept and motivation, it seems probable that these women did not perceive their parents' support as so important for pursuing their STEM career as they already had strong internal personal motives for their decision.

LIMITATIONS AND STRENGTHS OF THE STUDY

A strength of this study is its relatively large sample with 469 individuals of a "rare species" – women in STEM subjects, with a large number of them even studying subjects with a low proportion of females.

This study's limitations may perhaps be found in its methodology that uses a cross-sectional design instead of a longitudinal design. This has implications for drawing inferences and causalities. Thus, it is not possible to investigate whether relationships, for example, between school and family factors and motivational variables, are mediated by the STEM self-concept. To investigate such causalities, a longitudinal design would be desirable. However, achieving a sample size like the one in the present study with a longitudinal design that covers ideally primary to secondary and to tertiary education is nearly impossible. Nevertheless, the study gives new insights into the motivation of women in university studies that are connoted as typically male domains. An important point of the survey is that it does not consider these studies as a homogeneous domain but takes a differentiated view.

ETHICS STATEMENT

This study was performed in accordance with the 1964 Declaration of Helsinki and the American Psychological Association's Ethics Code. In accordance with the national and institutional requirements, review and approval were not required for this study. Participants gave consent to participate in the study by submitting the online questionnaire.

AUTHOR CONTRIBUTIONS

SL, MP, and BE have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

Parts of this manuscript were funded by the EU (LLP-Program, Project SESTEM 505437-llp-2009-GR-KA1-KA1SCR). The publication was funded by University College of Teacher Education Styria, Graz, Austria.

ACKNOWLEDGMENTS

Special thanks and acknowledgement to Ms. Sog Yee Mok for her support in implementing this study.

REFERENCES

Adya, M., and Kaiser, K. M. (2005). Early determinants of women in the IT workforce: a model of girls' career choices. *Inf. Technol. People* 18, 230–259. doi: 10.1108/09593840510615860

Aeschlimann, B., Herzog, W., and Makarova, E. (2016). How to foster students' motivation in mathematics and science classes and promote students' STEM career choice. A study in Swiss high schools. *Int. J. Educ. Res.* 79, 31–41. doi: 10.1016/j.ijer.2016.06.004

Altermatt, E. R., Pomerantz, E. M., Ruble, D. N., Frey, K. S., and Greulich, F. K. (2002). Predicting changes in children's self-perceptions of academic competence: a naturalistic examination of evaluative discourse among classmates. *Dev. Psychol.* 38, 903–917. doi: 10.1037/0012-1649.38.6.903

Bhanot, R., and Jovanovic, J. (2005). Do parents' academic gender stereotypes influence whether they intrude on their children's homework? *Sex Roles* 52, 597–607. doi: 10.1007/s11199-005-3728-4

Blickenstaff, J. (2005). Women and science careers: leaky pipeline or gender filter? *Gend. Educ.* 17, 369–386. doi: 10.1080/09540250500145072

Buchmann, M., Kriesi, I., Pfeifer, A., and Sacchi, S. (2002). *Halb drinnen – halb draussen. Analysen zur Arbeitsmarktintegration von Frauen in der Schweiz. [Half inside – half outside. Analyses on the integration of women into the labor market in Switzerland.].* (Zürich: Rüegger).

Bussek, K., and Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychol. Rev.* 106, 676–713. doi: 10.1037/0033-295X.106.4.676

Center of Excellence Women and Science. (2014). *Studentinnenanteil in Mathematik/Naturwissenschaften und Ingenieurwissenschaften (ISCED 5-6) im internationalen Vergleich (2011).* [Proportion of female students in mathematics/sciences and engineering (ISCED 5-6) in an international comparison]. Retrieved from: http://www.gesis.org/cews/fileadmin/cews/www/statistiken/08_d.gif (Accessed April 02, 2014).

Dasgupta, N., and Stout, J. G. (2014). Girls and women in science, technology, engineering and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights Behav. Brain Sci.* 1, 21–29. doi: 10.1177/2372732214549471

Deci, E. L., and Ryan, R. M. (1991). "A motivational approach to self: Integration in personality" in *Nebraska symposium on motivation: Perspectives on motivation.* ed. R. Dienstbier, vol. 38 (Lincoln, NE: University of Nebraska Press), 237–288.

Destatis (Statistisches Bundesamt) (2013). *Bildung und kultur. Studierende an hochschulen. Wintersemester 2012/2013.* [Education and culture. Students at universities. Winter term 2012/2013]. (Wiesbaden: Statistisches Bundesamt).

Dick, T. P., and Rallis, S. F. (1991). Factors and influences on high school students' career choices. *J. Res. Math. Educ.* 22, 281–292. doi: 10.2307/749273

Dickhäuser, O., and Meyer, W.-U. (2006). Gender differences in young children's math ability attributions. *Psychol. Sci.* 48, 3–16.

Dresel, M., Schober, B., and Ziegler, A. (2007). "Golem und Pygmalion. Scheitert die Chancengleichheit von Mädchen im mathematisch-naturwissenschaftlich-technischen Bereich am geschlechtsstereotypen Denken der Eltern? [Golem and Pygmalion. Do equal opportunities for girls in mathematical-scientific-technical domains fail because of their parents' gender-stereotype thinking?]" in *Erwartungen in himmelblau und rosarot. Effekte, Determinanten und Konsequenzen von Geschlechterdifferenzen in der Schule*. eds. P. H. Ludwig, and H. Ludwig (Weinheim: Juventa), 61–81.

Ellis, J., Fosdick, B. K., and Rasmussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: lack of mathematical confidence a potential culprit. *PLoS One* 11:e0157447. doi: 10.1371/journal.pone.0157447

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Ertl, B., Luttenberger, S., and Paechter, M. (2014). Stereotype als Einflussfaktoren auf die Motivation und die Einschätzung der eigenen Fähigkeiten bei Studentinnen in MINT-Fächern. [Stereotypes as influencing factors on motivation and assessment of one's own skills of female students in STEM-subjects]. *Gruppendynamik und Organisationsberatung* 45, 419–440. doi: 10.1007/s11612-014-0261-3

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Frenzel, A. C., Goetz, T., Pekrun, R., and Watt, H. M. G. (2010). Development of mathematics interest in adolescence: influences of gender, family, and school context. *J. Res. Adolesc.* 20, 507–537. doi: 10.1111/j.1532-7795.2010.00645.x

Geiser, C. (2013). *Data analysis with mplus*. (New York, NY: The Guilford Press).

Guay, F., Chanal, J., Ratelle, C. F., Marsh, H. W., Larose, S., and Boivin, M. (2010). Intrinsic, identified, and controlled types of motivation for school subjects in young elementary school children. *Br. J. Educ. Psychol.* 80, 711–735. doi: 10.1348/000709910X499084

Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles* 66, 153–166. doi: 10.1007/s11199-011-9996-2

Hoferichter, F., Lätsch, A., Lazarides, R., and Raufelder, D. (2018). The big-fish-little-pond effect on the four facets of academic self-concept. *Front. Psychol.* 9:1247. doi: 10.3389/fpsyg.2018.01247

Ihsen, S., Höhle, E. A., and Baldin, D. (2013). "Spurensuche!: Entscheidungskriterien für Natur-bzw. Ingenieurwissenschaften und mögliche Ursachen für frühe Studienabbrüche von Frauen und Männern an TU9-Universitäten. [Tracking!: decision criteria for science and engineering and possible causes for early dropouts of women and men at TU9 universities.]" in *TUM gender- und diversity-studies*, vol. 1 (Berlin: LIT).

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., and Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Dev.* 73, 509–527. doi: 10.1111/1467-8624.00421

Kessels, U. (2015). Bridging the gap by enhancing the fit: how stereotypes about STEM clash with stereotypes about girls. *Int. J. Gender, Sci. Technol.* 7, 280–296.

Lazarides, R., and Ittel, A. (2012). Instructional quality and attitudes toward mathematics: do self-concept and interest differ across students' patterns of perceived instructional quality in mathematics classrooms? *Child Dev. Res.* 2012. doi: 10.1155/2012/813920

Luttenberger, S., Wimmer, S., and Paechter, M. (2018). Spotlight on math anxiety. *Psychol. Res. Behav. Manag.* 11, 311–322. doi: 10.2147/PRBM.S141421

Macher, D., Paechter, M., Papousek, I., Ruggeri, K., Freudenthaler, H. H., and Arendasy, M. (2013). Statistics anxiety, state anxiety during an examination, and academic achievement. *Br. J. Educ. Psychol.* 83, 535–549. doi: 10.1111/j.2044-8279.2012.02081.x

Marsh, H. W., and Craven, R. G. (2005). "A reciprocal effects model of the causal ordering of self-concept and achievement: new support for the benefits of enhancing self-concept" in *New frontiers for self research*. eds. H. W. Marsh, R. G. Craven, and D. McInerney, vol. 2 (Greenwich, CT: Information Age), 15–52.

Marsh, H. W., and Scalas, L. F. (2011). "Self-concept in learning: reciprocal effects model between academic self-concept and academic achievement" in *Social and emotional aspects of learning*. ed. S. Järvela (Amsterdam: Elsevier), 191–197.

Murphy, M. C., Steele, C. M., and Gross, J. J. (2007). Signaling threat: how situational cues affect women in math, science, and engineering settings. *Psychol. Sci.* 18, 879–885. doi: 10.1111/j.1467-9280.2007.01995.x

Nagy, G., Watt, H. M. G., Eccles, J. S., Trautwein, U., Lüdtke, O., and Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: different countries, different trajectories? *J. Res. Adolesc.* 20, 482–506. doi: 10.1111/j.1532-7795.2010.00644.x

Nosek, B. A., and Smyth, F. L. (2011). Implicit social cognitions predict sex differences in math engagement and achievement. *Am. Educ. Res. J.* 48, 1125–1156. doi: 10.3102/0002831211410683

OECD (2015). *The ABC of gender equality in education: Aptitude, behaviour, confidence*. (Paris: PISA, OECD Publishing).

Paechter, M., Rebmann, K., Schlömer, T., Mokwinski, B., Hanekamp, Y., and Arendasy, M. (2013). Development of the Oldenburg Epistemic Beliefs Questionnaire (OLEQ), a German questionnaire based on the Epistemic Belief Inventory (EBI). *Curr. Issues Educ.* 16, 1–18. Retrieved from http://cie.asu.edu/ojs/index.php/cieatasu/article/view/1035

Parsons, J. E., Adler, T., and Meece, J. L. (1984). Sex differences in achievement: a test of alternate theories. *J. Pers. Soc. Psychol.* 46, 26–43. doi: 10.1037/0022-3514.46.1.26

Robnett, R. D., and Leaper, C. (2012). Friendship groups, personal motivation, and gender in relation to high school students' STEM career interest. *J. Res. Adolesc.* 23, 652–664. doi: 10.1111/jora.12013

Schunk, D. H., Meece, J. L., and Pintrich, P. R. (2014). *Motivation in education: Theory, research, and applications.* 4th edn. (Boston: Pearson Education).

Senler, B., and Sungur, S. (2009). Parental influences on students' self-concept, task value beliefs, and achievement in science. *Span. J. Psychol.* 12, 106–117. doi: 10.1017/S1138741600001529

Sonnert, G. (2009). Parents who influence their children to become scientists: effects of gender and parental education. *Soc. Stud. Sci.* 39, 927–941. doi: 10.1177/0306312709335843

Stevenson, H. W., and Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading. *Child Dev.* 57, 646–659. doi: 10.2307/1130343

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *J. Educ. Psychol.* 92, 144–151. doi: 10.1037/0022-0663.92.1.144

Valentine, J. C., DuBois, D. L., and Cooper, H. (2004). The relation between self-beliefs and academic achievement: a meta-analytic review. *Educ. Psychol.* 39, 111–133. doi: 10.1207/s15326985ep3902_3

Van Soom, C., and Donche, V. (2014). Profiling first-year students in STEM programs based on autonomous motivation and academic self-concept and relationship with academic achievement. *PLoS One* 9:e112489. doi: 10.1371/journal.pone.0112489

Vermeer, H. J., Boekaerts, M., and Seegers, G. (2000). Motivational and gender differences: sixth-grade students' mathematical problem-solving behavior. *J. Educ. Psychol.* 92, 308–315. doi: 10.1037/0022-0663.92.2.308

Wang, M.-T., Degol, J., and Ye, F. (2015). Math achievement is important, but task values are critical, too: examining the intellectual and motivational factors leading to gender disparities in STEM careers. *Front. Psychol.* 6. doi: 10.3389/fpsyg.2015.00036

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th through 11th grade Australian students. *Child Dev.* 75, 1556–1574. doi: 10.1111/j.1467-8624.2004.00757.x

Watt, H. M. G., Hyde, J. S., Petersen, J., Morris, Z. A., Rozek, C. S., and Harackiewicz, J. M. (2017). Mathematics: a critical filter for STEM-related career choices? A longitudinal examination among Australian and US adolescents. *Sex Roles* 77, 254–271. doi: 10.1007/s11199-016-0711-1

West, S. G., Finch, J. F., and Curran, P. J. (1995). "Structural equation models with nonnormal variables: problems and remedies" in *Structural equation modeling: Concepts, issues, and applications*. ed. R. H. Hoyle (Thousand Oaks, CA: Sage), 56–75.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 27 June 2019
doi: 10.3389/fpsyg.2019.01432

The Impacts of Gender and Subject on Experience of Competence and Autonomy in STEM

Sabrina Sobieraj* **and Nicole C. Krämer**

Social Psychology: Media and Communication, University of Duisburg-Essen, Duisburg, Germany

In most societies, women are less likely to choose a science, technology, engineering and mathematics (STEM)-related study program than men. This problem persists despite numerous initiatives aimed at fostering the uptake of STEM subjects by women, who represent an underutilized source of talent in a time of great need for STEM professionals. Many reasons for women's avoidance of the path into STEM-related areas have been discussed, including weaker mathematical skills, implicit gender stereotypes or structural deficits in school education. One variable which is presumably at the core of decisions regarding a specific study subject is motivation. We aim to look in greater depth at the basis for motivation by referring to self-determination theory (SDT). Here, we specifically focus on the needs for *competence* and *autonomy* which represent pivotal sources of motivation, effective performance and psychological well-being and are assumed to be positively correlated with academic achievement and perseverance. In line with previous SDT research, we assume that self-perceptions during STEM studies contribute to experiences of competence and autonomy and may be responsible for gender disparities. To examine whether and how a sex-specific perception of autonomy and competence influences decisions regarding STEM subjects, we conducted a survey study of Master's students (N = 888; 461 female, 427 male), who were enrolled either in STEM or non-STEM subjects, and asked about students' motivations, perceived competence (e.g., self-efficacy) and autonomy (e.g., volitional decision for a study major). The results revealed several main effects of study major and only a small number of interaction effects of sex and subject. For example, non-STEM students were more likely to enroll due to their stronger interest in their subject, signifying higher autonomy, while STEM students were more likely to select their subject according to their families' wishes. The comparison between female and male STEM students revealed that males perceived more self-efficacy and reported more leadership aspirations while female STEM students have lower perceptions of their own competence, especially regarding perceived future competences.

Keywords: gender, STEM – science technology engineering mathematics, competence, autonomy, motivation

OPEN ACCESS

Edited by:
Manuela Paechter,
University of Graz, Austria

Reviewed by:
Iris Marušić,
Institute for Social Research in Zagreb, Croatia
Kathy Ellen Green,
University of Denver, United States

Correspondence:
Sabrina Sobieraj
Sabrina.Sobieraj@uni-due.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *08 January 2019*
Accepted: *04 June 2019*
Published: *27 June 2019*

Citation:
Sobieraj S and Krämer NC (2019) The Impacts of Gender and Subject on Experience of Competence and Autonomy in STEM. Front. Psychol. 10:1432. doi: 10.3389/fpsyg.2019.01432

INTRODUCTION

Despite economists' repeated calls for more professionals in science, technology, engineering and mathematics (STEM) in both the short and medium term, the number of students deciding to enroll in STEM still does not meet the economic needs (Dasgupta and Stout, 2014). Moreover, Dasgupta and Stout (2014) report that women are still underrepresented in STEM, resulting in an underutilized source of talent. Indeed, even if women do start an academic career in science or engineering, they do not go on to achieve the highest positions: In the academic sector in the United States, only 21% of full professors in science and only 5% of full professors in engineering are female (Shen, 2013). This is a long standing problem, with Wickware (1997) having acknowledged over 20 years ago that women leave research more often than men.

Many reasons have been discussed regarding why women avoid the path into STEM, such as weaker mathematical skills, structural deficits in school education, and gender stereotypes (see Mavriplis et al., 2010; Smeding, 2012; Wang and Degol, 2013). Gender stereotypes are probably one of the most intensively debated of the proposed reasons. For example, it has been found that the picture of a scientist is more congruent with the stereotype of a man than of a woman (e.g., Carli et al., 2016). This stereotype has often been found among adults, and even more impressively, it has also been demonstrated in children: Using the draw-a-scientist method, in which children are instructed to draw a scientist, children consistently draw Caucasian males (e.g., Finson, 2002; Buldu, 2006; Losh et al., 2008; Miller et al., 2018). Such findings reflect how at least scientists or engineers (e.g., Fralick et al., 2009) are perceived. Based on this, it is conceivable that the implicit knowledge about STEM professionals and gender stereotypes influences students' decisions for or against STEM. For instance, women might be less motivated to decide in favor of STEM because according to stereotypes, they are expected to be less talented and interested in STEM professions.

One seminal theory which strives to explain human motivation is Self-Determination Theory (SDT). The theory states that there are external drivers/motivations (e.g., monetary incentives) and internal drivers/motivations (e.g., personal interests) for human behavior, and that internal motivations in particular can lead to persistence in goal achievement, such as graduating in a STEM major. SDT assumes that the motivation to engage in a specific behavior is dependent on whether this behavior is perceived to satisfy the needs for *competence, autonomy* and *relatedness*. The satisfaction of these basic psychological needs leads to high levels of intrinsic motivation, effective performance and well-being (Ryan and Deci, 2014). In this research, we suggest that specifically the perceptions of competence and autonomy are relevant for the selection of a study subject and might contribute to decisions for or against STEM. We do not focus on relatedness as we were more interested in the rather "self-centered" needs of competence and autonomy as potential drivers of gender differences regarding choice of study. Particularly if there are gender differences in perceived competence and autonomy in the context of study-related self-perceptions (operationalized as self-efficacy, self-esteem, leadership orientation and causality orientation regarding competence and autonomy), and in study interest and reasons for choosing a study subject, this may help to detect possible reasons for the gender gap in STEM. In the present study, we assess these constructs in students with STEM subjects in contrast to non-STEM subjects, with a special focus on differences between female and male undergraduates. By choosing a broad approach including motivations, interests as well as autonomy- and competence-related self-perceptions, we hope to gain further insights into the reasons why women are underrepresented in STEM.

Stereotype Reasons

Stereotypes about women and men exist in all cultures (e.g., Prentice and Carranza, 2002; Cuddy et al., 2008). In Western cultures, they include the assumption or expectation that women have a warm nature, are caring, gentle, and friendly, and act in a communal manner, while men are expected to be competitive, competent, goal-oriented, and mathematically skilled, and to act agentically (e.g., Cuddy et al., 2008). Importantly, individuals strive to fit into their gender role, because violating existing norms can have detrimental consequences such as discrimination and harm (e.g., Burgess and Borgida, 1999). In turn, this can have reinforcing effects on the individual who behaves unconventionally and is penalized, or behaves gender-congruently and is rewarded or at least not penalized. Possible effects of this may be that individuals adapt their behavior and/or change their self-related cognitions to fit into their gender roles.

The stereotype of a scientist is suggested to encompass many of the attributes that are also associated with stereotypical beliefs about males (Carli et al., 2016; Hand et al., 2017; Ramsey, 2017), such as being more agentic and less communal. However, Carli et al. (2016) note that the higher the proportion of female scientists in a particular field, the more the stereotype in that field mirrors stereotypical beliefs about women.

A meta-analysis of five decades of draw-a-scientist studies revealed that the scientist stereotype seems to evolve in childhood (elementary school, middle school) and increasingly strengthens with age (high school; Miller et al., 2018). The authors attributed this to children's/teenager's observation from their environment (e.g., textbooks, extracurricular experiences) in which female scientists are still underrepresented. This interpretation is supported by findings that male scientists are more often represented in TV formats than female scientists (e.g., Long et al., 2010; Steinke and Tavarez, 2018). Similar stereotypical beliefs apply for STEM professionals in general and have been demonstrated in high-school students and their teachers (e.g., Hand et al., 2017), college students (e.g., Piatek-Jimenez et al., 2018) and even among STEM professionals themselves (e.g., Farrell and McHugh, 2017). Notably, while most investigated groups show an implicit pro-male STEM bias, female STEM professionals demonstrate a slight pro-female gender bias (Farrell and McHugh, 2017). Nevertheless, the majority of findings point in favor of men.

Given that stereotypes, such as gender stereotypes, can affect actual perception and behavior (Bem, 1993; Master et al., 2017), it is easily conceivable that stereotypes about scientists can also influence women's and men's behavior, for example when they decide whether or not to engage in science, technology, engineering and math. Master and Meltzoff (2016) argued that such stereotypes can act as barriers preventing girls from studying STEM subjects. Based on this assumption, they conducted two experimental studies in which they varied the classroom environment, making stereotypes about computer scientists either salient (e.g., Star Trek posters) or not (e.g., nature posters). The authors found that boys' interest in computer science was not affected by the classroom design, but girls' interest was affected: The girls were three times more interested in computer science when they were in the non-stereotypical compared to the stereotypical environment. This is in line with findings regarding the "stereotype threat," which claims that stigmatized groups are threatened by stereotypes and perform less effectively when stereotypes are primed. For example, Spencer et al. (1999) found that women performed worse in solving a mathematical problem when they were made aware of the stereotypical belief that females are less mathematically skilled than men. In addition, Steffens et al. (2010) demonstrated that the knowledge about STEM stereotypes (here, math-gender stereotype) can become influential from the age of nine. The authors found that this knowledge affects academic achievements, enrolment preferences and the academic self-concept of girls but not of boys.

There is a reasonable amount of available knowledge concerning the stereotypes about STEM professionals and concerning which structural features can influence the decision for STEM study subjects, for example sufficient financial coverage or 12th-grade math achievements (Wang, 2013). However, less is known about the self-perceptions of STEM professionals or those who wish to become STEM professionals (STEM students). Moreover, women and men in STEM may perceive their environment, such as the campus climate, differently (Gayles and Ampaw, 2014), which can hinder their graduation. Wang and Degol (2013) stated that individuals make rational choices based on their abilities in order to maximize their outcomes relative to their costs, and suggested that these individual expectations to succeed influence their decisions regarding STEM or any other academic path. In sum, individuals' introspection concerning their motivation, abilities and capabilities might be a decisive factor influencing women's and men's paths into STEM professions. An influential theory which specifically focuses on motivations in order to be able to explain people's behaviors and choices is SDT which was therefore chosen as a theoretical framework to specifically target people's competence and autonomy.

Self-Determination Theory

Deci and Ryan (2011) proposed SDT as a macro theory consisting of six different mini-theories, such as *cognitive evaluation theory*, which focuses on the increase in intrinsic motivation, or *basic psychological needs theory*, which represents the core of SDT (for an overview see Deci and Ryan, 2011). Deci and Ryan (2000) described the needs for *competence, autonomy* and *relatedness,* which can be satisfied or thwarted by social contexts (Ryan and Deci, 2014). These needs are pivotal sources of motivation, effective performance and psychological well-being. We assume that the perceptions of specifically *competence* and *autonomy* as more "self-centered" needs may serve as valuable sources for making decisions in favor of or against STEM and that they might develop differentially in women and men.

Autonomy refers to the feeling of volitional self-regulation of behavior (Legate and Ryan, 2014): "Autonomy concerns acting from interest and integrated values. When autonomous, individuals experience their behavior as an expression of the self, such that, even when actions are influenced by outside sources, the actors concur with those influences, feeling both initiative and value with regard to them." (Ryan and Deci, 2002, p. 8). Furthermore, autonomy has been shown to positively contribute to interest in a subject (Black and Deci, 2000).

Competence refers to the experience of capability and effectiveness to achieve desired goals (Ryan and Deci, 2014). It "... leads people to seek challenges [...] and to persistently attempt to maintain and enhance those skills and capacities through activity. Competence is not, then, an attained skill or capability, but rather is a felt sense of confidence and effectance in action." (Ryan and Deci, 2002, p.7). According to Feltman and Elliot (2014), the perception of competence influences achievement outcomes, which can have an approaching character (success) or an avoiding character (failure). Striving for an approaching outcome can lead to creativity, optimal performance attainment and persistent interest. Avoiding outcomes, by contrast, lead to the opposite and to a higher probability of seeking easy rather than difficult goals/challenges. Jang et al. (2009) revealed that perceived competence is positively associated with academic achievements, and Wang and Degol (2013) summarized that girls who are confronted with STEM tasks are particularly vulnerable to perceiving themselves as lacking capability or as less competent.

Steele and Aronson (2005) stated that competence has a fragile nature, because a person who is confident in one task is not necessarily confident in another task (e.g., being good at art but a poor athlete). This fragility also reveals itself in the impact of stereotype threat (e.g., Spencer et al., 1999; Shapiro and Williams, 2012), as it shows that a specific competence (e.g., mathematical skills) can be shaken by mere priming.

One factor that shapes the perception of competence and autonomy refers to how individuals attribute the results of different actions. Deci and Ryan (1985) distinguished three types of causality orientations, which are uniquely represented in each individual: *autonomy orientation*, *controlled orientation* and *impersonal orientation*. Autonomy orientation refers to an internal attribution style in which the individual feels that he/she is in charge of his/her actions and consequences. Controlled orientation refers to an external attribution style in which environmental cues are held accountable for actions and outcomes. Impersonal orientation refers to the feeling that actions and consequences are beyond the individual's control, and is associated with anxiety regarding incompetence (Ryan and Deci, 2014). High ratings in autonomy orientation are positively correlated with professional satisfaction and academic achievements (Stipek and Weisz, 1981; Findley and Cooper, 1983; Ng et al., 2006); controlled orientation is associated with rigidity and lower levels of wellness (Ryan and Deci, 2014) and shows a small negative correlation with academic performance (Lane et al., 2004). Both concepts are measured in the present study in order to assess the individual perception of competence and autonomy.

Research has shown that women tend to possess an attribution style in which especially success is attributed to external cues, while men tend to have an internal attribution style in which especially success is attributed to oneself (e.g., ability; Burgner and Hewstone, 1993). However, the attribution style depends on the domain: "with boys citing more competency in traditionally masculine activities (sports and math) and girls citing more competency beliefs in traditionally feminine activities (reading and music...)." (Mezulis et al., 2004, p. 714). Given that STEM is assumed to reflect a male domain, it is conceivable that based on gender stereotypes and STEM stereotypes, women in STEM possess a rather self-derogating attribution style (e.g., low autonomy orientation), in contrast to their male counterparts. Early research on female engineering majors revealed that one reason for dropping out may be that female students tend to attribute their successes to external causes rather than to their own capabilities (Nauta et al., 1999).

Competence and Autonomy Perception

Another variable which is classically related to gender differences in male-dominated fields such as STEM (Zeldin et al., 2008) is self-efficacy, which contributes to the perception of competence. Bandura (1982) defined self-efficacy as the individual's belief about her/his ability to solve a task or reach a goal. It determines whether or not an individual wishes to engage in an activity and the degree of effort and perseverance the individual invests. Furthermore, self-efficacy is predictive of academic development in terms of academic aspirations, performance and persistence (Multon et al., 1991; Bandura, 1993; Lane et al., 2004; Zajacova et al., 2005). Focusing on self-efficacy in mathematical problem solving, Pajares and Miller (1994) revealed that self-efficacy beliefs are more important than, for instance, prior experience with mathematics. In a longitudinal study, Marra et al. (2009) found that women with high self-efficacy ratings are more often willing to persist in male-dominated fields such as STEM. Moreover, Zeldin et al. (2008) found that male STEM professionals gain their self-efficacy beliefs from prior mastery experience and previous successes, while women rely on vicarious experiences (e.g., observing similar others in performing) and verbal persuasion (e.g., positive encouragement from others).

While self-efficacy is a broad concept, academic self-efficacy refers to goal achievement in the academic context and captures the individual's specific beliefs about his/her confidence in solving academic tasks (Bong and Skaalvik, 2003). Similar to self-efficacy, academic self-efficacy is positively correlated with, for instance, academic performance and coping behavior (Chemers et al., 2001; Robbins et al., 2004; Gore, 2006). Zeldin et al. (2008) found that girls indicated having lower levels of science self-efficacy, which can predict lower interest in science. In a longitudinal study of STEM undergraduates who are underrepresented in STEM education and professions, MacPhee et al. (2013) reported that at the time of admission, women perceived themselves to have lower academic skills than their male counterparts, even though they did not show a weaker performance. By the time of graduation, however, women's academic self-concept was equal to that of men – at least for the specific sample that was enrolled in a mentoring program. Referring to previous findings and stereotype research, it can be assumed that women, in contrast to men, lack self-efficacy beliefs (general and academic) and that this in turn may lead to a decreased desire to enter a STEM study program and less confidence in one's success as a STEM professional.

Self-esteem represents another concept which can contribute to perceptions of competence and particularly to autonomy, because it – unlike self-efficacy - does not directly refer to capabilities. Self-esteem more broadly describes personal attitudes concerning one's self-worth (Rosenberg et al., 1995). While Lane et al. (2004) revealed a positive correlation between self-esteem and academic success, other authors, such as Baumeister et al. (2003), reported that self-esteem is not a good predictor of academic success but does sometimes predict job performance. However, these authors acknowledged that high self-esteem scores can facilitate further engagement after failure.

A further variable which is – albeit more loosely - associated with one's own perception of competence are leadership aspirations (Bass and Stogdill, 1990). As described by Ryan and Deci (2002) the perception of competence leads people to seek challenges. With regard to the challenge of aspiring leadership positions in STEM fields, however, there seems to be a distinct difference between men and women. The general lack of women in leadership positions (Amon, 2017) is even more pronounced in STEM fields (McCullough, 2011). On the one hand, this is caused by the lower number of women in STEM *per se* (McCullough, 2011), and on the other hand, perceived stereotypes and gender roles might also play a role. Amon (2017) found that female STEM graduates and postdoctoral researchers perceive various barriers concerning leadership in STEM, especially regarding the role conflict between being a woman and being in a leadership position, which is perceived as challenging. In particular, the effort that is perceived to be required for role transitioning (Amon, 2017) might lead women to balance costs and benefits, leading to fewer leadership aspirations than in men.

In conclusion, stereotype research suggests that stereotypes about STEM professionals resemble stereotypes about men and stand broadly in contrast to beliefs about women. These beliefs might be crucial, especially for women, when making decisions for or against STEM professions, for which choosing STEM subjects at university/college represents an important step. We draw on SDT as a theory on basic psychological needs which are described as pivotal sources of motivation and effective performance. We therefore measure the orientation specifically toward the here relevant "self-centered" needs of autonomy and competence by scales developed by Ryan and Deci (2014) and additionally assess the related constructs of self-efficiency, self-esteem and leadership orientation. Most importantly, we directly assess students' perception of their motivation to choose their major and their study motivation. Adding to self-determination-theory measures, we rely on measures stemming from expectancy-value models of motivation

(Kosovich et al., 2015). Also, as it was shown that decision making for educational trajectories represent complex processes (Becker and Hecken, 2009), that comprise, besides interests and abilities, motives for status maintenance (e.g., that families expect their children to follow their occupation), we added these aspects to classical motivation scales.

In a first step, we test whether there are general differences between STEM and non-STEM students. Therefore, we pose the following research question:

RQ1: How do STEM students (in contrast to non-STEM students) perceive their competence and autonomy? How do they perceive their self-efficacy, self-esteem and leadership aspiration and what study motivation and motivation for choosing their major do they report?

Moreover, research indicates that women in STEM face gender-STEM stereotypes, which can result, for instance, in stereotype threats or potential role conflicts (Amon, 2017), potentially affecting women's self-perceptions and well-being. In view of such negative outcomes, it seems reasonable to ask whether female STEM students differ not only from female non-STEM students but also from male STEM and male non-STEM students – potentially showing greater motivation, competence and autonomy than the other groups. We therefore ask:

RQ2: How do female STEM students differ from female non-STEM students, male STEM students and male non-STEM students concerning their perceived competence and autonomy? How do they perceive their self-efficacy, self-esteem and leadership aspiration and what study motivation and motivation for choosing their major do they report compared to the other groups?

In order to scrutinize gender differences specifically within the group of STEM students, we ask the following research question:

RQ3: How do female and male STEM students differ in their perceived competence and autonomy? How do they differ regarding their self-efficacy, self-esteem and leadership aspiration and do they report different study motivations and motivations for choosing their major?

MATERIALS AND METHODS

Procedure

To address our research questions, we conducted a survey study among Master's students of the University of Duisburg-Essen, Germany. Approximately 43.001 students are enrolled (48% women; retrieved from uni-due.de, 2019) in at least one study program at the university, which ranks number eight in German universities concerning its number of students. Approximately 9600 master students were enrolled in all disciplines at the moment of conduct (2017). The disciplinary canon of faculties of the university comprises: humanities, social sciences, educational sciences, economic sciences, business administration, mathematics, physics, chemistry, biology, engineering, and medicine.

According to national guidelines no ethics vote was required for the present survey; however, the study adheres to ethical standards, which are made transparent through the following descriptions. Participants were fully debriefed about the purpose of the study and participants did not face any consequences if they canceled the survey. However, since the study was issued as an official survey by the university it was presented, discussed and approved by the rectorate of the University and the data protection officer. The study was part of a larger survey study focusing on the careers of young academics, which also included post-docs and Ph.D. candidates (these two groups are not focused on in the current paper). The students were invited to take part in an online survey to gather data on their study motivation, their perceptions of their own competence and autonomy, as well as sociodemographic characteristics (more details below). At the beginning of the survey participants were informed about the aim of the survey (to collect data about individual perceptions of their study programs and their career aspirations) and that their data will be saved anonymously. By clicking a check box they gave informed consent that they are of age and permit to capture their data, afterward they got access to the rest of the survey. Those not accepting these requirements were free to quit the survey without any consequences. After completing the survey participants were fully debriefed about the focus on gender differences. We did not inform about this in advance in order to prevent priming of gender stereotypical beliefs, which could have influenced responses. Participants did not receive any immediate payment but had the opportunity to take part in a prize draw. The potential winnings included minor amounts of money (e.g., 10 Euros up to 50 Euros, in total 500 Euros) and the chance to win a book allowance worth 1.000 Euros (split into 2 × 500 Euros). Students needed approximately 30 min to complete the survey.

Measures

Sociodemographic Characteristics

The survey began with questions concerning students' sociodemographic characteristics such as sex (biological category), nationality, and age. Furthermore, we asked about their school career, professional career, and academic career as well as of that their parents. All categories are based on the German school, academic and professional system and can only be partially compared to other countries' systems. In cases in which no international equivalent exists we will present translations. The participants could only choose one category at a time.

Firstly, participants were asked at which type of school they received their university entrance exams/level (e.g., "Gymnasium [secondary school], "Gesamtschule [comprehensive school]," "received in a foreign country"). Afterward they were asked "where did you receive your first university degree (e.g., Bachelor's degree)?" and could respond by the choices "at the University of Duisburg-Essen," "at another university in Germany" and "at another foreign university." In addition, we assessed in which field of study they received their first university degree, for instance, "Humanities," "Engineering," "Law Studies." Finally, we asked for their final grade ("≤1.4; 1.5 – 1.9; 2.0 – 2.4; 2.5 – 2.9; 3.0 – 3.4; 3.5 – 4.0"; the lower the number, the better the grade with 1,0 representing the best grade and 4.0 representing the worst pass grade in the German tertiary education).

Concerning the educational and professional trajectories of participants' parents we asked for the following choices within the school trajectory sections: "mittlere Reife" [graduating from a medium-track school], "Fachhochschulreife oder Allg. Hochschulreife oder Ähnliches [graduading from a higher-track school/qualification for university entrance or something comparable]," "no graduation," "something else," "I don't know." The categories of the academic education comprised: "university degree," "doctorate"; "no academic degree," "something else," "I don't know." The categories in the professional career are "abgeschlossene Berufsausbildung [completed vocational training]," "weiterführende berufsqualifizierende Ausbildungsgänge [advanced vocational trainings]," "no completed vocational training," "something else" and "I don't know."

Reasons for Choosing Majors

We asked for specific reasons (eight items) why students made the decision for their majors, including intrinsic (e.g., "my subject comes easily to me") and extrinsic orientations (e.g., "I can make a lot of money") and reasons, which stem from educational studies and seem to be decisive for study program choice alike (e.g., Ramseier, 2006; Kretschmann et al., 2017). In addition, we created ad-hoc items to assess motives for status maintenance (Becker and Hecken, 2009) and the autonomy of the decision. With this, we aimed to potentially show (1) a broader range of reasons for participants' decisions and (2) to examine whether there are differences depending on gender and major. These items are such as "I chose my major because my family wanted me to" or "I chose my major because I did not get a place on my preferred study program" (for all items, means and standard deviation see **Table 1**). Items were rated on a 5-point rating scale (1 = not at all; 5 = absolutely). Due to their heterogeneity, for further analysis we consider the single items; no factors were formed.

Study Interest

To indicate their specific interest in their subject, students completed the Fragebogen zum Studieninteresse [Study Interest Questionnaire] by Schiefele et al. (1993). This short scale consists of nine items (α = 0.727), such as "If I had enough time, I would be more concerned with certain issues of my studies, beyond the exams" or "Even before my studies, the subject had a special significance for me." Items were rated on a 4-point rating scale (0 = not at all; 3 = absolutely).

TABLE 1 | Descriptive statistics of reasons for study subject.

Reasons	*M*	*SD*	*N*
I know exactly what I will do as a professional after graduation	2.32	1.12	888
It reflects my interests	3.57	0.63	886
My subjects comes easily to me	3.25	0.70	887
I did not know what else to study	1.74	0.94	888
I can make a lot of money after graduation	2.26	0.98	887
I will have a lot of professional opportunities after graduation	2.82	0.91	886
My family wanted me to choose this major	1.37	0.72	888
I did not get a place on my preferred study program	1.42	0.82	888

Study Motivation

To gather data on study motivation in terms of expectancy, value, and cost of study, we adapted the Expectancy-Value-Cost (EVC) scale by Kosovich et al. (2015) to refer to students' study programs. Each of the three dimensions was measured with two items. For example, expectancy (α = 0.695) was measured with the statement "I believe that I can be successful in my study program," value (α = 0.861) was measured with "I value my major," and costs of study (α = 0.667) was measured with "I have to give up too much to do well in my major." Items were rated on a 4-point scale (1 = strongly disagree; 4 = strongly agree).

Causality Orientation

To examine participants' causality orientation, we used three selected situational vignettes by Deci and Ryan (1985, retrieved from selfdeterminationtheory.org, 2017), which are oriented toward achievement situations. Participants were asked to imagine a fictitious situation from day-to-day life and to make suggestions about their feelings and thoughts. For example: "You had a job interview several weeks ago. In the mail you received a form letter which states that the position has been filled. It is likely that you might think:. . .", with the three responses "It's not what you know, but who you know" (controlled orientation); "I'm probably not good enough for the job." (impersonal orientation) and "Somehow they didn't see my qualifications as matching their needs" (autonomy orientation). Participants had to make a choice on each statement on 7-point scales (1 = very unlikely that the participants would respond in this way; 4 = moderately likely; 7 = very likely), which represented one of three causal dimensions. Afterward the scores for the respective statements were added into these subscales (controlled orientation, α = 0.280; autonomy orientation, α = 0.360; impersonal orientation, α = 0.506) resulting in ratings from 3 (low manifestation within this orientation) to 21 (high manifestation within this orientation).

Self-Efficacy

General self-efficacy was measured with 10 items from the Skala zur Allgemeinen Selbstwirksamkeitserwartung [Generalized Self-Efficacy Scale] by Jerusalem and Schwarzer (1999), such as "I feel comfortable with difficulties because I can always rely on my abilities." Items were rated on a 4-point rating scale (1 = disagree; 4 = agree). Scores were summed up into one scale (α = 0.812) ranging from 10 (low self-efficacy) to 40 (high self-efficacy).

We captured academic self-efficacy by using the BWS Skala [Occupational Self-Efficacy Scale] by Abele et al. (2000); we adapted the scale to refer to academia. The scale focuses on how to solve requirements and difficulties related to study, such as "I do not know if I really have the skills to study." Items were rated on a 5-point rating scale (1 = not at all; 5 = absolutely; α = 0.765).

Self-Esteem

Self-esteem was measured using the 10-item scale by Collani and Herzberg (2003), a German version of the Rosenberg Self-Esteem Scale. Items comprise statements such as "On the whole, I am satisfied with myself"; statements were rated on a 4-point rating scale (1 = strongly disagree; 4 = strongly agree). Scores of items were added into a sum score resulting in a scale ranging from 10 (low self-esteem) to 40 (high self-esteem; α = 0.864).

Leadership Aspirations

Leadership aspirations were captured by asking students the question (Powell and Butterfield, 2013) of which positions they would wish to hold in future: top management, middle management, lower management or employee without a leadership position.

Sample

The final sample comprised 888 Master's students (461 female, 427 male), who would be graduating either in STEM (physics, chemistry, biology, engineering, mathematics) or non-STEM (humanities, social sciences, educational sciences) subjects. We excluded nine participants because they refrained from indicating their gender or categorized themselves as transsexual/intersexual/queer. A further three were excluded due to insufficient data quality (less than 60 percent of questions answered).

Students ranged in age from 20 to over 49; the majority were 25–29 years old (57.3%), followed by the group of 20–24 year-olds (30.3%). Eighty percent indicated that they were German nationals, while the remaining 20% came, for instance, from India, Turkey, or European countries. Students graduated (Bachelor's degree or something comparable) from different disciplines: humanities (n = 190), social sciences (n = 148), art, music, design (n = 6), economic science (n = 26), natural science (n = 150), law studies (n = 1), engineering (n = 355), others (8). Four persons refrained from answering this question.

With regard to students' family background, the largest proportion of students' mothers had graduated from a medium-track school of the German three-tier school system (48.6%), followed by a higher-track school, corresponding to university entrance level (42.3%). Most of the mothers did not hold a university degree (65.3%), while approximately one third were university graduates (29.1%). More than half of the mothers had received vocational training (54.4%). By contrast, the largest proportion of students' fathers had graduated from a higher-track school (50.1%), followed by a medium-track school (40%). Approximately half of the fathers did not have an academic degree (51.6%), while 42.3% were university graduates and 42.4% had received vocational training.

RESULTS

Before testing the research questions, we checked for assumptions (e.g., normally distributed data and moderate correlations of dependent variables, see Field, 2013) to run appropriate analyses. To examine RQ1-RQ3, we conducted MANOVAs, ANOVAs as well as Chi^2 tests; detailed descriptions of tests and results will be consecutively reported. We used the whole sample (N = 888) for testing RQ1 and RQ2, while for testing RQ3 we refer to a smaller sample consisting of STEM students (n = 529). We will sequentially report results from RQ1 to RQ3.

Effects of Subject

Effects of Subject on Reasons for Choosing Study Subjects

Regarding reasons for choosing their study subjects, we conducted 2 × 2 ANOVAs with subject (STEM vs. non-STEM) and gender (female, male) as independent variables and reasons as dependent variables (see **Table 2**), because assumptions for MANOVA testing were not fully met (e.g., low or no correlations between variables). Moreover, for dependent variables which did not meet assumptions for ANOVA testing we conducted Chi^2 tests.

The single ANOVA referring to the statement that "their subject comes easily to them" revealed that non-STEM students agreed slightly stronger [M = 3.39; SD = 0.65; Levene's test $F(3,883) = 0.43$, $p = 0.730$, for all variance terms see **Table 2**] with the statement than the STEM students (M = 3.16; SD = 0.72).

The ANOVA concerning the statement "I can make a lot of money after graduation" is significant (see **Table 2**). Although the Levene's test is significant, $F(3,883) = 6.30$, $p < 0.001$, we adhere to the analysis because the ANOVA is fairly robust to violation of homogeneity (Bortz and Schuster, 2010; Field, 2013). STEM students agreed with the statement more often than non-STEM students ($M_{\mathrm{STEM}} = 2.60$; $SD = 0.92$; $M_{\mathrm{non-STEM}} = 1.76$; $SD = 0.82$).

In addition, STEM students agreed more often to the reasoning that they will have a lot of professional opportunities after graduation [$M_{\mathrm{STEM}} = 3.09$; $SD = 0.85$; $M_{\mathrm{non-STEM}} = 2.44$; $SD = 0.86$; Levene's test $F(3,882) = 3.86$, $p = 0.009$]. No significant difference was found for the item "I know exactly what I will do as a professional after graduation" [Levene's test $F(3,884) = 11.08$, $p < 0.001$].

Before conducting Chi^2 tests for the remaining reasons we checked for distributions of answers in each cell (2 × 5) resulting from answers in 5 response categories (1 = not at

TABLE 2 | Effects of subjects, gender and subject*gender on reasons for choosing study subject.

Reasons	*df*	*F*	*p*	η^2
My subject comes easily to me				
Subject	1,883	19.86	< 0.001	0.022
Gender		0.00	0.948	< 0.001
Subject × Gender		1.03	0.310	0.001
I can make a lot of money after graduation				
Subject	1,883	135.78	< 0.001	0.133
Gender		4.52	0.034	0.005
Subject × Gender		1.28	0.259	0.001
I will have a lot of professional opportunities after graduation				
Subject	1,882	88.46	< 0.001	0.091
Gender		2.51	0.114	0.003
Subject × Gender		4.97	0.026	0.006
I know exactly what I will do as a professional after graduation				
Subject	1,884	0.39	0.530	< 0.001
Gender		1.33	0.250	0.001
Subject × Gender		2.46	0.117	0.003

all; 5 = absolutely) and two subjects (STEM vs. non-STEM). Noteworthy, distributions to the remaining reasons are left skewered and no answers in category 5 (absolutely) were provided resulting in a maximum of 2 × 4 cells. In cases in which cell counts were below 5 we collapsed the respective category with the next adjacent category, e.g., when response counts in category 4 were below 5 we collapsed it with category 3.

The Chi^2 test is not significant, $\chi^2(3) = 5.81$, $p = 0.121$, $\varphi = 0.08$, for the reason "I did not get a place on my preferred study program" indicating that STEM and non-STEM students do not differ in their agreement to this statement.

Referring to the reason that the study subject reflects participants' interest we merged categories 1 and 2, because of counts less than 5 in category 1. The Chi^2 test was significant, $\chi^2(2) = 8.81$, $p = 0.012$, $\varphi = 0.100$. There is no difference between STEM (6.1%) and non-STEM (4.5%) students in choosing category 2; indicating that there is no difference depending on student's major concerning low agreement to the statement. However, STEM students more often decided for category 3 ($35.0\%_{\text{STEM}}$; $26.8\%_{\text{non-STEM}}$) indicating that they more often agreed moderately to the statement that they chose their major due to their interest. By contrast, non-STEM students decided more often for a stronger agreement (category 4) to the reason than STEM students ($58.9\%_{\text{STEM}}$; $68.7\%_{\text{non-STEM}}$) indicating that non-STEM students rather chose their study subject based on their interests than STEM students.

Concerning the reason "I did not know what else to study" the test, $\chi^2(3) = 14.47$, $p = 0.002$, $\varphi = 0.13$, revealed that non-STEM students (60.5%) significantly decided more often for category 1, indicating more disagreement with this reason, than STEM students (50.9%).

Moreover, STEM students more often chose category 2 ($23.6\%_{\text{STEM}}$; $21.7\%_{\text{non-STEM}}$) and category 4 ($7.9\%_{\text{STEM}}$; $2.8\%_{\text{non-STEM}}$), while no significant difference for category 3 can be observed ($17.6\%_{\text{STEM}}$; $15.0\%_{\text{non-STEM}}$). These results point in the direction that STEM students agree more often to the statement that they chose their study program because they did not know what else to study than their non-STEM counterparts.

For the reason "my family wanted me to study the subject" we merged category 4 with category 3 because of zero counts in single cells of category 4. The analysis revealed a significant Chi^2 test, $\chi^2(2) = 45.15$, $p < 0.001$, $\varphi = 0.23$. STEM students significantly less often chose category 1 ($66.5\%_{\text{STEM}}$; $85.8\%_{\text{non-STEM}}$) but more often category 2 ($19.8\%_{\text{STEM}}$; $10.9\%_{\text{non-STEM}}$) and category 3 ($13.6\%_{\text{STEM}}$; $3.3\%_{\text{non-STEM}}$). This pattern indicates that STEM students agreed more often that they chose their study subject because their families wanted them to than their non-STEM counterparts.

Effects of Subject on Study Interest

Regarding study interest, we conducted a 2 × 2 ANOVA that revealed that non-STEM students ($M_{\text{non-STEM}} = 3.15$; $SD = 0.47$) indicated more interest than did STEM students [$M_{\text{STEM}} = 3.02$; $SD = 0.45$; Levene's test $F(3,872) = 0.45$, $p = 0.719$, variance terms see **Table 2**].

Effects of Subject on Study Motivation

To test for potential differences between STEM and non-STEM students concerning their study motivation we checked for assumptions to conduct a MANOVA with the scales three sub-dimensions (*value*, *expectancy,* and *cost*), however, these were not fully met (e.g., partly low correlations; no normal distributions of residuals). Therefore, we conducted Chi^2 tests for *expectancy* and *value* and an ANOVA for *cost*. For *expectancy* as well as *value* (2 = not at all, 8 = absolutely, resulting in categories 2–8) we had to collapse categories 2 and 3 (each below 5 counts) with category 4 to reach a critical number of counts. The test was not significant for *expectancy*, $\chi^2(4) = 5.04$, $p = 0.283$, $\varphi = 0.08$ indicating that STEM students and non-STEM students do not differ in their expectancies that they can be successful in their study program.

By contrast, the Chi^2 test for *value* $\chi^2(4) = 18.72$, $p < 0.001, \varphi = 0.15$ is significant. STEM students less often chose category 4 (1.5%) compared to non-STEM students (2.8%), in contrast STEM students more often decided for category 8 (37.5%) than non-STEM students (21.4%). Students did not differ by subject concerning the remaining categories (category 5: $2.0\%_{\text{STEM}}$, $2.1\%_{\text{non-STEM}}$; category 6: $11.5\%_{\text{STEM}}$, $9.7\%_{\text{non-STEM}}$; category 7: $7.1\%_{\text{STEM}}$, $4.2\%_{\text{non-STEM}}$). This pattern of results indicates that STEM students value their study program more than their non-STEM counterparts.

For the subdimension of *cost* we run an ANOVA resulting in a significant effect [Levene's test $F(3,882) = 1.61$, $p = 0.186$; variance terms see **Table 2**]. Means show that STEM students attribute more costs ($M = 5.55$, $SD = 1.58$) than non-STEM students ($M = 4.78$, $SD = 1.45$).

Effects of Subject on Causality Orientation

In respect to causality orientation we conducted a 2 × 2 MANOVA with the *control* and *impersonal orientation* as dependent variables after checking for assumptions. Since *autonomy* is not correlated to both other dimensions, we conducted a separate 2 × 2 ANOVA for this dimension.

The MANOVA revealed no significant differences concerning subject [Box's test $F(9,827532) = 8.65$, $p = 0.474$; variance terms see **Table 2**] for *impersonal* and *control orientations*, as neither did the ANOVA for the *autonomy orientation* [Levene's test $F(3,879) = 0.54$, $p = 0.658$; variance terms see **Table 2**]. This indicates that STEM and non-STEM students do not differ in the way they attribute the results of different actions.

Effects of Subject on Self-Esteem

Moreover, we conducted a 2 × 2 ANOVA with *self-esteem* as dependent variable. STEM students showed lower self-esteem ratings ($M_{\text{STEM}} = 31.87$; $SD = 5.62$) than did non-STEM students [$M_{\text{non-STEM}} = 32.78$; $SD = 5.55$; Levene's test $F(3,868) = 1.33$, $p = 0.268$, for variance terms see **Table 3**].

Effects of Subject on (Academic) Self-Efficacy

In addition, we run a 2 × 2 MANOVA with self-efficacy and academic self-efficacy [Box's test $F(9,813227) = 8.12$, $p = 0.57$], which resulted in insignificant effects of subject for both subdimensions (for variance terms see **Table 3**). STEM and

TABLE 3 | Effects of subjects, gender and subject*gender on study interest, study motivation, causality orientation, self-efficacy, academic self-efficacy, and self-esteem.

Construct	*df*	*F*	*p*	η^2
Study interest				
Subject	1,872	13.44	< 0.001	0.015
Gender		0.00	0.999	0.000
Subject × Gender		0.81	0.369	0.001
Study motivation				
Cost				
Subject	1,882	49.44	< 0.001	0.053
Gender		0.01	0.938	< 0.001
Subject × Gender		11.71	0.001	0.013
Causality orientation				
Autonomy				
Subject	1,879	0.95	0.329	0.001
Gender		10.73	< 0.001	0.012
Subject × Gender		2.62	0.106	0.003
Controlled				
Subject	1,877	3.61	0.058	0.004
Gender		22.03	< 0.001	0.025
Subject × Gender		0.85	0.357	0.001
Impersonal				
Subject	1,877	0.27	0.602	< 0.001
Gender		35.04	< 0.001	0.038
Subject × Gender		0.23	0.633	< 0.001
Self-efficacy				
Self-efficacy				
Subject	1,855	0.18	0.669	< 0.001
Gender		14.85	< 0.001	0.017
Subject × Gender		0.13	0.669	< 0.001
Academic self-efficacy				
Subject	1,855	3.45	0.064	0.004
Gender		2.65	0.104	0.003
Subject × Gender		1.14	0.285	0.001
Self-esteem				
Subject	1,868	7.28	0.007	0.008
Gender		1.09	0.297	0.001
Subject × Gender		2.36	0.125	0.003

non-STEM students seem not to differ concerning their perceived self-efficacy and academic self-efficacy.

Effects of Subject on Leadership Aspirations

Finally, we conducted a Chi^2 test to examine whether STEM students differ from non-STEM students regarding their leadership aspirations (4 response categories). The test yielded a significant difference, $\chi^2(3) = 24.93$, $p < 0.001$, $\varphi = 0.17$: STEM and non-STEM students did not differ regarding striving for positions as employees without leadership duties (8.9%$_{\text{STEM}}$; 7.5%$_{\text{non-STEM}}$) and lower-management positions (15.7%$_{\text{STEM}}$; 19.3%$_{\text{non-STEM}}$). However, non-STEM students were found to strive significantly more often for positions in middle management (59.5%$_{\text{non-STEM}}$; 48.2%$_{\text{STEM}}$), while STEM students strive significantly more often for top management positions (27.1%$_{\text{STEM}}$; 13.7%$_{\text{non-STEM}}$). In sum, therefore, STEM students aspire higher leadership positions than non-STEM students.

Interaction Effects of Subject × Gender

Interaction Effects of Subject × Gender on Reasons for Choosing Study Subjects

With regard to RQ2 (interactions of subject and gender), the ANOVAs revealed one significant interaction effect for the reason "I will have a lot of professional opportunities after graduation" (variance term see **Table 2**), indicating that male STEM students ($M_{\text{STEM}} = 3.18$; $SD = 0.79$) see the most opportunities after graduation, even more than their female counterparts ($M_{\text{STEM}} = 2.92$; $SD = 0.94$), while there was no difference in the lower ratings of female non-STEM students ($M_{\text{non-STEM}} = 2.45$; $SD = 0.88$) and male non-STEM students ($M_{\text{non-STEM}} = 2.40$; $SD = 0.78$). The other interaction effects were not significant (see **Table 2**).

Moreover, the Chi^2 tests revealed three effects for subject and gender. Ratings were given from 1 = "not at all" to 5 = "absolutely," mirroring in 5 response categories. Owing the fact that no answers were provided in category 5 we downsized response categories to a maximum of 4 categories.

For the reason "It reflects my interests" we merged response categories 1–3 due to lower counts than 5 in single cells, resulting in the two response categories: 3 and 4. The analysis revealed a significant effect, $\chi^2(3) = 11.98$, $p = 0.007$, $\varphi = 0.12$. Male STEM students (43%) more often chose category 3 than female non-STEM students (29.5%). By contrast, female non-STEM students (70.5%) select more often category 4 than male STEM students (57%). No other differences emerged (category 3: 37.6%$_{\text{female STEM}}$, 37.6%$_{\text{male non-STEM}}$; category 4: 62.4%$_{\text{female STEM}}$, 62.7%$_{\text{male non-STEM}}$). This response pattern indicates that the groups of female non-STEM students and male STEM students significantly differ from each other, while the other groups share some similarities. Female non-STEM students seem to choose their major more often because it reflects their interests, while male STEM students more often rather disagree with this reason.

To conduct the analysis for the reason "I did not get a place on my preferred study subject" we had to merge categories 4 through 2 because of low cell counts in categories 3 and 4, resulting in two response categories: 2 and 1. The test is significant, $\chi^2(3) = 13.52$, $p = 0.004$, $\varphi = 0.12$. Male non-STEM students chose more often category 1 (86.9%) than their male STEM counterparts (68.8%), while it is opposite for category 2 (13.1%$_{\text{male non-STEM}}$; 31.2%$_{\text{male STEM}}$). No further differences for the other categories emerged (category 1 = 76%$_{\text{female non-STEM}}$; 76.9%$_{\text{female STEM}}$; category 2: 24%$_{\text{female non-STEM}}$; 23.1%$_{\text{female STEM}}$). It seems that male STEM students agree slightly more to the reason that they chose their study program because they did not get a place on their preferred study program than their male non-STEM counterparts, while no difference concerning the female groups can be observed.

We collapsed categories 3 and 4, resulting in 3 response categories (see **Table 4**) for the item "My family wanted me to," to run the Chi^2 test. The analysis revealed a significant

TABLE 4 | Crosstable for the reason "My family wanted me to."

	Category	Female non-STEM	Female STEM	Male non-STEM	Male STEM
1,00	Count	241_a	$137_{b,c}$	$67_{a,c}$	215_b
	Expected count	204.4	138.2	62.4	254.9
	% within	87.6%	73.7%	79.8%	62.7%
	2,00 Count	27_a	$30_{a,b}$	$12_{a,b}$	75_b
	Expected count	44.6	30.2	13.6	55.6
	% within	9.8%	16.1%	14.3%	21.9%
3,00	Count	7_a	19_b	$5_{a,b}$	53_b
	Expected count	26.0	17.6	7.9	32.4
	% within	2.5%	10.2%	6.0%	15.5%

*Each subscript letter denotes a subset of Gender*Subject categories whose column proportions do not differ significantly from each other at the 0.05 level.*

effect, $\chi^2(6) = 55.17$, $p < 0.001$, $\varphi = 0.18$. Referring to the column's proportions (see **Table 4**) female non-STEM students chose category 1 more often than female STEM students and male STEM students. In addition, the male non-STEM students more often disagree with this reason than the male STEM students. These results indicate that the groups of female and male non-STEM students more often disagree with the statement that they chose their study program because their families wanted them to than male STEM students. Within category 2 the only difference occurred between male STEM students and female non-STEM students, indicating that male STEM students chose this category more often. Concerning category 3 male STEM students are more often in this category compared to female non-STEM students. In addition, female STEM students chose this category more often than female non-STEM students. This pattern of results indicates that predominantly male STEM students, but also female STEM students chose their study program because their families wanted them to in contrast to the group of female non-STEM students.

Moreover, the analysis of the statement "I did not know what else to study" was not significant, $\chi^2(6) = 10.69$, $p = 0.099$, $\varphi = 0.08$, indicating that none of the groups differed in the agreement with this statement.

Interaction Effects of Subject × Gender on Study Motivation

The ANOVA concerning the subdimension of *costs* within the study motivation construct revealed an interaction effect for subject and gender (variance term see **Table 3**). The means indicated that men in STEM scored highest regarding expected costs ($M = 5.68$; $SD = 1.54$), followed by women in STEM ($M = 5.30$; $SD = 1.54$), women in non-STEM ($M = 4.88$; $SD = 1.45$) and men in non-STEM ($M = 4.47$; $SD = 1.42$).

Moreover, we analyzed the interaction of subject and gender for the *value* subdimension after collapsing categories 2 through 5 (2 = not at all, 8 = absolutely), resulting in the response categories 5–8. The Chi^2 test is significant, $\chi^2(9) = 23.70$, $p = 0.005$, $\varphi = 0.09$. Referring to the column's proportions female and male non-STEM students more often decide for category 5 than female STEM students and male STEM students, indicating that non-STEM students more often agree on a lower moderate level concerning study program value.

TABLE 5 | Crosstable for the subdimension value of study motivation scale.

	Category	Female non-STEM	Female STEM	Male non-STEM	Male STEM
5,00	Count	33_a	8_b	12_a	23_b
	Expected count	23.6	16.0	7.1	29.3
	% within	12.0%	4.3%	14.5%	6.7%
6,00	Count	$62_{a,b}$	$40_{a,b}$	24_b	62_a
	Expected count	58.4	39.5	17.6	72.6
	% within	22.5%	21.5%	28.9%	18.1%
7,00	Count	28_a	17_a	9_a	46_a
	Expected count	31.0	21.0	9.4	38.6
	% within	10.2%	9.1%	10.8%	13.5%
8,00	Count	$152_{a,b}$	121_c	38_b	$211_{a,c}$
	Expected count	162.0	109.6	48.9	201.5
	% within	55.3%	65.1%	45.8%	61.7%

*Each subscript letter denotes a subset of Gender*Subject categories whose column proportions do not differ significantly from each other at the 0.05 level.*

Moreover, in category 6 the only difference emerged for male STEM students and male non-STEM students indicating that male non-STEM students are more often in this category. Within category 7 no difference between groups proportions occurred.

For category 8 the analysis revealed that female STEM students chose this answer more often than female non-STEM students and male non-STEM students. In addition, male STEM students are more often in this category than male non-STEM students (for proportions see **Table 5**). No difference emerged between female and male STEM students. This pattern of results indicates that STEM students, especially female STEM students differ from the non-STEM groups in the sense that they value their study subject more.

The analysis for the subdimension of *expectancy* is not significant, $\chi^2(6) = 9.04$, $p = 0.171$, $\varphi = 0.07$, indicating that the groups do not differ in their expected study success.

Interaction Effects of Subject × Gender on Study Interest, Causality Orientation, Self-Esteem, and Self-Efficacy

The interaction effects of gender and subject are not significant for the (sub-) dimensions of study interest, causality orientation, self-esteem and self-efficacy (for variance terms see **Table 2**), indicating that the groups cannot be distinguished by these variables.

Gender Differences Within STEM Students

To examine RQ3, we compared female and male STEM students referring to the same methods as for RQ1 and RQ2.

Effects of STEM Student's Gender on Reasons for Choosing Study Subject

The analyses identified several differences between female and male STEM students regarding the reasons why they chose their STEM subjects. Concerning the reason "My subject comes easily to me" the ANOVA revealed no significant effect, $F(1,527) = 0.60$, $p = 0.440$, $\eta^2 = 0.001$ [Levene's test $F(1,526) = 0.45$, $p = 0.833$], indicating that female and male STEM students do not differ in this regards.

By contrast, the ANOVA for "I can make a lot of money" is significant, $F(1,526) = 7.01$, $p = 0.008$, $\eta^2 = 0.013$ [Levene's test $F(1,526) = 6.94$, $p = 0.009$] indicating that male STEM students ($M_{male} = 2.68$; $SD = 0.87$) agreed slightly more often than female STEM students ($M_{female} = 2.46$; $SD = 0.99$).

The same pattern of responses emerged for "I will have lot of opportunities after graduation," $F(1,525) = 7.56$, $p = 0.001$, $\eta^2 = 0.020$ [Levene's test $F(1,525) = 3.16$, $p = 0.076$]. Male STEM students ($M_{male} = 3.18$; $SD = 0.79$) agree slightly stronger than female STEM students ($M_{female} = 2.92$; $SD = 0.04$).

The ANOVA for the item "I know exactly what I will do professionally after graduation" points in the same direction, $F(1,527) = 6.10$, $p = 0.014$, $\eta^2 = 0.011$ [Levene's test $F(1,527) = 0.36$, $p = 0.546$] that male STEM students ($M_{male} = 2.39$; $SD = 1.02$) agreed slightly more than their female counterparts ($M_{female} = 2.16$; $SD = 1.08$).

After collapsing categories 1 and 2, resulting in response categories 2–4, of the item "it reflects my interests," due to low cell counts, the Chi^2 test was not significant, $\chi^2(2) = 1.44$, $p = 0.487$, $\varphi = 0.05$, indicating that women and men in STEM do not differ in their response behavior.

Moreover, none of the remaining reason analyses were significant: "My family wanted me to," $\chi^2(3) = 6.77$, $p = 0.080$, $\varphi = 0.11$, "I did not know what else to study," $\chi^2(3) = 0.25$, $p = 0.968$, $\varphi = 0.02$, and "I did not get a place on my preferred study program," $\chi^2(3) = 4.04$, $p = 0.258$, $\varphi = 0.09$, indicating that male and female STEM students cannot be discriminated by these variables.

Effects of STEM Student's Gender on Study Interest

The ANOVA focusing on study interest as dependent variable was not significant, $F(1,524) = 0.61$, $p = 0.436$, $\eta^2 = 0.001$ [Levene's test $F(1,524) = 0.31$, $p = 0.861$], indicating that female and male STEM students show the same interest in their study program.

Effects of STEM Student's Gender on Study Motivation

We run an ANOVA with the subdimension *cost* as dependent variable, which revealed a significant effect, $F(1,526) = 7.69$, $p = 0.006$, $\eta^2 = 0.014$ [Levene's test $F(1,527) = 0.06$, $p = 0.815$] indicating that male STEM students ($M = 5.30$, $SD = 1.54$) score higher on this scale than female STEM students ($M = 5.68$; $SD = 1.54$).

For the Chi^2 tests of *expectancy* and *value* we collapsed categories 2 through 5, resulting in response categories 5–8, due to low cell counts. The analyses revealed no significant effects, neither for *expectancy*, $\chi^2(2) = 0.12$, $p = 0.941$, $\varphi = 0.02$, nor for *value*, $\chi^2(3) = 4.01$, $p = 0.260$, $\varphi = 0.09$, indicating that female and male STEM students have similar expectancies toward their study success and the value of their study program.

Effects of STEM Student's Gender on Causality Orientation

The ANOVA concerning the subdimension *autonomy* is not significant, $F(1,523) = 2.00$, $p = 0.157$, $\eta^2 = 0.004$ [Levene's test $F(1,523) = 0.16$, $p = 0.693$] indicating that female and male STEM students do not differ in the way they attribute results of actions to themselves.

Moreover, the MANOVA with its subdimensions *impersonal* and *control* [Box's test $F(3,4157256) = 5.08$, $p = 0.168$] revealed significant effects for both dimensions [impersonal: $F(1,523) = 30.23$, $p < 0.001$, $\eta^2 = 0.055$; control: $F(1,523) = 22.33$, $p < 0.001$, $\eta^2 = 0.041$]. The analysis showed that female STEM students score higher ($M = 16.23$; $SD = 2.80$) on *control orientation* than male STEM students ($M = 14.95$; $SD = 3.05$) indicating that female STEM students assume environmental cues more accountable for their actions and outcomes than male STEM students. The same response pattern occurred for the *impersonal orientation* ($M_{female} = 12.45$; $SD = 3.80$; $M_{male} = 10.62$; $SD = 3.56$), indicating that female STEM students have an increased feeling that actions and consequences are beyond their control in contrast with male STEM students.

Effects of STEM Student's Gender on Self-Esteem

The analysis referring to STEM student's self-esteem was not significant, $F(1,520) = 0.17$, $p = 0.687$, $\eta^2 < 0.001$ [Levene's test $F(1,520) = 2.80$, $p = 0.095$]. These result suggest that male and female STEM students do not differ in their self-esteem perception.

Effects of STEM Student's Gender on Self-Efficacy

The MANOVA concerning self-efficacy and academic self-efficacy [Box's test $F(3,3828798) = 6.14$, $p = 0.107$] revealed a significant effect for self-efficacy, $F(1,508) = 12.76$, $p < 0.001$, $\eta^2 = 0.025$, showing that male STEM students ($M = 29.54$; $SD = 4.21$) score higher on the scale than their female counterparts ($M = 28.10$; $SD = 4.61$). There is no significant effect for academic self-efficacy, $F(1,508) = 0.23$, $p = 0.631$, $\eta^2 < 0.001$, indicating no difference between female and male STEM students in this variable.

Effects of STEM Student's Gender on Leadership Orientation

Finally, the Chi^2 test revealed a significant difference in the leadership aspirations of female and male STEM students, $\chi^2(3) = 42.16$, $p < 0.001$, $\varphi = 0.28$. The frequencies show that female and male STEM students do not differ in striving for positions as employees without leadership duties ($10.8\%_{females}$; $7.9\%_{males}$). However, female STEM students more frequently strive for positions within lower management ($21.6\%_{females}$; $12.6\%_{males}$) and middle management ($57.3\%_{females}$; $43.3\%_{males}$), while male STEM students more often strive for top management positions ($36.3\%_{males}$; $10.3\%_{females}$) indicating higher leadership aspiration for males.

DISCUSSION

The current study focused on self-perceptions of STEM and non-STEM students with a special focus on female STEM students in terms of their study motivation, competence and autonomy. SDT suggests that confidence in one's own abilities and capabilities (competence) and acting volitionally (autonomy) can determine an individual's interest in a subject and academic achievements. The results of our survey study revealed that STEM and non-STEM students differ in their motivation, perceived competence and autonomy. A first indication of the different perceptions of STEM and non-STEM students emerged concerning their motivation.

First, STEM students scored lower than their non-STEM counterparts on general interest in their majors. In line with Black and Deci (2000), this lower interest might indicate that STEM students' need for autonomy is met to a lesser degree compared to non-STEM students. A consideration of the reasons for choosing a study subject supports this assumption: STEM students did not choose their subjects based on their general interest; instead, their choice was determined more by a lack of ideas about what else they should study, because their family wished for them to choose STEM more often.

Second, the analyses showed that STEM students valued their subjects to a greater degree than did non-STEM students, and third, they expected to incur more costs in order to be successful in their subjects. However, the two groups did not differ in their general expectancies of successfully graduating. It seems that STEM students are aware of the costs and efforts which they need to invest in order to succeed in STEM. However, this awareness does not prevent them from choosing a STEM subject. At first glance, this appears to be in contrast to previous suggestions that perceived barriers might discourage students from deciding for a STEM subject, insofar as students seem to be aware of the effort but nevertheless see potential gains and advantages, which can be interpreted as a rational decision (Wang and Degol, 2013). On the other hand, there might be a relation of costs and perceived values based on cognitive dissonance theory: As students expect greater costs, they need to value the subject more.

In sum, these results point in the direction that STEM students' need for autonomy could be satisfied to a lesser degree compared to non-STEM students. Nevertheless, STEM students value their subject more. This partly surprising pattern should be addressed in future studies which scrutinize the relation of autonomy and satisfaction with a choice.

Differences between STEM and non-STEM students also emerged regarding competence-related variables, but the pattern was not consistent. Competence refers to the perception of capability and effectiveness (Ryan and Deci, 2014). In this regard, non-STEM students, for instance, more frequently endorsed the statement that they chose their subjects based on their talents than did STEM students. In addition, non-STEM students showed higher self-esteem ratings than did STEM students. Both of these findings indicate that non-STEM students tend to attribute more competence to themselves than do STEM students. On the other hand, STEM students were more likely to report having selected their study subjects because they predicted that they would be able to earn a lot of money after graduation and because they wished to achieve a top management position, while non-STEM students were more likely to strive for middle management positions. This could be attributed as high competence perceptions in STEM students. Whether these mixed results allow conclusions to be drawn about the academic achievements of especially STEM students, such as the findings of Jang et al. (2009), is unclear. They do, however, reflect the notion put forward by Steele and Aronson (2005) that competence has a fragile character. It is possible that while STEM students are positively disposed toward their future and their future competencies after graduation (leadership; money), they lack confidence in their current competence (self-esteem). Furthermore, the analyses did not reveal any differences between STEM and non-STEM students in terms of attribution style, self-efficacy and academic self-efficacy. This indicates that both student groups are not particularly characterized by one of these constructs.

While the current results indicate that STEM students experience different kinds of competence, and less autonomy than do non-STEM students, we were especially interested in potential gender differences which might help to explain the lack of women in STEM subjects (Shen, 2013; Dasgupta and Stout, 2014).

There were, however, no interaction effects with regard to motivations to choose the specific subject of study. Concerning motivation and interest in study subject, on the other hand, analyses interestingly show that especially female STEM students differ from the non-STEM groups in the sense that they value their study subject more. This might indicate that those students who – against their gender role – choose a study subject are all the more determined to evaluate their choice positively. This is in conflict, however, with the finding that there are no differences in the reasons for choosing the study subject.

Regarding the expected investments needed, male STEM students, in contrast to their female counterparts, expect most costs, in terms of invested effort, to succeed in their study programs, followed by female non-STEM students and male non-STEM students. However, male STEM students scored higher regarding their expected professional opportunities after graduation compared to female STEM students, while no such difference was found between female and male non-STEM students. Male STEM students in particular seem to direct their focus toward future competence after graduation. No further interaction effects emerged. The interaction effects do not give reason to assume that female STEM students, in contrast to all other groups and particularly to their female non-STEM counterparts, can be characterized by, for instance, lower competence perceptions in terms of self-efficacy or a detrimental attribution style. However, the deeper analyses of female and male STEM students revealed several gender differences, which discriminate female STEM students at least from male STEM students.

Male STEM students were more likely to choose their study programs because they knew what they wanted to do after graduation, they saw a lot of professional opportunities after graduation and were more confident that they would be able to make a lot of money after graduation. Furthermore, compared to female STEM students, they were more likely to wish to work in top management. These results indicate more intense (future) competence perceptions compared to those of female STEM students, which are particularly directed at the future after graduation. Only one finding clearly showed that male STEM students currently perceive more competence than do female STEM students: The males showed higher self-efficacy scores, which is in line with previous findings (e.g., Zeldin et al., 2008). One construct which may contribute to the perception of competence is how people attribute their successes and failures (Ryan and Deci, 2014). For instance, attributing success to one's own ability could be beneficial, while attributing the same result to luck would be less beneficial. In line with previous research (Burgner and Hewstone, 1993), we found that women in STEM score higher on controlled orientation and impersonal orientation, but that there is no difference in autonomy orientation. In sum, this indicates that while women do not necessarily perceive less competence and autonomy, and women and men attribute the results of their behavior equally to internal sources, women are also more likely to attribute results to external sources and to have a greater feeling of powerlessness regarding the results of their behavior, as reflected in their lower values for self-efficacy. In contrast to previous findings (MacPhee et al., 2013), our analysis did not reveal any differences regarding perceived academic self-efficacy or self-esteem. Although MacPhee et al. (2013) reported a gender difference in perceived academic skills, they acknowledged that by the time of graduation, women had reached equal levels of academic self-efficacy to those of men. Given that we surveyed Master's students, it is reasonable to assume that female STEM students had already reached this point of equality.

Still, in sum there are several distinct differences between men and women regarding perceived competence and study motivation. How can this be traced back to the influence of stereotypes as discussed above? While stereotype research would not predict a STEM subject for women, it would do so all the more for men (e.g., Carli et al., 2016). Socialization in such beliefs could motivate males to choose a STEM study program, and render them confident in experiencing competence after graduation. Female STEM students, in contrast, are partly lacking the perception of competence. In line with previous stereotype research, this may also have been induced by stereotypical beliefs about women and STEM, with female STEM students assuming that their own effort, capabilities and abilities might be worth less compared to male STEM students, because STEM is a male domain (e.g., Hand et al., 2017). However, the assumption of links between stereotypical beliefs and the present findings also leads to our limitations.

Limitations

We did not capture participants' stereotypical beliefs about women and men. Such a consideration might have provided valuable insights into the degree to which participants' behavior and perceptions may have been affected by these social norms. Future studies should address this issue. Moreover, we captured participants' belonging to the binary (biological) category of sex, but used the term "gender" within this paper, due to the association with stereotypical beliefs which refer to social identity (social gender, Bem, 1993).

In addition, it would be useful to focus on social identity in order to examine whether differences in students' perceptions could be better predicted by a more female or male identity instead of a fixed binary category. It could be possible that women in STEM and men in non-STEM have slightly different gender identities, because they decided for gender-incongruent subjects. However, we found numerous differences with regard to the reasons why people selected their majors. Although we suggest that these additional ad hoc items yielded valuable results, ratings were mostly given on the lower parts of the scales, indicating that the chosen reasons were only relevant for a minority of students. Therefore, future studies should also capture additional reasons beyond classical motivations and interest items. Moreover, some of the scales we used, especially the causality orientation scale, obtained rather low internal consistency. In the case of the causality scale we have used a reduced set of the scale's original number of vignettes by Deci and Ryan (1985) to reduce the length of the survey. This lower number could be one reason for the low consistency (Cortina, 1993; Peterson, 1994). Nevertheless, using the full set of vignettes could be one way to increase the quality of the scale; another way could be to use scales which do not refer to vignettes.

In sum, results need to be considered with caution. As we have a comparatively large sample, we obtain significant differences which might not always be associated with high effect sizes. Therefore, implications for everyday life should only be derived from those results which also have high or at least medium effect sizes.

Furthermore, we gathered data of students from one university; interpretations might be exclusive to this particular group of students. To provide further evidence, future studies should be conducted with a student sample stemming from various universities.

CONCLUSION

The bigger picture shows that there are some differences between STEM and non-STEM students in terms of their competence and autonomy perceptions, which are in favor of non-STEM students. One major contribution of the paper is the finding that STEM students seem to direct their competence perceptions toward future competence, such as achieving a position in top management, seeing numerous professional opportunities or earning large amounts of money. This also holds true for the interaction of subject and gender, revealing that men in STEM score highest, followed by their female counterparts, while this future-related competence does not seem to be important to non-STEM students (either female or male). Furthermore, male and female STEM students do not differ in their autonomous behavior. Therefore, it could be expected that both women and men will achieve their academic goals and might become professionally satisfied (e.g., Ng et al., 2006). However, there are some differences in perceived competence ratings, such as self-efficacy and leadership aspirations, which is in line with previous research (e.g., Pajares and Miller, 1994; Amon, 2017). In conclusion, it seems that female STEM students are equally autonomy-oriented as male STEM students, but have lower perceptions of their own competence, especially regarding the expectation of future success.

ETHICS STATEMENT

This study was neither approved nor rejected by an ethical committee of the University due to lack of Central Ethics Commission in the University. However, the study idea and questionnaire were read, revised, and approved by different boards of the university: University board of the University of Duisburg-Essen, [Prof. Radke (rector); Prof. van Ackeren (vice-rector); Prof. Dr. Spitzley (vice-rector), Prof. Ziegler (former vice-rector), the (former) vice-rector for Diversity and Equality (Prof. Ziegler), the Equality Committee (e.g., former Prof. Krämer; former Dr. Mense, in total including members of all status groups of the University] and the University's data protection officer (Dr. Lose) who ensured anonymity of participants and data protection. We hope that the inclusion and approval of all status groups of the University serve as a sufficient approval that participants of the current survey study were treated with respect and did not face unethical conduct. We conducted the study in all conscience.

AUTHOR CONTRIBUTIONS

SS and NK contributed to conception and design of the study. SS organized the database, performed the statistical analysis, and wrote the manuscript. Both authors read, revised, and approved the submitted version.

REFERENCES

Abele, A. E., Stief, M., and Andrä, M. S. (2000). Zur ökonomischen Erfassung beruflicher selbstwirksamkeitserwartungen–neukonstruktion einer BSW-Skala. *Zeitschrift für Arbeits-und Organisationspsychologie* 44, 145–151. doi: 10.1026//0932-4089.44.3.145

Amon, M. J. (2017). Looking through the glass ceiling: a qualitative study of STEM women's career narratives. *Front. Psychol.* 8:236. doi: 10.3389/fpsyg.2017.00236

Bandura, A. (1982). Self-efficacy mechanism in human agency. *Am. Psychol.* 37, 122–147.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educ. Psychol.* 28, 117–148. doi: 10.1207/s15326985ep2802_3

Bass, B. M., and Stogdill, R. M. (1990). *Bass & Stogdill's Handbook of Leadership: Theory, Research, and Managerial Applications*, 3rd Edn. New York, NY: Simon and Schuster.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., and Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychol. Sci. Public Inter.* 4, 1–44. doi: 10.1111/1529-1006.01431

Becker, R., and Hecken, A. E. (2009). Higher education or vocational training? An empirical test of the rational action model of educational choices suggested by Breen and Goldthorpe and Esser. *Acta Sociol.* 52, 25–45. doi: 10.1177/0001699308100632

Bem, S. L. (1993). *The Lenses of Gender: Transforming the Debate on Sexual Inequality*. New Haven, CT: Yale University Press.

Black, A. E., and Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: a self-determination theory perspective. *Sci. Educ.* 84, 740–756. doi: 10.1002/1098-237X(200011)84:6<740::AID-SCE4<3.0.CO;2-3

Bong, M., and Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: how different are they really? *Educ. Psychol. Rev.* 15, 1–40. doi: 10.1023/A:1021302408382

Bortz, J., and Schuster, C. (2010). *Statistics: For Human and Social Scientists*, 7th Edn. Berlin: Springer.

Buldu, M. (2006). Young children's perceptionsof scientists: a preliminary study. *Educ. Res.* 48, 121–132. doi: 10.1080/00131880500498602

Burgess, D., and Borgida, E. (1999). Who women are, who women should be: descriptive and prescriptive gender stereotyping in sex discrimination. *Psychol. Public Policy law* 5:665. doi: 10.1037/1076-8971.5.3.665

Burgner, D., and Hewstone, M. (1993). Young children's causal attributions for success and failure:'Self-enhancing'boys and 'self-derogating'girls. *Br. J. Dev. Psychol.* 11, 125–129. doi: 10.1111/j.2044-835X.1993.tb00592.x

Carli, L. L., Alawa, L., Lee, Y., Zhao, B., and Kim, E. (2016). Stereotypes about gender and science: women≠ scientists. *Psychol. Women Q.* 40, 244–260. doi: 10.1177/0361684315622645

Chemers, M. M., Hu, L. T., and Garcia, B. F. (2001). Academic self-efficacy and first year college student performance and adjustment. *J. Educ. Psychol.* 93:55.

Collani, V. G., and Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen skala zum selbstwertgefühl von Rosenberg. *Zeitschrift für differentielle und diagnostische Psychologie* 24, 3–7.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *J. Appl. Psychol.* 78, 98–104. doi: 10.1037/0021-9010.78.1.98

Cuddy, A. J., Fiske, S. T., and Glick, P. (2008). Warmth and competence as universal dimensions of social perception: the stereotype content model and the BIAS map. *Adv. Exp. Soc. Psychol.* 40, 61–149. doi: 10.1016/S0065-2601(07)00002-0

Dasgupta, N., and Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights Behav. Brain Sci.* 1, 21–29. doi: 10.1177/2372732214549471

Deci, E. L., and Ryan, R. M. (1985). The general causality orientations scale: self-determination in personality. *J. Res. Personal.* 19, 109–134. doi: 10.1016/0092-6566(85)90023-6

Deci, E. L., and Ryan, R. M. (2000). The" what" and" why" of goal pursuits: human needs and the self-determination of behavior. *Psychol. Inquiry* 11, 227–268. doi: 10.1207/S15327965PLI1104_01

Deci, E. L., and Ryan, R. M. (2011). Self-determination theory. *Handbook Theor. Social Psychol.* 1, 416–433.

Farrell, L., and McHugh, L. (2017). Examining gender-STEM bias among STEM and non-STEM students using the implicit relational assessment procedure (IRAP). *J. Context. Behav. Sci.* 6, 80–90. doi: 10.1016/j.jcbs.2017.02.001

Feltman, R., and Elliot, A. (2014). "Competence," in *Encyclopedia of Quality of Life and Well-Being Research*, ed. A. C. Michalos (Dordrecht: Springer), doi: 10.1007/978-94-007-0753-5

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA: Sage.

Findley, M. J., and Cooper, H. M. (1983). Locus of control and academic achievement: a literature review. *J. Pers. Soc. Psychol.* 44:419. doi: 10.1037/0022-3514.44.2.419

Finson, K. D. (2002). Drawing a scientist: what we do and do not know after fifty years of drawings. *Sch. Sci. Math.* 102, 335–345. doi: 10.1111/j.1949-8594.2002.tb18217.x

Fralick, B., Kearn, J., Thompson, S., and Lyons, J. (2009). How middle schoolers draw engineers and scientists. *J. Sci. Educ. Technol.* 18, 60–73. doi: 10.1007/s10956-008-9133-3

Gayles, J. G., and Ampaw, F. (2014). The impact of college experiences on degree completion in STEM fields at four-year institutions: does gender matter? *J. High. Educ.* 85, 439–468. doi: 10.1080/00221546.2014.117 77336

Gore, P. A. Jr. (2006). Academic self-efficacy as a predictor of college outcomes: two incremental validity studies. *J. Care. Assess.* 14, 92–115.

Hand, S., Rice, L., and Greenlee, E. (2017). Exploring teachers' and students' gender role bias and students' confidence in STEM fields. *Soc. Psychol. Educ.* 20, 929–945. doi: 10.1007/s11218-017-9408-8

Jang, H., Reeve, J., Ryan, R. M., and Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? *J. Educ. Psychol.* 101, 644.

Jerusalem, M., and Schwarzer, R. (1999). *Skala zur allgemeinen Selbstwirksamkeitserwartung. Skalen zur Erfassung von Lehrer-und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Berlin: Freie Universität Berlin.

Kosovich, J. J., Hulleman, C. S., Barron, K. E., and Getty, S. (2015). A practical measure of student motivation: establishing validity evidence for the expectancy-value-cost scale in middle school. *J. Early Adolesc.* 35, 790–816. doi: 10.1177/0272431614556890

Kretschmann, J., Gronostaj, A., Schulze, A., and Vock, M. (2017). Wenn sich die masterfrage stellt: soziale herkunftseffekte auf die übergangsintention nach dem bachelorstudium. *Zeitschrift für Empirische Hochschulforschung* 1, 76–92. doi: 10.3224/zehf.v1i1.05

Lane, J., Lane, A. M., and Kyprianou, A. (2004). Self-efficacy, self-esteem and their impact on academic performance. *Soc. Behav. Personal.* 32, 247–256. doi: 10.2224/sbp.2004.32.3.247

Legate, N., and Ryan, R. M. (2014). "Individual autonomy," in *Encyclopedia of Quality of Life and Well-Being Research*, ed. A. C. Michalos (Dordrecht: Springer), 3233–3236. doi: 10.1007/978-94-007-0753-5_140

Long, M., Steinke, J., Applegate, B., Knight Lapinski, M., Johnson, M. J., and Ghosh, S. (2010). Portrayals of male and female scientists in television programs popular among middle school-age children. *Sci. Commun.* 32, 356–382. doi: 10.1177/1075547009357779

Losh, S. C., Wilke, R., and Pop, M. (2008). Some methodological issues with "Draw a Scientist Tests" among young children. *Int. J. Sci. Educ.* 30, 773–792. doi: 10.1080/09500690701250452

MacPhee, D., Farro, S., and Canetto, S. S. (2013). Academic self-efficacy and performance of underrepresented STEM majors: gender, ethnic, and social class patterns. *Anal. Soc. Issues Public Policy* 13, 347–369. doi: 10.1111/asap.12033

Marra, R. M., Rodgers, K. A., Shen, D., and Bogue, B. (2009). Women engineering students and self-efficacy: a multi-year, multi-institution study of women engineering student self-efficacy. *J. Eng. Educ.* 98, 27–38. doi: 10.1002/j.2168-9830.2009.tb01003.x

Master, A., Cheryan, S., Moscatelli, A., and Meltzoff, A. N. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *J. Exp. Child Psychol.* 160, 92–106. doi: 10.1016/j.jecp.2017.03.013

Master, A., and Meltzoff, A. N. (2016). Building bridges between psychological science and education: cultural stereotypes, STEM, and equity. *Prospects* 46, 215–234. doi: 10.1007/s11125-017-9391-z

Mavriplis, C., Heller, R., Beil, C., Dam, K., Yassinskaya, N., Shaw, M., et al. (2010). Mind the gap: women in STEM career breaks. *J. Technol. Manag. Innov.* 5, 140–151. doi: 10.4067/S0718-272420100001 00011

McCullough, L. (2011). Women's leadership in science, technology, engineering and mathematics: barriers to participation. *Forum Public Policy Online* 2011,

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., and Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychol. Bull.* 130:711. doi: 10.1037/0033-2909.130.5.711

Miller, D. I., Nolla, K. M., Eagly, A. H., and Uttal, D. H. (2018). The development of children's gender-science stereotypes: a meta-analysis of 5 decades of US Draw-A-scientist studies. *Child Dev.* 89, 1943–1955. doi: 10.1111/cdev.13039

Multon, K. D., Brown, S. D., and Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: a meta-analytic investigation. *J. Counsel. Psychol.* 38:30.

Nauta, M. M., Epperson, D. L., and Waggoner, K. M. (1999). Perceived causes of success and failure: are women's attributions related to persistence in engineering majors? *J. Res. Sci. Teach.* 36, 663–676. doi: 10.1002/(SICI)1098-2736(199908)36:6<663::AID-TEA5>3.0.CO;2-F

Ng, T. W., Sorensen, K. L., and Eby, L. T. (2006). Locus of control at work: a meta-analysis. *J. Organ. Behav.* 27, 1057–1087.

Pajares, F., and Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: a path analysis. *J. Educ. Psychol.* 86:193.

Peterson, R. A. (1994). A meta-analysis of Cronbach's coefficient alpha. *J. Consum. Res.* 21, 381–391.

Piatek-Jimenez, K., Cribbs, J., and Gill, N. (2018). College students' perceptions of gender stereotypes: making connections to the underrepresentation of women in STEM fields. *Int. J. Sci. Educ.* 40, 1432–1454. doi: 10.1080/09500693.2018.1482027

Powell, G. N., and Butterfield, D. A. (2013). Sex, gender, and aspirations to top management: who's opting out? Who's opting in? *J. Vocat. Behav.* 82, 30–36. doi: 10.1016/j.jvb.2012.11.003

Prentice, D. A., and Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: the contents of prescriptive gender stereotypes. *Psychol. Women Q.* 26, 269–281. doi: 10.1111/1471-6402.t01-1-00066

Ramseier, E. (2006). Gründe für die wahl eines studienfaches an der hochschule. *Higher Educ. Stud.* 2006:41.

Ramsey, L. R. (2017). Agentic traits are associated with success in science more than communal traits. *Personal. Indiv. Differ.* 106, 6–9.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., and Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychol. Bull.* 130:261.

Rosenberg, M., Schooler, C., Schoenbach, C., and Rosenberg, F. (1995). Global self-esteem and specific self-esteem: different concepts, different outcomes. *Am. Sociol. Rev.* 60, 141–156.

Ryan, R. M., and Deci, E. (2014). "Self-determination theory," in *Encyclopedia of Quality of Life and Well-being Research*. ed. A. C. Michalos (Dordrecht: Springer), 5755–5760. doi: 10.1007/978-94-007-0753-5_2630

Ryan, R. M., and Deci, E. L. (2002). "Overview of self-determination theory: an organismic dialectical perspective," in *Handbook of Self-Determination Research*. eds E. L. Deci and R. M. Ryan (Rochester, NY: University of Rochester Press), 3–33.

Schiefele, U., Krapp, A., Wild, K. P., and Winteler, A. (1993). Der" Fragebogen zum Studieninteresse"(FSI). *Diagnostica* 39, 335–351.

Shapiro, J. R., and Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles* 66, 175–183. doi: 10.1007/s11199-011-0051-0

Shen, H. (2013). Mind the gender gap. *Nature* 495, 22–24.

Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): an investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles* 67, 617–629. doi: 10.1007/s11199-012-0209-4

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35, 4–28. doi: 10.1006/jesp.1998.1373

Steele, C. M., and Aronson, J. (2005). "Stereotypes and the fragility of academic competence, motivation, and self-concept," in *Handbook of Competence and Motivation*, eds A. J. Elliot and C. S. Dweck (New York, NY: Guilford Publications), 436–455.

Steffens, M. C., Jelenec, P., and Noack, P. (2010). On the leaky math pipeline: comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *J. Educ. Psychol.* 102, 947–963.

Steinke, J., and Tavarez, P. M. P. (2018). Cultural representations of gender and STEM: portrayals of female STEM characters in popular films 2002-2014. *Int. J. Gen. Sci. Technol.* 9, 244–277.

Stipek, D. J., and Weisz, J. R. (1981). Perceived personal control and academic achievement. *Rev. Educ. Res.* 51, 101–137.

Wang, M. T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy–value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001

Wang, X. (2013). Why students choose STEM majors: motivation, high school learning, and postsecondary context of support. *Am. Educ. Res. J.* 50, 1081–1121. doi: 10.3102/0002831213488622

Wickware, P. (1997). Along the leaky pipeline. *Nature* 390, 202–203. doi: 10.1038/36641

Zajacova, A., Lynch, S. M., and Espenshade, T. J. (2005). Self-efficacy, stress, and academic success in college. *Res. High. Educ.* 46, 677–706. doi: 10.1007/s11162-004-4139-z

Zeldin, A. L., Britner, S. L., and Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *J. Res. Sci. Teach.* 45, 1036–1058. doi: 10.1002/tea.20195

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 27 June 2019
doi: 10.3389/fpsyg.2019.00990

Adolescents' Motivational Profiles in Mathematics and Science: Associations With Achievement Striving, Career Aspirations and Psychological Wellbeing

Helen M. G. Watt, Micaela Bucich and Liam Dacosta*

School of Education and Social Work, The University of Sydney, Sydney, NSW, Australia

OPEN ACCESS

Edited by:
*Bernhard Ertl,
Universität der Bundeswehr München, Germany*

Reviewed by:
*Linda Wirthwein,
Technische Universität Dortmund, Germany*
*Alejandro Veas,
University of Alicante, Spain*
*Heike Itzek-Greulich,
Thomas Morus High School, Germany*

***Correspondence:**
*Helen M. G. Watt
helen.watt@sydney.edu.au*

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *10 January 2019*
Accepted: *15 April 2019*
Published: *27 June 2019*

Citation:
Watt HMG, Bucich M and Dacosta L (2019) Adolescents' Motivational Profiles in Mathematics and Science: Associations With Achievement Striving, Career Aspirations and Psychological Wellbeing. Front. Psychol. 10:990. doi: 10.3389/fpsyg.2019.00990

Pertinent to concern in Australia and elsewhere regarding shortages in STEM fields, motivational expectancies and values predict STEM study and career aspirations. Less is known about how "cost" values may deter, and how expectancies/values and costs combine for different profiles of learners to predict achievement aspirations and psychological wellbeing outcomes. These were the aims of the present study using established measures of perceived talent, intrinsic and utility values, and a new multidimensional "costs" measure as the platform to explore a typology of mathematics/science learners. Grade 10 Australian adolescents (N = 1,172; 702 girls) from 9 metropolitan Sydney/Melbourne schools completed surveys early 2012/2013. Latent profile analyses educed profiles within each of mathematics and science: "Positively engaged" scored high on positive motivations, low on costs; "Struggling ambitious" were high for both positive motivations and costs; "Disengaged" exhibited generally low scores on positive motivations but high costs. MANOVAs examined mathematics/science profile differences on clustering variables, experienced learning environments, achievement background and striving, career aspirations and psychological wellbeing. Positively engaged/Struggling ambitious were distinguished by high costs perceived by Struggling ambitious, associated with debilitated psychological wellbeing, but not eroding achievement striving. A greater proportion of boys was in this risk type. Disengaged students reported lowest STEM-related career aspirations, aimed marks and history of results; in mathematics, a greater proportion of girls was in this risk type. Profiles could be conceptualized along dimensions of achievement striving and psychological wellbeing. Similar profiles for mathematics and science, and coherent patterns of antecedents and outcomes, suggest several theoretical and educational implications.

Keywords: expectancy-value, profiles, mathematics, science, motivations, costs, aspirations, psychological wellbeing

INTRODUCTION

There is concern, in Australia and elsewhere, regarding shortages in "STEM" fields (science, technology, engineering and mathematics). STEM skills contributed 65% toward Australia's economic growth between 1964 and 2005 and are required by three-quarters of the fastest growing occupations (Office of the Chief Scientist [OCS], 2014a). Global shortages of STEM capable workforces are predicted to worsen and impact economic and social development (Organisation for Economic Co-operation and Development [OECD], 2006; Office of the Chief Scientist [OCS], 2014b). Over the past two decades, STEM participation has declined in many western nations, including Australia, where there has been a concerning trend away from high school advanced mathematics, physics and chemistry (Federation of Australian Scientific and Technological Societies [FASTS], 2002; Ainley et al., 2008; Dobson, 2012). Women are less likely to choose STEM careers, and more likely to leave once entered (NSCRC, 2013; Cheryan et al., 2017). Adolescence is the crucial time when most students make choices whether to concentrate on STEM in the future (Maltese and Tai, 2011), when course selections can foreclose future educational and career pathways (Watt, 2006). In fact, much of the disparity in both gender and ethnic representation in STEM fields at university could be accounted for by differences in secondary school course-taking (Riegle-Crumb and King, 2010). Although convenient to refer collectively to "STEM", this can mask different patterns between, as well as within those fields (e.g., physics, chemistry, and biology within science). Our study examines each of mathematics and science as core curricula studied until the end of grade 10 in Australia, after which students choose their courses for upper secondary school.

Theoretical Framework

Eccles et al.'s (1983) expectancy-value theory (EVT; Eccles, 2005) offers a comprehensive framework to explain achievement-related choices. Initially developed to explain gender differences in high school mathematics enrolments, EVT has become a foremost motivation framework to understand how youths' beliefs predict educational choices (Jacobs and Simpkins, 2005). At its core, the model posits that expectancy-related beliefs interact with different kinds of values, to predict achievement behaviours and choices. Expectancies and values are contextualised in a developmental framework drawing on decision, achievement goal, attribution and self-worth theories to provide an integrated framework accounting for origins stemming from childhood. The components of task value include intrinsic value or interest, attainment value (referring to the personal importance of doing well), utility value, and cost. Most studies have focused on the first three positive values, at times combining attainment and utility values into an "importance value"; some researchers have measured additional value factors (e.g., Gaspard et al., 2015 distinguished 11 task values). A wealth of studies has collectively established that expectancy-related beliefs (e.g., perceived competence, perceived talent, self-concept) and intrinsic/utility/importance values predict achievement-related choices, including enrolments and career aspirations (for reviews see Watt, 2010, 2016; Wigfield and Cambria, 2010).

The bulk of literature has been in relation to mathematics because it was identified as a "critical filter" (Sells, 1980) limiting girls' and women's access to certain high-status and high-income fields of education and career. Gendered mathematics values and ability beliefs predict advanced participation over and above achievement background, for mathematical enrolments (Updegraff et al., 1996; Simpkins et al., 2012; Watt et al., 2012) and career aspirations (Watt et al., 2012), as well as science enrolment intentions (Ethington, 1991; Atwater et al., 1995), scientific career aspirations (Watt et al., 2017a,b), and pursuit of a science career in adulthood (Farmer et al., 1999). Science expectancies and values (including costs) predicted planned completion of a STEM major among college students (Perez et al., 2014).

Researchers within this framework have recently targeted the negative cost value for empirical attention. From the outset (Eccles et al., 1983; Eccles, 1987), Eccles described cost in terms of all the negative aspects of engaging in a task including emotional states, such as anxiety, and the amount of effort required in order to succeed, which can impact opportunities to pursue other valued activities (for a review see Wigfield and Eccles, 1992). She posited that positive values (such as interest and enjoyment) should interact with perceived costs to impact task choices, in increasingly complex ways as the choices between different activities become more numerous with age. Wigfield and Eccles (1992) emphasised the importance of examining how values *work together* in influencing students' task choices and highlighted the conflict that may occur when values are not in synchrony. Although studies have not yet investigated these factors altogether, task-specific asynchronous profiles, in which positive values but also costs are high, would seem likely to exert deleterious effects on general psychological wellbeing.

Until recently, there had been little empirical attention to the measurement of costs within EVT. First studies included the single facet of opportunity cost (Conley, 2012; Trautwein et al., 2012), single factors which mixed facets together (Luttrell et al., 2010; Jiang et al., 2018), a combination of single and mixed factors (opportunity cost versus other mixed costs, Gaspard et al., 2017), or were unable to discern theorised multiple dimensions resulting in an omnibus cost factor (Battle and Wigfield, 2003). In each of those studies, cost was factorially distinct from and inversely correlated with positive task values. Cost negatively predicted planned (Battle and Wigfield, 2003) or actual (Luttrell et al., 2010) enrolments; and explained variance additional to self-efficacy and an aggregate task value factor on disorganisation, procrastination and avoidance intentions, even when controlling for pre-test construct scores (Jiang et al., 2018).

Battle and Wigfield's (2003) study was conducted among female college students regarding intentions to attend graduate school, rather than among school students in mathematics as were the other studies. Building on their groundwork, Perez et al. (2014) adapted items designed to tap each of effort cost, psychological cost, and opportunity cost among college students in STEM, and were able to confirm their theorised structure. Each cost factor positively predicted intentions to quit a STEM major, most strongly for effort cost. It is this scale that forms the basis for the multidimensional cost measure developed for the present study, adapted from the college context to suit adolescents. A similar goal was expressed by Flake et al. (2015), despite their studies being conducted among college students, who proposed an additional fourth "outside effort" cost (example item: "I have so many other responsibilities that I am unable to put in the effort that is necessary for this class"). However, their factors showed extreme correlations: from 0.83 between outside effort and emotional costs, up to 0.95 between effort and opportunity costs (that they termed "loss of valued alternatives"), with a median correlation of 0.87. Costs inversely correlated with general expectations for success, an omnibus positive value factor, items tapping long-term interest, a single "overall motivation" item, and a final course grade in the college-level calculus course. A measure sensitive to the nuances of different costs is required to allow empirical examination of the tenets of expectancy-value theory with regard to how all values, including costs, work together and with expectancies, to influence choices before students already self-select into their chosen fields.

Motivation Profiles in the Expectancy-Value Framework

Although expectancies and values are posited to interactively predict choices, many studies have adopted main-effects variable-centred models. Little is known about how expectancies and positive/negative values combine for different profiles of learners, with implications for achievement striving, career aspirations or psychological wellbeing. Latent profile analyses allow the exploration of types of individuals having distinct motivation profiles, to link them with potential antecedents and outcomes. Such person-centred analyses focus on the individual as the unit of analysis, consistent with modern developmental theory (Magnussen, 2000). This approach was adopted in the present study based on the premise that different configurations of expectancies and values (including costs) should impact students' achievement striving, career aspirations, and potentially psychological wellbeing outcomes.

A small but growing literature has begun to adopt a typological approach informed by EVT constructs. These have either been concentrated in STEM domains (Conley, 2012; Lazarides et al., 2016a, 2018, 2019; Andersen and Chen, 2016; Chittum and Jones, 2017; Linnenbrink-Garcia et al., 2018; Perez et al., 2019), assessed at a domain-general level (Roeser et al., 1999; Tuominen-Soini et al., 2011, 2012), or across multiple domains (Viljaranta et al., 2009; Chow and Salmela-Aro, 2011; Chow et al., 2012; Lazarides et al., 2016b; Guo et al., 2018). A number of studies exist which compare motivations across domains using aggregate task values (combining component positive values into a single score) but exclude expectancy or cost measures (Viljaranta et al., 2009; Chow and Salmela-Aro, 2011; Chow et al., 2012; Guo et al., 2018); while others include aggregate values plus self-concept (Lazarides et al., 2016b). Emergent profiles from this body of work characterise relative valuing across subjects (e.g., mathematics and physical science versus English; Chow et al., 2012). While illuminating in their own right, our aim is to identify groups of students *within* each domain of mathematics and science, including the breadth of expectancy and values constructs. We review studies in this section which focused on STEM-specific motivational profiles informed by the EVT framework.

There has been a line of research focused on EVT motivational profiles in mathematics (Lazarides et al., 2016a, 2018, 2019), although not all of them have included multiple value factors and none included costs. Lazarides et al. (2018) found two profiles among Finnish first- and second-graders who showed consistently low or high levels of interest value, self-concept and performance, as well as a group with medium interest despite low self-concept and performance. A fourth group was characterised by low interest, yet medium levels of self-concept and performance. These profiles supported the authors' expectations that there would be mixed as well as consistent configurations of value and expectancy, although only intrinsic value was included. Lazarides et al. (2016a,b) identified four profiles among German adolescents based on a wider range of values (intrinsic, utility, and attainment) and self-concept: three showed consistent configurations (low, moderate, and high), and a fourth mixed group (containing more girls) showed moderate to high utility value but lower scores on other dimensions. Most recently, using a large sample (N = 6,020) of German 9th–10th graders from upper- and middle-track schools (PISA-I Plus study in 2003–2004; Prenzel et al., 2006). Lazarides et al. (2019) identified similar profiles: low, medium, high, and an infrequent mixed type (4% of sample) characterised this time by high self-concept, low interest and pronounced utility value. Similar profiles were identified for science motivation (Andersen and Chen, 2016; Chittum and Jones, 2017; who added teacher caring as a clustering variable). It seems that utility value can be high, independent of other kinds of internal values. Including negatively valenced constructs (i.e., costs) may yield more nuanced, asynchronous profiles.

Only three studies included cost (Conley, 2012; Bøe and Henriksen, 2013; Perez et al., 2019) among clustering variables in examining students' motivational profiles. Using selected variables from EVT and achievement goal theory, Conley (2012) identified 7 clusters of 7th-grade United States students who could be grouped at "low", "average", and "high" levels of mathematics motivation. Within these broad groups, clusters were further distinguished by perceptions of cost (only "opportunity cost" was measured). In the 3 clusters with "average" motivation, only the cluster with higher perceived cost differed on achievement and affective outcomes, showing lower levels of achievement, more negative and less positive affect in mathematics class. This highlights the value of including even just one type of cost in distinguishing types of students. Although increased cost did not lead to more negative affect in mathematics class when comparing two clusters with similarly high levels of task values, including multiple dimensions of cost may provide a more complete picture and better prediction of outcomes by tapping the potential strains among highly motivated students. Similarly to Conley (2012), Linnenbrink-Garcia et al. (2018) included EVT as well as achievement goal theory constructs. Unlike Conley (2012), task value (interest, attainment, and utility) was aggregated, which meant clusters were distinguished more by their achievement goals than patterns of values, and costs were not included.

In a sample of undergraduate students enrolled in gateway chemistry courses at an elite United States university, Perez et al. (2019) identified three profiles based on combinations of science competence beliefs (expectancies), values (attainment, utility, and interest) and two kinds of costs (opportunity and effort; validating their previously developed scale also in this independent sample; Perez et al., 2014). One profile characterised by lowest competence beliefs and values with highest costs was labelled "Moderate All" due to their moderate scores on all variables. Moderate scores on positive motivational variables may be expected for even the least motivated, among students at an elite university who already self-selected into a science major. The other two profiles exhibited higher values and perceived competence, with lower costs. These two profiles differed only in their level of values and effort cost as reflected in their names; "Very High Competence/Values-Low Effort Cost" and "High Competence/Values-Moderate Low Costs." Interestingly, despite equivalent STEM GPA at the end of their 1st and 4th years, students in the "High Competence/Values-Moderate Low Costs" profile completed fewer STEM courses by the end of both years than those in the "Very High Competence/Values-Low Effort Cost" profile. This highlighted the differentiating role of values, competence beliefs and effort cost (but not opportunity cost) in determining STEM participation among these two more motivated profiles. The "Moderate All" profile fared worst in terms of both STEM GPA and course completion at both timepoints, and, women and underrepresented minorities were more likely to fall into that profile. Perez et al. (2019, p. 19) acknowledged the limitation of only assessing two dimensions of cost and recommended inclusion of psychological cost in future research. We would also expect more variation among students not already self-selected into tertiary STEM studies.

Only one other study explicitly measured cost in terms of the negative aspects related to one educational choice compared with others (Bøe and Henriksen, 2013). While termed "relative cost", it seemed to tap cost in general, rather than a specific dimension (Perez et al., 2019). Students of physics were asked to retrospectively rate the importance in their choice to study physics, alongside expectancies, interest, attainment and utility values (Bøe and Henriksen, 2013). The three resultant profiles differed on positive motivation variables, but all showed similar low levels of cost, quite possibly because only students who had self-selected into physics (potentially in part due to low perceived cost) were sampled. Our study explored expectancies and values including costs at a timepoint preceding students' choice of STEM enrolments to overcome this limitation.

Akin to psychological cost, test anxiety was measured along with constructs closely related to EVT such as self-efficacy, competence and task value, among secondary school mathematics and science students in Singapore (Ng et al., 2016). Four clusters were educed: "low" (low on motivational beliefs and anxiety), "high" (high motivational beliefs and anxiety), "good" (high motivational beliefs and low anxiety), and "poor" (lowest motivational beliefs but high anxiety). Only achievement correlates were examined, precluding any directional inferences: low and high groups had moderate academic achievement in those subjects, "good" had the highest, and "poor" had the worst. It is interesting that the asynchronous group ("high") only had moderate achievement despite high motivation, conceivably related to their high anxiety.

There have been growing efforts to integrate wellbeing and achievement motivation-related constructs altogether when identifying typologies of students. An early study combining motivational and wellbeing variables was conducted by Roeser et al. (1999) who identified profiles of grade 8 students in the United States based on their perceived academic competence and task value (each averaged across mathematics and English to approximate "general" school beliefs; value as an aggregate of component values within EVT) and mental health. Analyses yielded four profiles: "well-adjusted" (positive on all three indicators), "multiple problems" (low on all), "poor motivation" (positive mental health, but low competence beliefs and value), and "poor mental health" (poor mental health, despite high competence beliefs and value). Those authors called for further studies at a domain-specific level, encompassing multiple dimensions of psychological wellbeing.

A recent study also aimed at integrating academic motivation and psychological wellbeing dimensions but did so among 15–16 years old Finnish students (Parhiala et al., 2018), based on four motivation factors (aggregate task value for each of mathematics and literacy, school enjoyment, task-focused behaviour) and four wellbeing factors (school burnout, self-esteem, externalising and internalising behaviours). Five identified profiles included three synchronous (high/average/low) as well as two asynchronous types (low motivations but average wellbeing, average motivation but low wellbeing), highlighting the fact that positive motivations need not accompany psychological wellbeing. Those authors retained domain-specific values (aggregated across interest, importance, and usefulness) for mathematics/literacy, but other motivation factors at the domain-general level (about schoolwork in general, which might not equally apply in all learning domains) and did not tap expectancies or costs. Our review could identify no previous study which taps the breadth of EVT constructs including values, expectancies and different types of cost among youth not yet self-selected into STEM studies.

Profiles Based in Other Motivational Theories

Four types of students were initially theorised and identified by Covington and Omelich (1991) in their seminal quadripolar model of need achievement, distinguished by their degree of success orientation and failure avoidance along two orthogonal continua. Although aligned with achievement goal rather than expectancy-value theory, their seminal work on motivational types is relevant to those we might expect in our study as the stresses tapped by our measure of psychological cost may function similarly to failure avoidance. Their first two groups were the classic success-oriented and failure-avoiding students: "Optimists" were high on success orientation and low on failure avoidance (cf. performance-approach oriented students: Nicholls, 1984; Elliot and Harackiewicz, 1996); "Self-protectors" were high on failure avoidance but low on success orientation (cf. performance-avoidant students: Nicholls, 1984; Elliot and Harackiewicz, 1996). The two other groups were "Failure acceptors" (low on both dimensions) and "Overstrivers" (high on both). Failure acceptors were indifferent to school achievement, having given up their efforts to avoid the implications of failure (Covington and Omelich, 1985). Overstrivers reflected an intense desire both to succeed and avoid failure. They perceived themselves as capable but feared that they may not be as worthy as their achievements indicated, exhibiting a hybrid quality of hope and fear (Covington and Omelich, 1991). Although a successful strategy for achievement in the short-term, in the long run Covington proposed their success would become an "intolerable burden" (1992, p. 89) because nobody can live up to perfection and avoid failure forever. Although anxiety may arouse overstrivers' abilities and efforts, Covington (1992) highlighted the risk to their wellbeing as a core issue that should be of concern to researchers. It is these 'negative' motivations that students can associate with studying mathematics and science, and a range of achievement-related as well as psychological wellbeing outcomes that the current study adds to previous literature using a typological approach. Regardless of the theoretical approach taken, it appears that when negatively valenced constructs are included, an asynchronous type emerges, with resemblance to Covington and Omelich's (1991) originally proposed overstrivers in terms of maladaptive outcomes. In this section we review studies that used a typological approach outside of EVT and included both positively and negatively valenced motivational constructs.

Achievement Goal Theory

Originally, achievement goal theory stipulated a dichotomy of goal orientations: mastery/task (striving for competence) or performance/ego (to demonstrate competence relative to others; Dweck, 1986; Ames, 1992). Performance goals were later divided into performance-avoidance (to avoid showing lack of competence) and performance-approach (to demonstrate relative competence) with evidence that only performance-avoidance had clear detrimental effects (Barron and Harackiewicz, 2001), and performance-approach may be adaptively paired with mastery goals within the "multiple goals framework" (e.g., Pintrich, 2000). There have been a number of person-oriented studies using achievement goal theory in domain-general, rather than STEM-specific ways (e.g., Tuominen-Soini et al., 2011, 2012). Among Finnish secondary school students, four achievement goal orientation profiles have been consistently found: (a) indifferent: those with scores close to the mean on all achievement goal orientations, therefore not displaying a tendency toward any particular orientation; (b) success-oriented: those with high levels of mastery and performance-related orientations; (c) mastery-oriented: having high mastery orientation and relatively low scores on all other orientations; and (d) avoidance-oriented: students low on mastery orientation who aim to minimise effort. In both those studies, aggregated single value measures at a domain-general level (i.e., school value) were included as a criterion. Mastery-oriented students had highest school value, followed by success-oriented students, with the other groups (avoidance-oriented and indifferent) lowest. Mastery-oriented students reported lowest feelings of inadequacy (other types did not differ from each other). Mastery- and success-oriented students had higher academic achievement; success-oriented and indifferent students had higher fear of failure (Tuominen-Soini et al., 2011). Success-oriented and indifferent students suffered greater exhaustion than other types; mastery- and avoidance-oriented students did not differ on exhaustion (Tuominen-Soini et al., 2012). Higher psychological wellbeing among the mastery- than success-oriented students, points to the negative effects of concerns regarding one's competence, as theorised by Covington (1992). The success-oriented group resembled a profile identified in an early study by Roeser et al. (2002) with high motivation and achievement, and at the same time, emotional distress. While we do not measure students' personal goal orientations in the current study, we do measure the perceived goal orientations of their learning environments (to what extent the teacher promotes a mastery or performance-oriented classroom).

Burnout and Engagement

There is a line of student burnout and engagement literature that focuses on the adverse effects of school-related stress and anxiety. A number of studies have identified profiles of students combining school engagement and burnout at a domain-general level (Tuominen-Soini and Salmela-Aro, 2014; Salmela-Aro and Read, 2017). The latter study found four groups; "engaged", "engaged-exhausted", "burned-out", and "cynical" (disengaged but lower on burnout dimensions than "burned-out"). Engaged students reported the highest school value, GPA and psychological wellbeing. Engaged-exhausted had the second highest level of school value and GPA but more depressive symptoms, lower self-esteem and greater preoccupation with possible failure than engaged and cynical students. Cynical and burned-out students scored lowest on both school value and GPA; their distinguishing feature was their general psychological wellbeing – burned-out students suffered the most depressive symptoms and lowest self-esteem; cynical students had higher psychological wellbeing than burned-out and engaged-exhausted students. Three profiles of Finnish lower-secondary school students were also identified based on burnout together with academic-general values (beliefs about the importance of school) and self-regulation variables (effort and preparation; Virtanen et al., 2018): "high-engagement/low-burnout" (a positive type), "low-engagement/high-burnout" (a negative type), and "average-engagement/average-burnout" (an asynchronous type). The burnout and engagement literature shows, at a domain-general level, that combining positively and negatively valenced constructs results in synchronous as well as asynchronous profiles, which link to academic outcomes. However, researchers are yet to examine such links in a STEM-specific way.

Self-Regulation/Coping and Psychological Wellbeing

A combined focus on motivation, self-regulation and psychological wellbeing identified four profiles of Australian undergraduate students based on a combination of adaptive and maladaptive motivational constructs (self-belief, valuing of education, learning focus and failure avoidance), psychological wellbeing (anxiety) and self-regulation variables (task management, persistence and planning; Elphinstone and Tinker, 2017). As well as two groups characterised by high/low engagement and study skills, respectively, they identified two groups with moderate engagement – one accompanied by high maladaptive constructs (anxiety, failure avoidance, and uncertain control) and the other low. Thus, including negatively valenced constructs was important in distinguishing students having otherwise similar levels of motivation and engagement.

In an influential large-scale German study among health professionals (Schaarschmidt and Fischer, 1997), four types (clustered on 11 self-reported coping dimensions) were linked to achievement striving and psychological wellbeing. "Good psychological health" and "Sparing" exhibited positive psychological health, but differed in their professional commitment and work efforts; "Excessively ambitious" and "Burnout" were both risk types for poor psychological health, but the former showed excessive commitment to work (resembling overstrivers, although grounded in different measures and theory), whereas the "Burnout" were exhausted with reduced commitment to work. There appear to be similarities in patterns of profiles across different theoretical approaches and contexts (both country settings and level of schooling/workforce). It will be important to see how STEM-specific expectancies and values combine and relate to achievement striving, career aspirations, psychological wellbeing, gender, prior achievement, and experienced learning environments, among a sample of youth not yet self-selected into their STEM studies.

Potential Outcomes: Achievement Striving and Career Aspirations

Achievement striving, career aspirations and dimensions of psychological wellbeing were expected to relate to expectancy-value motivational profiles. Some of the reviewed studies linked motivational profiles to career aspirations (Viljaranta et al., 2009; Chow et al., 2012; Chittum and Jones, 2017; Guo et al., 2018; Linnenbrink-Garcia et al., 2018) or achievement striving (Chittum and Jones, 2017; Ng, 2018) as criterion variables. Key findings, from studies which examined profiles within STEM domains together with STEM-related outcomes, linked profiles to aimed marks in a study framed by achievement goal theory (Ng, 2018), efforts exerted and STEM-related career aspirations in a study located within the expectancy-value framework (Chittum and Jones, 2017), and STEM career aspirations in a study drawing on both theories (Linnenbrink-Garcia et al., 2018).

In brief, "avoidant" Hong Kong secondary school students (high performance-avoidance, low mastery, and performance-approach) aimed for lowest grades, followed by "performance-anxious" (high performance-approach and avoidance); "all-goal" and "motivated" (relatively high mastery, performance-approach, and pro-social goals) students demonstrated similar high grade aspirations (Ng, 2018). United States fifth to seventh graders in a high interest/expectancy/utility value profile reported highest efforts and career aspirations in science, whereas a low interest/utility value profile reported the lowest, despite high success expectancies (Chittum and Jones, 2017). A United States college student "high intrinsic motivation and confidence" profile identified highest intentions for a science research career, while an "average" profile had the lowest intentions (Linnenbrink-Garcia et al., 2018).

Wellbeing-related factors were included as criterion variables in some of the already reviewed motivation profile studies (e.g., negative affect in mathematics class, Conley, 2012; exhaustion and feelings of inadequacy, Tuominen-Soini et al., 2012). There have been growing efforts to integrate psychological wellbeing and achievement motivation-related constructs, although not yet in a STEM-specific way.

Potential Antecedents

We included the role of gender, prior achievement background and experienced learning environments as key potential antecedents.

Achievement Background and Gender

Despite equivalent abilities, a large literature has documented higher mathematics ability-related beliefs for boys versus girls (Eccles et al., 1983, 1984; Stevenson and Newman, 1986; Else-Quest et al., 2010), and girls' lower beliefs in their capabilities for mathematical activities than boys' are well-established with an effect size of $d = 0.16$ in Hyde's (2005) meta-analysis. On the "values" side of EVT, adolescent girls and boys report similar beliefs about the utility/importance value of mathematics in Australian (Watt, 2004), United States and Canadian samples (Watt et al., 2012). However, boys consistently report higher interest in mathematics than girls (Updegraff et al., 1996; Watt, 2004; Frenzel et al., 2010), including in the PISA (Programme of International Student Assessment; Organisation for Economic Co-operation and Development [OECD], 2004) results, which showed higher mathematics interest and enjoyment for 15-year old boys than girls across all 41 participating countries. Less attention has been given to gender differences in *negative* motivational factors. An exception is mathematics anxiety, with an effect size showing worse anxiety for girls, $d = -0.15$ (Hyde, 2005), having obvious relevance to psychological or emotional cost. On this basis, more girls are expected to be in mathematics types showing low perceived talent and interest, as well as high psychological cost, but no hypothesis is advanced regarding gender composition of science profiles where there has not been the same volume of systematic study of gender differences, and where the domain ('science') encompasses a range of disciplines that may mask nuanced effects.

Learning Environments

Achievement goal theory (AGT) offers a framework within which to analyse students' motivational learning environments as potential antecedents to motivational profiles. Mastery-oriented classrooms (focused on learning and understanding) have generally been found to predict higher levels of interest (Midgley et al., 2001; Harackiewicz et al., 2008; Carmichael et al., 2017), whereas performance-oriented classrooms (focused on competition and grades) tend to be unassociated (Harackiewicz et al., 2008; Pantziara and Philippou, 2015; Carmichael et al., 2017) or even undermine learning (Crouzevialle and Butera, 2013). Mastery and performance-oriented learning environments have been linked to students' mathematical task values and career aspirations (Lazarides and Watt, 2015). There is large variation in students' perceptions of the same classroom environment (Wolters, 2004; Spearman and Watt, 2013), pointing to the importance of factors which filter and frame students' interpretations of their learning experiences, such as gender and motivational type, possibly fuelled by stereotypes that girls are not as naturally gifted at mathematics as boys. For example, Covington (1984) anticipated competitive learning environments to exacerbate stressors for overstrivers. We expected that experienced mastery environments would predominate among positive motivational types, and performance-oriented environments among a hybrid type characterised by high positive and negative cost motivations.

The Present Study

Our core aims were, first, to validate the new adolescent multidimensional cost measure in the context of the set of expectancy-value constructs within each of mathematics and science, and present a rich nomological network of associations with demographics, achievement background, experienced learning environments, achievement striving, STEM-related career aspirations and dimensions of psychological wellbeing. Our second aim was to discern theoretically coherent profiles, explicitly grounded in the Eccles et al. (1983; Eccles, 2005, 2009) expectancy-value model, to link with potential antecedents and outcomes, and to assess the degree of domain specificity versus generality.

Using established measures of perceived talent, intrinsic and utility values, and a new multidimensional measure for seldom-researched "costs", latent profile analysis (LPA) educed expectancy-value profiles among grade 10 students in each of mathematics and science. This person-centred approach allowed consideration of how motivational dimensions combined among different types of students to offer a nuanced understanding of the features and dynamics associated with particular profiles (Lawson and Lawson, 2013). Gender, experienced classroom learning environment and achievement influences were compared; and potential consequences for achievement striving, STEM-related career plans, and psychological wellbeing. It was hypothesised that profiles would be related but also distinct across mathematics/science, differentiated along positive/negative dimensions of achievement striving and psychological wellbeing.

Based on the preceding review, we advanced the following hypotheses:

(1) We expected to identify a positive profile (high perceived talent, intrinsic and utility values, low costs) with adaptive outcomes.
(2) We hypothesised that we would find an asynchronous type (high on all) with detrimental consequences for psychological wellbeing, although not for achievement striving or career aspirations.
(3) Finally, because the importance of mathematics and science are especially emphasised during school, we speculated that it may be unlikely that we identify a group low on all factors. Indeed, the studies by Lazarides et al. (2016a,b) detected a group for whom mathematics utility value was pronounced despite their low intrinsic value and self-concept. With cost factors simultaneously included, we speculated that the inconsistency between high utility value versus own low interests and self-beliefs may coincide with strain manifested in costs, leading to a mixed profile (high utility value, low perceived talent and intrinsic value, and potentially elevated costs) with lower achievement striving and STEM career aspirations, but less effect on psychological wellbeing.
(4) Experienced learning environments were expected to relate to motivational types, with performance climate more strongly experienced by an asynchronous type, mastery by the positive type, and no hypothesis was advanced for the third speculated type.

MATERIALS AND METHODS

Sample

Data were from the Study of Transitions and Education Pathways[1] (STEPS; Watt, 2004). Participants were grade 10 students (N = 1,172; 702 girls and 470 boys), before students have the option to opt out of mathematics and/or science for their final grades 11 and 12 of secondary schooling, recruited from 9 metropolitan/suburban schools in 2012/2013 from Melbourne and Sydney, Australia. Response rates ranged from 36 to 96% across schools (Mdn = 78%). Three schools were academically selective (i.e., students pass an achievement test to be able to enrol). School ICSEA (Index of Community Socio-Educational Advantage) scores ranged from 957 to 1,187 (Mdn = 1,128). Higher ICSEA scores indicate a higher level of educational advantage of students who attend that school, set at a national average value of 1,000. Students reported their parents' highest level of education from 1 (did not complete high school), 2 (completed high school), 3 (completed TAFE training [Technical And Further Education colleges in Australia]), 4 (completed a university degree). **Table 1** depicts the sample size for each cohort, percentage response rate, proportion of girls, school ICSEA and selective schools. Mean age was 15.79 in years (SD = 0.50). The sample was predominantly of English-speaking background (n = 754), with Chinese as the next frequent home language (n = 121), followed by Vietnamese (n = 92), Sinhalese (n = 23), and Korean (n = 18); other home language frequencies were 10 or fewer. Home language background was coded 1 (English) or 0 (other language).

Procedure

Principals of participating schools were sent an invitation letter by the first author outlining the study design and rationale,

[1] www.stepsstudy.org

TABLE 1 | Sample response rates and distribution per year by school, cohort and gender (N = 1,172).

	2012					2013					
School	1[a]	2	3	4	5	4	5	6[a]	7[a]	8	9
Cohort	1	2	3	4	5	6	7	8	9	10	11
N	50	74	186	54	74	70	78	178	182	159	67
% response	36	81	76	51	96	78	88	77	91	56	83
% girls	100	45	47	100	0	100	0	100	40	45	100
ICSEA[b]	1187	1157	957	1051	1078	1051	1078	1128	1123	1184	1163

[a]Selective schools. [b]ICSEA (Index of Community Socio-Educational Advantage) is a scale set at an average of 1,000 which allows for comparisons among schools with similar students. The higher the ICSEA, the higher the level of educational advantage of students who go to that school; the lower the ICSEA, the lower the level of students' educational advantage. For comparability the 2012 scores were used for all schools.

TABLE 2 | Expectancy-value constructs and items.

Construct (α) (math/science)	Item
Perceived talent (0.91/0.93): 1 = not at all, 7 = extremely	
	Compared with other students in your class, how talented do you consider yourself to be at maths/science?
	Compared with other students in your Year, how talented do you consider yourself to be at maths/science?
	Compared with your friends, how talented do you consider yourself to be at maths/science?
	Compared with other subjects at school, how talented do you consider yourself to be at maths/science?
Intrinsic value (0.95/0.94): 1 = not at all, 7 = extremely	
	How much do you like maths/science compared with your other subjects at school?
	How interesting do you find maths/science?
	How enjoyable do you find maths/science?
Utility value (0.89/0.89): 1 = not at all, 7 = extremely	
	How useful do you believe maths/science is?
	How useful do you think maths/science is in the everyday world?
	How useful do you think mathematical/scientific skills are in the workplace?
Cost-effort (0.77/0.86): 1 = strongly disagree, 7 = strongly agree	
	When I think about the hard work needed to get through in maths/science, I am not sure that it is going to be worth it in the end.
	Considering what I want to do with my life, studying maths/science is just not worth the effort.
	Achieving in maths/science sounds like it really requires more effort than I'm willing to put into it.
Cost-psychological (0.85/0.86): 1 = strongly disagree, 7 = strongly agree	
	It frightens me that maths/science courses are harder that other courses.
	I'm concerned that I won't be able to handle the stress that goes along with studying maths/science.
Cost-social (0.81/0.89): 1 = strongly disagree, 7 = strongly agree	
	I'm concerned that working hard in maths/science classes might mean I lose some of my close friends.
	I worry about losing some valuable friendships if I'm studying maths/science and my friends are not.

including required university, school system and departmental ethical approvals, followed up with a telephone call one week later. Meetings with each principal and selected senior school staff followed, following which the included nine schools agreed they were willing to facilitate the study. Invitation letters, explanatory statements and participation consent forms were distributed by school year coordinators to students, who were requested to return them to indicate their own and parents' consent to participate or otherwise. Surveys were administered by the first author and trained research assistants to participating students during negotiated class time in approximately the third month of the school year and took approximately 30 minutes to complete.

Measures

The second Study of Transitions and Education Pathways was designed to tap a range of students' motivational constructs and aspirations relevant to mathematics and science, building on previous work in the expectancy-value model (Watt, 2004), achievement background and aspirations, learning environments and psychological wellbeing.

Expectancies and Values

Motivational constructs were assessed in relation to each of mathematics and science using Eccles and colleagues' expectancy-value measures (Eccles and Wigfield, 1995), with contextualising modifications for the Australian sample. **Table 2** presents a complete items list and acceptable Cronbach's alpha measures of internal consistency. Students' perceptions of talent, intrinsic and utility task values were assessed by items adapted to the Australian setting (see Watt, 2002, 2004) on 7-point Likert scales ranging from 1 (not at all) to 7 (extremely). Perceived talent was tapped by four items, e.g., "Compared with other students in your class, how talented do you consider yourself to be at maths[/science]?". Intrinsic value was measured by three items, e.g., "How much do you like maths[/science], compared with your other subjects at school?". Utility value was measured by three items, e.g., "How useful do you believe maths[/science] is?"

A new multidimensional costs measure was developed for adolescents, adapted from Perez et al.'s (2014) college-level scale, tapping Effort cost (3 items, e.g., "Achieving in maths[/science] sounds like it really requires more effort than I'm willing to put into it"), Psychological cost (2 items, e.g., "I'm concerned that I won't be able to handle the stress that goes along with studying maths[/science]") and Social cost (2 items, e.g., "I'm concerned that working hard in maths[/science] classes might mean I lose some of my close friends"). The Social cost factor was renamed from "opportunity cost" referred to by Perez et al. (2014) to better reflect these items' content. All costs were rated from 1 = strongly disagree, to 7 = strongly agree. Validation of this new measure was established as part of the outlined confirmatory factor analyses that follow.

Achievement Background

Achievement background in mathematics was assessed by students' performance on the grade 9 national numeracy examination 'NAPLAN' (National Assessment Program –

Literacy and Numeracy) from Band 5 (lowest) to Band 10 (highest). In science, for which no common assessment is undertaken, students self-reported their 'usual' grade 9 score selected from 0 (<50%), 1 (50–54%), 2 (55–59%), 3 (60–64%), 4 (65–69%), 5 (70–74%), 6 (75–79%), 7 (80–84%), 8 (85–89%), 9 (90–94%), 10 (95–100%).

Perceived Classroom Learning Environment

Students' perceived mathematics/science classroom learning environments were assessed for performance-approach (e.g., "Our maths[/science] teacher points out those students who get good grades as an example to all of us") and mastery goal orientations (e.g., "Our maths[/science] teacher really wants us to enjoy learning new things"). All items were measured by 7-point Likert scales ranging from 1 (not at all) to 7 (extremely) using items from the Patterns of Adaptive Learning Scales (PALS; Midgley et al., 2000). For mastery orientation, three out of five items from PALS were administered in the interest of total survey length (αs = 0.85/0.88 for mathematics/science, respectively); for performance-approach, all three items were used (αs = 0.80/0.87).

Achievement Striving

Aimed marks were measured by the question "What mark do you (realistically) aim for in Year 10 maths[/science] exams?", selected from 0 (<50%), 1 (50–54%), 2 (55–59%), 3 (60–64%), 4 (65–69%), 5 (70–74%), 6 (75–79%), 7 (80–84%), 8 (85–89%), 9 (90–94%), 10 (95–100%). Effort exertion was tapped by two items (e.g., "How much effort do you put into maths[/science]?") rated from 1 = not at all, to 7 = extremely (αs = 0.77/0.86).

STEM Career Aspirations

Mathematics/science-related career aspirations were subjectively assessed by the question "How much would you like to have a maths[/science]-related career?", rated from 1 (not at all) to 7 (extremely). As well, students responded to an open-ended question: "What career are you mainly considering for your future?". Each career plan was subsequently objectively coded for the extent of mathematics/physics/chemistry/biology/engineering knowledge required for that occupation, using the United States Department of Labor O*NET 2016 (National Center for O*NET Development, 2016), based on data from workers and occupation experts. This yielded a continuous score for each STEM dimension (from 0 to 100) for knowledge required per occupation. For example, for physiotherapists: mathematics = 38, physics = 43, chemistry = 31, biology = 74, engineering = 17. If students listed more than one occupation, the one that they listed first was coded. Occupational data could be coded for 796 of the 824 participants who nominated a career plan.

Psychological Wellbeing

Psychological wellbeing was measured by the $DASS_{21}$ (Depression Anxiety Stress Scales; Lovibond and Lovibond, 1995), tapping 3 factors measured by 21 statements rated in relation to the past week (0: did not apply to me at all; 1: applied to me to some degree, or some of the time; 2: applied to me to a considerable degree, or a good part of the time; 3: applied to me very much, or most of the time). The DASS has been used in a range of previous studies to assess psychological wellbeing (e.g., Chu and Richdale, 2009; Asante, 2012; Giallo et al., 2013; Larcombe et al., 2013; Sharp et al., 2016). Sample items were "I felt that I had nothing to look forward to" for Depression, "I felt I was close to panic" for Anxiety, and "I tended to over-react to situations" for Stress (see **Table 3**).

Gender was coded 0 for boys, 1 for girls.

TABLE 3 | Depression, anxiety, stress (DASS) constructs and items.

Depression (α = 0.88)	
	I found it hard to wind down.
	I couldn't seem to experience any positive feeling at all.
	I found it difficult to work up the initiative to do things.
	I felt that I had nothing to look forward to.
	I felt down-hearted and blue.
	I was unable to become enthusiastic about anything.
	I felt I wasn't worth much as a person.
	I felt that life was meaningless.
Anxiety (α = 0.84)	
	I was aware of dryness of my mouth.
	I experienced breathing difficulty (e.g., excessively rapid breathing, breathlessness in the absence of physical exertion).
	I experienced trembling (e.g., in the hands).
	I was worried about situations in which I might panic and make a fool of myself.
	I felt I was close to panic.
	I was aware of the action of my heart in the absence of physical exertion (e.g., sense of heart rate increase, heart missing a beat).
	I felt scared without any good reason.
Stress (α = 0.85)	
	I felt that I was rather touchy.
	I tended to over-react to situations.
	I felt that I was using a lot of nervous energy.
	I found myself getting agitated.
	I found it difficult to relax.
	I was intolerant of anything that kept me from getting on with what I was doing.

0 = Did not apply to me at all; 1 = applied to me to some degree, or some of the time; 2 = applied to me to a considerable degree, or a good part of time; 3 = applied to me very much, or most of the time ("over the past week").

Analyses

Confirmatory factor analysis for each of the mathematics and science latent motivation constructs (perceived talent, intrinsic and utility values, together with the new items tapping effort/psychological/social costs), and psychological wellbeing constructs (depression, anxiety, and stress) were specified, with items as indicators for only their assigned latent constructs, no cross-loadings or error covariances, using listwise deletion for low missing data initially. Fit adequacy was assessed following Hu and Bentler's (1999) two-index strategy, according to the comparative fit index (CFI ≥ 0.95) and standardised root mean square residual (SRMR ≤ 0.09), as well as the root mean square error of approximation (RMSEA < 0.06) and Tucker-Lewis Index (TLI close to 0.95). Data showed an acceptable fit for

TABLE 4 | Confirmatory factor analysis factor loadings (LX) and measurement errors (TD) for mathematics and science expectancy-value constructs and Cronbach alpha reliabilities.

	Subscale α	Item	LX	TD
Talent	0.91/0.93	TAL1*	0.813/0.892	0.627/0.361
		TAL2	0.847/0.902	0.698/0.411
		TAL3	0.832/0.899	0.674/0.405
		TAL4	0.879/0.858	0.576/0.583
Intrinsic	0.95/0.94	INT1*	0.929/0.917	0.364/0.418
		INT2	0.942/0.911	0.298/0.441
		INT3	0.930/0.933	0.369/0.347
Utility	0.89/0.89	UTIL1*	0.847/0.888	0.633/0.503
		UTIL2	0.895/0.855	0.457/0.687
		UTIL3	0.849/0.817	0.518/0.859
Effort	0.77/0.86	EFF1*	0.907/0.950	0.325/0.186
		EFF2	0.961/0.940	0.145/0.227
Cost–effort	0.77/0.86	EFFC1*	0.751/0.837	1.116/0.816
		EFFC2	0.776/0.828	0.971/0.929
		EFFC3	0.682/0.786	1.488/1.111
Cost–psychological	0.85/0.86	PSYC1*	0.866/0.900	0.830/0.626
		PSYC2	0.850/0.832	0.947/1.011
Cost–social	0.81/0.89	SOCC1*	0.854/0.926	0.507/0.264
		SOCC2	0.811/0.887	0.635/0.391

**Item set as reference per subscale. Coefficients are reported adjacent for mathematics (left side) and science models (right side).*

TABLE 5 | Confirmatory factor analysis factor loadings (LX) and measurement errors (TD) for DASS constructs and Cronbach alpha reliabilities.

	Subscale α	Item	LX	TD
Depression	0.88	DASS1*	0.492	0.694
		DASS3	0.727	0.401
		DASS5	0.607	0.658
		DASS10	0.746	0.413
		DASS13	0.810	0.299
		DASS16	0.778	0.326
		DASS17	0.759	0.362
		DASS21	0.704	0.461
Anxiety	0.84	DASS2*	0.479	0.774
		DASS4	0.615	0.473
		DASS7	0.689	0.397
		DASS9	0.655	0.629
		DASS15	0.777	0.306
		DASS19	0.683	0.417
		DASS20	0.712	0.379
Stress	0.85	DASS6*	0.658	0.587
		DASS8	0.741	0.424
		DASS11	0.750	0.412
		DASS12	0.728	0.475
		DASS14	0.669	0.493
		DASS18	0.670	0.459

**Item set as reference per subscale. Standardised error covariance between DASS17 and DASS18 TD = 0.397.*

TABLE 6 | ICC and design effect for clustering variables.

	Mathematics		Science	
	ICC	*deff*	ICC	*deff*
Perc. talent	0.34	6.07	0.15	3.23
Intrinsic value	0.31	0.69	0.21	0.79
Utility value	0.12	0.88	0.19	0.81
Effort cost	0.14	0.86	0.17	0.83
Psych. cost	0.07	0.93	0.13	0.87
Social cost	0.08	0.92	0.06	0.94

deff, design effect.

each of the mathematics (χ^2 = 691.31, df = 131, CFI = 0.96, TLI = 0.95, RMSEA = 0.06, SRMR = 0.04) and science models (χ^2 = 764.77, df = 131, CFI = 0.96, TLI = 0.95, RMSEA = 0.07, SRMR = 0.04), with details reported in **Table 4**. The fit was marginal for the psychological wellbeing model (χ^2 = 1283.72, df = 186, CFI = 0.91, TLI = 0.89, RMSEA = 0.08, SRMR = 0.05), subsequently improved through correlating measurement errors for Depression items "I felt I wasn't worth much as a person" (DASS17) and "I felt that life was meaningless" (DASS21) (TD = 0.40 standardised estimate; revised model fit χ^2 = 1140.29, df = 185, CFI = 0.92, TLI = 0.91, RMSEA = 0.07, SRMR = 0.05); see **Table 5**.

Latent profile analysis (LPA) educed profiles within each of mathematics and science, among the 1,172 Australian grade 10 students' perceived talent, intrinsic and utility values, and effort/psychological/social costs. The intraclass correlation (ICC) reveals the extent to which individual responses are attributable, in this case, to mathematics/science classroom membership (calculated by dividing the between-cluster variance by the between- plus within-cluster variance; Raudenbush and Bryk, 2002). ICC $\geq$ 0.05 may provide evidence of a classroom membership effect (LeBreton and Senter, 2008). A more definitive indicator is the design effect (*deff*), a function of the ICC and average cluster size [approximately equal to 1 + (average cluster size − 1) × ICC; Kish, 1965]. *deff* $\geq$ 2 indicates that the clustered structure should be considered to avoid biassed estimates of standard errors (Maas and Hox, 2004). *deff* values on the clustering variables were well below 2.0 (see **Table 6**), except for students' perceived talent in each of mathematics/science, presumably due to achievement-streamed classes in secondary schooling. We therefore accounted for non-independence of observations[2] of students within each of their 74 mathematics and 74 science classes (see Asparouhov, 2005, 2006; Maas and Hox, 2005) by employing the robust maximum likelihood estimator and design-based correction for standard errors and chi-square test of model fit available within *Mplus* version 8.1 (Type = Mixture Complex). Mean numbers of student respondents across classes were 15.78 in mathematics (*SD* = 6.60) and 15.80 in science (*SD* = 7.64). Cases that had missing values for all variables (*n* = 4 in mathematics, 15 in science) or for the classroom variable (*n* = 4 in mathematics, 3 in science) were excluded from each analysis and FIML was employed to address the low remaining missing data.

[2]In actuality, only 2 students were differently classified (in the mathematics LPA) when single-level analyses were conducted.

A series of analyses compared one to five latent profile models whose fit indices are presented in **Table 8**. The optimal number of profiles was decided based on a range of widely used statistical criteria (Nylund et al., 2007), supported by substantive interpretability: the Akaike information criterion (AIC; Akaike, 1974), Bayesian information criterion (BIC; Schwarz, 1978), sample-size-adjusted Bayesian information criterion (aBIC), entropy, adjusted Lo–Mendell–Rubin Likelihood Ratio Test (LMR LRT), and Vuong–Lo–Mendell–Rubin Likelihood Ratio Test (VLMR LRT). The optimised solution is expected when values for the AIC/BIC/aBIC are lowest – or when the plotted curve begins to flatten (Masyn, 2013) – indicating that little further information would be gained through additional profiles; entropy >0.80 (Rost, 2006); and LMR/VLMR LRT p-values indicate when the $k-1$ class null model should be rejected in favour of k classes.

Subsequently, using SPSS (version 24) and listwise deletion for missing data, initial ANOVAs compared achievement backgrounds by profiles in mathematics and science, and potential dependency on selective schools in mathematics. Seven M/ANOVAs with Bonferroni protected p-values were then conducted comparing each of mathematics (with and without the NAPLAN achievement covariate) and science profiles, on each of (i) the motivation constructs from which profiles were educed (i.e., perceived talent, intrinsic and utility values, effort/psychological/social costs), (ii) achievement background, (iii) experienced learning environments (mastery and performance goal structures), (iv) aimed marks, (v) effort exerted and subjective mathematics/science career aspiration, (vi) objectively coded STEM career plans, and (vii) psychological wellbeing (depression, anxiety, stress). Chi-square tested associations of profiles with gender, language background, and mothers'/fathers' level of education.

RESULTS

To address our first central aim, associations were examined among mathematics/science expectancy-value motivations including costs and potential antecedent (demographics, achievement background, and learning environment) and outcome factors (achievement striving, career aspirations, and psychological wellbeing). This yielded a rich nomological network for the new cost factors, including psychological wellbeing together with achievement-related outcomes. For our second aim, latent profiles of students were educed in each domain, cross-domain membership was examined, and potential antecedents and outcomes were compared.

Associations Among Motivations, Potential Antecedents, Achievement Striving, Career Aspirations, and Psychological Wellbeing

Latent correlations among all latent and observed constructs (see **Table 7**) were obtained from a combined CFA including demographic factors (gender, English home language, mother and father highest level of education, selective school), mathematics and science achievement background (grade 9 NAPLAN band and usual science score, respectively), expectancy-value constructs in each of mathematics and science (perceived talent, intrinsic and utility values, effort/psychological/social costs), mathematics and science class mastery and performance climates, aimed marks and effort exerted, mathematics/science-related career aspirations, and psychological wellbeing (depression, anxiety, and stress). In this model, items were specified only as indicators of their respective theorised latent constructs, measurement errors were covaried between the same items pertaining to mathematics/science parallel constructs (due to their parallel wording), the measurement error covariance between depression items was retained (DASS 17, 21), and all latent correlations between constructs were estimated using full information maximum likelihood to account for missing data. The data fitted the model acceptably, $\chi^2 = 6481.79$, df = 2978, CFI = 0.95, TLI = 0.93, RMSEA = 0.03. Relationships between core study constructs are highlighted below (**Table 7** shows all relationships for completeness).

Student Gender

In mathematics, gender was unrelated to NAPLAN achievement background, achievement striving in the form of aimed marks and effort exerted. Yet, gender was weakly negatively associated with mathematical motivations of perceived talent, intrinsic and utility values, and social cost (boys higher), and positively with psychological cost (girls higher). Gender negatively related to mathematics mastery and performance-oriented classroom learning environments. Similarly, for science, gender was weakly positively associated with aimed marks, but not effort exertion. Gender was positively associated with social cost and negatively associated with psychological cost, but unassociated with any other science motivation. Gender negatively related to performance-oriented science learning environments (unrelated to mastery), and subjectively reported mathematical career aspirations (but not science career). For objectively scored STEM-related career plans according to O*NET, biological career aspirations were positively related to gender; but physics, mathematics, and engineering-related careers showed negative relationships. For dimensions of psychological wellbeing (depression, anxiety, and stress), gender positively related to stress (girls higher).

Achievement Background

Grade 9 mathematics (nationally assessed NAPLAN score) and science (self-reported 'usual' mark) achievement backgrounds were strongly positively associated. Mathematics achievement background significantly associated with all mathematics motivation factors: positively with perceived talent, intrinsic and utility values; negatively with all costs. Mathematics achievement background was unrelated to mathematics classroom learning environments. Students with higher mathematics achievement background aimed for substantially higher marks, exerted more effort, aspired to higher STEM careers, and reported lower depression, anxiety, and stress. Similarly, for science, achievement background was positively related to perceived

TABLE 7 | Latent correlations between study constructs.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
(1) Gender	–																		
(2) Language	−0.04	–																	
(3) MoEd level	−0.02	0.15**	–																
(4) FaEd level	0.03	0.07†	0.54**	–															
(5) Selective school	0.21**	−0.16**	0.14**	0.14**	–														
(6) M NAPLAN ach.	0.04	−0.12**	0.22**	0.22**	0.47**	–													
(7) S ‘usual’ ach.	0.14**	−0.11**	0.14**	0.12**	0.43**	0.54**	–												
(8) M talent	−0.18**	−0.13**	0.08†	0.13	0.01*	0.45**	0.34**	–											
(9) M intrinsic	−0.17**	−0.19	0.07†	0.12**	0.25**	0.39**	0.32**	0.73**	–										
(10) M utility	−0.15**	−0.08*	0.05	0.08†	0.14**	0.16**	0.24**	0.44**	0.57**	–									
(11) M cost-eff	0.02	0.03	−0.08†	−0.09*	−0.27**	−0.31**	−0.36**	−0.46**	−0.56**	−0.54**	–								
(12) M cost-psy	0.09*	0.06	−0.08†	−0.09*	−0.16**	−0.23**	−0.18**	−0.31**	−0.34**	−0.13**	0.52**	–							
(13) M cost-soc	−0.20**	−0.05	−0.08†	−0.07†	−0.17**	−0.13**	−0.16**	0.03	−0.05	−0.08†	0.40**	0.44**	–						
(14) M mastery	−0.17**	−0.02	−0.02	0.04	0.04	0.01*	0.05	0.22**	0.33**	0.29**	−0.18**	−0.04	−0.05	–					
(15) M perf.	−0.30**	0.17**	0.05	0.03	−0.30**	−0.06	−0.19**	0.01*	0.05	0.04	0.08†	0.03	0.21**	0.21**	–				
(16) M aimed%	0.02	−0.25**	0.11**	0.15**	0.35**	0.52**	0.55**	0.60**	0.56**	0.31**	−0.42**	−0.23**	−0.06	0.12**	−0.09*	–			
(17) M effort	−0.06	−0.04	0.01*	0.07	0.09*	0.17**	0.23**	0.47**	0.55**	0.37**	−0.38**	−0.11**	−0.02	0.26**	0.03	0.37**	–		
(18) S talent	−0.08	0.04	0.14**	0.06†	0.06	0.30**	0.55**	0.44**	0.28**	0.26**	−0.27**	−0.15**	−0.02	0.10*	0.04	0.27**	0.25**	–	
(19) S intrinsic	0.00	−0.07†	0.01*	0.04	0.24**	0.28**	0.53**	0.30**	0.35**	0.36**	−0.35**	−0.17**	−0.10*	0.15**	−0.10*	0.28**	0.27**	0.71**	–
(20) S utility	0.06	−0.12**	0.09	0.01*	0.30**	0.33**	0.48**	0.33**	0.38**	0.42**	−0.36**	−0.16**	−0.14**	0.13**	−0.14**	0.37**	0.27**	0.52**	0.75**
(21) S cost-effort	−0.05	0.08*	−0.10*	−0.10*	−0.32**	−0.31**	−0.49**	−0.23**	−0.26**	−0.31**	0.54**	0.32**	0.30**	−0.11*	0.16**	−0.26**	−0.22**	−0.52**	−0.66**
(22) S cost-psych	0.15**	−0.03	−0.14**	−0.10*	−0.10*	−0.12**	−0.23**	−0.12**	−0.11**	−0.10*	0.27**	0.55**	0.30**	−0.07†	0.03	−0.04	−0.07*	−0.33**	−0.31**
(23) S cost-social	−0.21**	−0.04	−0.01*	−0.09*	−0.20**	−0.14**	−0.22**	0.00	−0.03	−0.07†	0.30**	0.26**	0.71**	−0.04	0.24**	−0.07†	−0.04	−0.10*	−0.19**
(24) S mastery	0.04	0.04	0.04	−0.01	0.11**	0.07	0.18**	0.10*	0.08†	0.15**	−0.14**	−0.05	−0.14**	0.29**	−0.02	0.03	0.13**	0.27**	0.43**
(25) S perf.	−0.30**	0.10	−0.02	−0.03	−0.35**	−0.15**	−0.19**	0.02	−0.03	−0.02	0.10*	0.07	0.22**	0.08†	0.59**	−0.18**	−0.07†	0.09*	−0.05
(26) S aimed%	0.14**	−0.13**	0.14	0.12**	0.33**	0.50**	0.77**	0.36**	0.32**	0.24**	−0.34**	−0.15**	−0.14**	0.07†	−0.16**	0.68**	0.25**	0.56**	0.56**
(27) S effort	0.05	−0.03	0.09*	−0.01	0.18	0.20**	0.43**	0.25**	0.26	0.26	−0.27**	−0.07**	−0.07	0.09*	−0.09*	0.24**	0.46**	0.59**	0.71**
(28) M career	−0.19**	−0.17**	−0.01	0.02	0.13**	0.26**	0.24**	0.57**	0.64**	0.51**	−0.48**	−0.23**	−0.01	0.19**	0.09*	0.40**	0.33**	0.23**	0.24**
(29) S career	0.03	−0.12**	0.06†	0.04	0.32**	0.32**	0.50**	0.33**	0.36**	0.36**	−0.40**	−0.14**	−0.11*	0.14**	−0.09*	0.37**	0.25**	0.55**	0.71**
(30) O*NET biology	0.18**	−0.12**	0.02	0.04	0.20**	0.14**	0.28**	0.18**	0.20**	0.15**	−0.23**	−0.07†	−0.07†	0.02	−0.12*	0.19**	0.16**	0.28**	0.39**
(31) O*NET chem.	0.01	−0.15**	0.03	0.03	0.26**	0.19**	0.31**	0.24**	0.29**	0.26**	−0.32**	−0.15**	−0.05	0.09†	−0.11*	0.24**	0.21**	0.30**	0.44**
(32) O*NET physics	−0.25**	−0.05**	0.01	0.01	0.17**	0.15**	0.22**	0.22**	0.28**	0.27**	−0.30**	−0.19**	−0.01	0.14**	0.03	0.20**	0.19**	0.21**	0.31**
(33) O*NET math	−0.29**	−0.13**	0.03	−0.01	0.20**	0.14**	0.19**	0.26**	0.37**	0.33**	−0.37**	−0.20**	0.02	0.16**	0.04	0.22**	0.23**	0.18**	0.29**
(34) O*NET engineer	−0.34**	−0.04	0.00	0.02	0.08	0.13**	0.10*	0.13**	0.17**	0.22**	−0.21**	−0.11*	0.01	0.11*	0.07†	0.12**	0.10*	0.05†	0.12**
(35) Depression	0.02	−0.02	−0.08†	0.00	−0.04	−0.12**	−0.12**	−0.11**	−0.12**	−0.10†	0.25**	0.32**	0.25**	−0.09†	0.05	−0.07†	−0.16**	−0.12**	−0.12**
(36) Anxiety	0.01	−0.07†	−0.10*	−0.01	−0.07†	−0.14**	−0.13**	−0.07†	−0.05	−0.01	0.20**	0.32**	0.28**	−0.02	0.07	−0.06	−0.06	−0.08†	−0.07†
(37) Stress	0.11**	−0.05	−0.06	0.02	0.00	−0.09*	−0.06	−0.05	−0.03	0.02	0.11*	0.34**	0.21**	−0.02	0.04	0.00	−0.02	−0.07†	−0.06

(Continued)

TABLE 7 | Continued

	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
(21) S cost-effort	−0.56**	–																
(22) S cost-psych	−0.14**	0.58**	–															
(23) S cost-social	−0.23**	0.45**	0.43**	–														
(24) S mastery	0.32**	−0.23**	−0.09*	−0.13**	–													
(25) S perf.	−0.08†	0.17**	0.05	0.30**	0.10*	–												
(26) S aimed%	0.53**	−0.47**	−0.16**	−0.18**	0.21**	−0.17**	–											
(27) S effort	0.58**	−0.48**	−0.13**	−0.01	0.36**	−0.01	0.48**	–										
(28) M career	0.29**	−0.24**	−0.08†	0.04	0.05	0.05	0.23**	0.20**	–									
(29) S career	0.67**	−0.66**	−0.21**	−0.18**	0.23**	−0.07†	0.51**	0.47**	0.38**	–								
(30) O*NET biology	0.38**	−0.40**	−0.04	−0.08†	0.08†	−0.12**	0.28**	0.28**	0.17**	0.56**	–							
(31) O*NET chem.	0.41**	−0.45**	−0.10*	−0.04	0.10*	−0.11*	0.28**	0.30**	0.29**	0.58**	0.80**	–						
(32) O*NET physics	0.27**	−0.32**	−0.17**	0.01	0.07†	−0.02	0.18**	0.19**	0.34**	0.38**	0.31**	0.56**	–					
(33) O*NET math	0.24**	−0.27**	−0.15**	0.05	0.08†	−0.03	0.16**	0.22**	0.44**	0.31**	0.20**	0.54**	0.67**	–				
(34) O*NET engineer	0.09*	−0.14**	−0.15**	0.02	0.03	0.02	0.07†	0.08†	0.27**	0.11**	−0.20**	0.20**	0.69**	0.58**	–			
(35) Depression	−0.05	0.23**	0.25**	0.25**	−0.12**	0.09*	−0.09*	−0.12**	−0.03	−0.10*	−0.04	−0.07	−0.09†	−0.12**	−0.04	–		
(36) Anxiety	−0.02	0.18**	0.26**	0.26**	−0.07†	0.12**	−0.10*	−0.05	0.01	−0.08**	−0.06	−0.06	−0.05	−0.05	0.02	0.81**	–	
(37) Stress	0.03	0.14**	0.28**	0.18**	−0.07	0.07†	−0.03	−0.05	−0.01	−0.05	−0.01	−0.02	−0.06	−0.07	−0.04	0.85**	0.93**	–

*Gender coded 1 = girls; Language coded 1 = English; Selective school coded 1 = Selective. 'M' denotes mathematics, 'S' science factors. †$p < 0.05$, *$p < 0.01$, **$p < 0.001$.*

talent, intrinsic, and utility values, and negatively related to costs. Science achievement background positively related to mastery-, and negatively to performance-oriented learning environments. Science achievement positively associated with aimed marks, effort exertion, STEM career aspirations, and lower levels of depression and anxiety, but not stress.

Motivation Factors

Among mathematics motivation factors, there were strong positive correlations between perceived talent, intrinsic, and utility values. The new cost factors negatively related to those positive motivations (social cost only related to utility value). The cost factors were moderately positively intercorrelated. Similarly, for science, there were strong positive correlations among perceived talent, intrinsic, and utility values. Cost factors were inversely related to these, and positively intercorrelated themselves. Between mathematics and science, positive motivation factors (perceived talent, intrinsic, and utility values) were all positively, moderately intercorrelated. The new cost factors also positively intercorrelated – strongly between corresponding mathematics and science costs, moderately between others.

Motivations, Achievement Striving, and Career Aspirations

Within each of mathematics and science, positive motivation factors (perceived talent, intrinsic, and utility values) were positively associated with aimed marks, effort exertion and STEM career aspirations; cost factors were negatively associated (social cost weakly negatively correlated only with subjectively rated science career and objectively scored biology career). In each of mathematics and science, higher achievement striving (aimed marks and effort exerted) related to STEM career aspirations. Subjectively rated mathematics and science career aspirations were moderately positively correlated. Objectively scored STEM career aspiration dimensions were highly correlated for biology and chemistry; also strongly among physics, engineering, mathematics, and chemistry; weakly between biology and others.

Experienced Learning Environments

Experienced mathematics mastery-oriented learning environment positively related to mathematics achievement striving; performance-oriented environments negatively related to aimed marks. Mastery environments positively related to all three positive motivations; and, negatively related to effort and psychological costs (unrelated for social cost). Conversely, mathematics performance environments negatively associated with intrinsic and utility values; positively with effort and social costs (unrelated for psychological cost). Mathematics mastery environments positively related to STEM-related career aspirations (except biology), whereas performance environments showed a mixed pattern of relationships.

In science, mastery environments positively related to achievement striving; performance environments negatively related to aimed marks. Science mastery environments positively related to all positive science motivations; negatively with all costs. Science performance-oriented environments weakly positively associated with perceived talent, negatively with utility value, and positively with effort and social costs. Science mastery environments positively related to subjectively rated science career aspirations and objectively scored chemistry, physics and mathematics careers. Science performance environments negatively associated with subjectively rated science careers and objectively scored biology and chemistry careers.

Psychological Wellbeing

Depression, anxiety, and stress were highly positively inter-correlated. Mathematics intrinsic and utility values were negatively associated with depression; perceived talent negatively associated with depression and anxiety. Depression negatively associated with mathematics aimed marks and effort exertion. The three mathematics costs positively (moderately) associated with depression, anxiety, and stress. In science, perceived talent and intrinsic value (but not utility value) negatively associated with depression and anxiety; perceived talent additionally negatively associated with stress. Aimed marks in science negatively associated with depression and anxiety; effort exertion associated with reduced anxiety. Science costs positively associated with depression, anxiety, and stress. Students who scored higher on depression aspired to less subjectively rated science careers and objectively scored physics and mathematics careers. Students who scored higher on anxiety aspired to lower subjectively rated science careers. Mastery learning environments exhibited a weak negative association with depression; science mastery environments additionally associated with reduced anxiety. Only science performance environments (not mastery) associated with higher depression, anxiety, and stress.

Motivational Profiles

To move beyond patterns of association between variables and achieve our aim of identifying person-centred patterns, latent profile analysis (LPA) discerned three profiles in each of mathematics and science, based on the range of fit indices reported in **Table 8**. The profiles were named "Positively engaged" (high scores on the 3 positive motivation factors, low on the 3 negative costs), "Disengaged" (low on perceived talent and intrinsic value, high utility value and costs), and "Struggling ambitious" (high scores both on positive and negative factors). For mathematics, 638 students were Positively engaged (54.8%), 169 were Struggling ambitious (14.5%) and 357 were Disengaged (30.7%). In science, 726 students were Positively engaged (62.9%), 186 were Struggling ambitious (16.1%) and 242 were Disengaged (21.0%). Mathematics and science profile memberships were significantly associated, $\chi^2(N = 1147, df = 4) = 306.365, p < 0.001$. However, sizeable off-diagonal numbers indicated a substantial degree of non-overlap (i.e., domain specificity), most pronounced between Struggling ambitious and Disengaged profiles.

Profile memberships differed by gender as anticipated. In mathematics, boys were overrepresented among the Struggling ambitious (20.5% boys versus 10.5% girls), and girls among the Disengaged (22.1% boys versus 36.4% girls), $\chi^2(N = 1164, df = 2) = 39.397, p < 0.001$. In science, boys were again overrepresented among Struggling ambitious (22.9% boys versus 11.5% girls), but girls among the Positively engaged (57.6% boys versus 66.5% girls), $\chi^2(N = 1154, df = 2) = 26.574, p < 0.001$.

TABLE 8 | Model fit indices for 1—5 latent profiles.

	Number of latent profiles				
	1	2	3	4	5
Mathematics					
AIC	24528.45	23341.18	22781.97	22495.83	22349.24
BIC	24589.16	23437.31	22913.52	22662.80	22551.63
aBIC	24551.05	23376.96	22830.94	22557.98	22424.57
Entropy		0.774	0.841	0.782	0.813
aLMR		1177.442	561.837	294.189	157.406
p aLMR		0.000	0.013	0.024	0.193
VLMR		−12252.20	−11651.60	−11365.00	−11214.90
p VLMR		0.000	0.012	0.022	0.187
Science					
AIC	24720.41	23111.20	22509.57	22133.62	21850.41
BIC	24781.02	23207.17	22640.90	22300.30	22052.45
aBIC	24742.90	23146.82	22558.32	22195.48	21925.40
Entropy		0.843	0.889	0.815	0.849
aLMR		1590.968	603.403	382.215	291.302
p aLMR		0.000	0.001	0.017	0.006
VLMR		−12348.20	−11536.60	−11228.80	−11033.80
p VLMR		0.000	0.000	0.016	0.006

Profiles were associated with home language background (English versus other), both in mathematics [$\chi^2(N = 1123, df = 2) = 27.249, p < 0.05$] and science [$\chi^2(N = 1110, df = 2) = 8.766, p < 0.05$], although patterns of association differed by domain. In mathematics, non-English language background (NELB) students were overrepresented among the Positively engaged (62.7% NELB versus 51.1% ELB) and English native speakers among the Disengaged (20.4% NELB versus 35.6% ELB). In science, NELB students were overrepresented among the Struggling ambitious (19.8% NELB versus 13.9% ELB) and English native speakers again among the Disengaged (17.4% NELB versus 22.6% ELB).

Parents' levels of education were associated with profile memberships. In mathematics only father's education background associated with student profiles, $\chi^2(N = 1016, df = 6) = 12.700, p = 0.048$. In science, associations were significant both for fathers, $\chi^2(N = 1009, df = 6) = 15.346, p < 0.05$, and mothers, $\chi^2(N = 1067, df = 2) = 25.287, p < 0.05$. In all cases, tertiary parent education levels (i.e., university or TAFE) associated with more Positively engaged profiles. Students whose mothers/fathers had not completed high school were overrepresented in Struggling ambitious profiles, and those whose parents had completed high school as their highest level of education were overrepresented among the Disengaged. **Table 9** summarises demographic characteristics for each profile, in mathematics and science.

Achievement and Striving

Mathematics achievement background

Mathematics NAPLAN bands significantly differed according to mathematics profiles, $F(2,858) = 57.35, p < 0.001$ (all paired differences $p < 0.01$ in Tukey *post hoc* tests). Positively engaged students scored higher bands ($M = 9.20, SD = 1.09$),

TABLE 9 | Latent profile summary demographic characteristics.

	Mathematics							Science						
	Positively engaged		Struggling ambitious		Disengaged			Positively engaged		Struggling ambitious		Disengaged		
	n	%	*n*	%	*n*	%	Total	*n*	%	*n*	%	*n*	%	Total
Mathematics strand (Melbourne)														
General	343	52.77	93	14.31	214	32.92	650	–	–	–	–	–	–	–
Methods	112	77.78	19	13.19	13	9.03	144	–	–	–	–	–	–	–
Mathematics strand (Sydney)														
Standard	0	0	5	41.67	7	58.33	12	–	–	–	–	–	–	–
Intermediate	35	28.46	19	15.45	69	56.10	123	–	–	–	–	–	–	–
Advanced	130	64.68	29	14.43	42	20.90	201	–	–	–	–	–	–	–
[a]Mother highest education														
Part high school	67	49.63	23	17.04	45	33.33	135	76	56.30	29	21.48	30	22.22	135
High school	113	51.36	37	16.82	70	31.82	220	120	55.30	47	21.66	50	23.04	217
TAFE	63	57.27	39	35.45	8	7.27	110	76	70.37	8	7.41	24	22.22	108
University	360	59.11	77	12.64	172	28.24	609	418	68.86	77	12.69	112	18.45	607
Father highest education														
Part high school	55	46.61	21	17.80	42	35.59	118	76	65.52	24	20.69	16	13.79	116
High school	91	48.66	26	13.90	70	37.43	187	104	55.91	37	19.89	45	24.19	186
TAFE	47	56.62	9	10.84	27	32.53	83	55	66.27	10	12.05	18	21.69	83
University	372	59.24	83	13.22	173	27.55	628	422	67.63	79	12.66	123	19.71	624
Selective schooling														
Selective	283	69.70	36	8.87	87	21.43	406	325	79.85	26	6.39	56	13.76	407
Non-selective	355	46.83	133	17.55	270	35.62	758	401	53.68	160	21.42	186	24.90	747
Home language														
NELB	234	62.73	63	16.89	76	20.38	373	231	62.77	73	19.84	64	17.39	368
English	383	51.07	100	13.33	267	35.60	750	471	63.48	103	13.88	168	22.64	742

[a]*No significant association between mother's education and mathematics profiles. All other associations were significant $p < 0.05$.*

followed by Struggling ambitious ($M = 8.66$, $SD = 1.37$), and Disengaged scored lowest ($M = 8.20$, $SD = 1.42$). When selective school membership was added into the model as a second fixed factor, the main effect of mathematics profile remained significant, $F(2,855) = 22.60$, $p < 0.001$, along with a significant main effect of selective school attendance, $F(1,858) = 150.25$, $p < 0.001$. However, Bonferroni pairwise comparisons revealed that Positively engaged and Struggling ambitious profiles had equivalent NAPLAN achievement ($p = 0.14$) once selective school attendance was taken into account, although the other paired comparisons remained significant ($p < 0.05$). **Table 8** shows proportions of students in each profile, who attended selective schools: 69.7% of Positively engaged, 8.9% of Struggling ambitious, and 21.4% of Disengaged. Consequently, construct comparisons by mathematics profiles were conducted with, and without, NAPLAN achievement as a covariate, to understand the unique role of mathematics profile membership on achievement striving, career aspirations, psychological wellbeing, and experienced learning environments.

Science achievement background

For science, self-reported 'usual' grade 9 results differed by science profiles, $F(2,1066) = 145.75$, $p < 0.001$ (all paired differences $p < 0.01$ in Tukey *post hoc* tests). Positively engaged students reported higher usual scores ($M = 7.58$, $SD = 2.11$), followed by Struggling ambitious ($M = 5.34$, $SD = 2.68$), and lowest scores were reported for Disengaged ($M = 4.75$, $SD = 2.95$). Because science assessments may not be comparable across schools or even classes, we decided against controlling for usual reported science scores in subsequent science construct comparisons.

Aimed marks and effort exerted

Positively engaged students aimed for higher marks in mathematics than Struggling ambitious (**Table 10A**), but their aimed marks were equivalent once NAPLAN achievement was included as a covariate (**Table 10B**). Disengaged students aimed for lowest marks in mathematics. All profiles differed significantly on effort exerted – Positively engaged were highest, followed by Struggling ambitious, and Disengaged were lowest. For science, Positively engaged students reported highest aimed marks, followed by Struggling ambitious, and then Disengaged; the same pattern occurred for reported effort.

Perceived Classroom Learning Environment

Both in mathematics and science, Positively engaged students experienced classroom learning environments characterised by highest mastery orientation, whereas Struggling ambitious and Disengaged students scored similarly and lower. Conversely, Struggling ambitious students experienced the most performance focused classroom learning environments,

TABLE 10A | Mathematics and science latent profile descriptive statistics and significant differences.

	Mathematics									Science								
	Positively engaged		Struggling ambitious		Disengaged		*F*	df	*p*	Positively engaged		Struggling ambitious		Disengaged		*F*	df	*p*
	M	*SD*	*M*	*SD*	*M*	*SD*				*M*	*SD*	*M*	*SD*	*M*	*SD*			
Perc. talent	5.07	0.98	4.64	1.17	3.02	1.14	411.27	2,1106	<0.001	5.02	1.01	4.18	1.10	2.91	1.24	333.25	2,1094	<0.001
Intrinsic value	5.26	0.96	4.40	1.27	2.64	1.09	675.15	2,1106	<0.001	5.72	0.89	4.41	1.34	2.85	1.16	676.64	2,1094	<0.001
Utility value	5.86	0.95	5.31	1.14	4.27	1.39	213.76	2,1106	<0.001	5.89	0.96	4.72	1.27	3.66	1.40	362.08	2,1094	<0.001
Effort cost	2.56	1.03	4.07[a]	1.20	4.07[a]	1.20	256.07	2,1106	<0.001	2.44	1.07	4.33[a′]	1.12	4.48[a′]	1.38	373.93	2,1094	<0.001
Psych. cost	3.08	1.52	5.05	1.28	4.38	1.62	147.95	2,1106	<0.001	3.04	1.59	4.74	1.15	4.10	1.74	101.65	2,1094	<0.001
Social cost	1.43	0.61	4.31	1.01	1.66	0.85	967.53	2,1106	<0.001	1.39	0.62	4.40	0.97	1.60	0.80	1224.61	2,1094	<0.001
Achievement background	9.20	1.09	8.66	1.37	8.20	1.42	57.35	2,858	<0.001	7.58	2.11	5.34	2.68	4.75	2.95	145.75	2,1066	<0.001
Performance	3.02[a]	1.68	3.76	1.66	3.01[a]	1.54	14.60	2,1146	<0.001	2.75[a′]	1.54	3.95	1.55	2.87[a′]	1.62	42.30	2,1114	<0.001
Mastery	5.22	1.37	4.89[a]	1.35	4.60[a]	1.59	21.39	2,1146	<0.001	5.54	1.25	4.89[a′]	1.48	4.66[a′]	1.60	43.99	2,1114	<0.001
Mark-aimed	8.44	1.54	7.44	2.27	5.90	2.64	170.63	2,1145	<0.001	8.23	1.62	6.51	2.54	5.52	2.73	175.85	2,1110	<0.001
Effort exerted	5.49	1.07	5.04	1.27	4.33	1.43	100.72	2,1127	<0.001	5.53	1.01	4.70	1.25	3.69	1.40	234.89	2,1109	<0.001
STEM career (subjective)	4.48	1.55	3.96	1.76	2.33	1.44	215.34	2,1127	<0.001	5.33	1.65	3.54	1.91	2.21	1.47	337.78	2,1109	<0.001
STEM career (O*NET)*																		
Biology	43.32[a]	33.87	32.90[a]	32.86	27.65[a]	29.40	3.07	2,773	0.047	45.38	33.908	27.18[a′]	30.19	17.83[a′]	20.69	30.18	2,773	<0.001
Chemistry	39.71[a]	24.62	30.55[a]	23.73	23.47	21.50	9.62	2,773	<0.001	39.76	24.94	26.61[a′]	20.89	18.56[a′]	17.71	29.62	2,773	<0.001
Physics	34.73[a]	23.61	27.62[ab]	20.44	22.21[b]	19.39	9.52	2,773	<0.001	33.79	23.68	25.48[a′]	18.59	20.95[a′]	18.65	12.72	2,773	<0.001
Mathematics	60.77[a]	17.187	55.41[a]	17.84	47.89	16.15	17.76	2,773	<0.001	58.63[a′]	17.75	55.39[a′]	15.93	48.68	17.70	7.15	2,773	0.001
Engineering	36.35[a]	28.00	32.53[ab]	23.61	26.63[b]	21.30	5.10	2,773	0.006	34.20[a′]	27.43	31.42[a′]	22.37	29.83[a′]	22.83	0.99	2,773	0.373
Depression	0.68	0.60	1.16	0.79	0.94	0.77	34.76	2,1038	<0.001	0.70	0.63	1.20	0.82	0.92	0.71	38.84	2,1032	<0.001
Anxiety	0.57	0.55	0.99	0.78	0.75	0.72	27.98	2,1038	<0.001	0.59[a′]	0.58	1.05	0.82	0.68[a′]	0.62	34.92	2,1032	<0.001
Stress	0.80	0.68	1.16	0.82	0.95	0.76	16.11	2,1038	<0.001	0.82[a′]	0.71	1.16	0.85	0.93[a′]	0.68	14.57	2,1032	<0.001

[a,b]*Paired letters within each row denote non-significant differences. All other means within each row significantly differ at $p < 0.05$.*
**O*NET scores were compared within a single MANOVA that included mathematics and science profiles as independent variables (n.s. interaction effect).*

TABLE 10B | Mathematics latent profile descriptive statistics and significant differences – with NAPLAN covariate.

	Positively engaged		Struggling ambitious		Disengaged		*F*	df	*p*
	M	***SD***	***M***	***SD***	***M***	***SD***			
Perc. Talent	5.12[a]	0.99	4.79[a]	1.14	3.07	1.13	237.53	2, 817	<0.001
Intrinsic value	5.28	0.98	4.47	1.25	3.07	1.13	387.96	2, 817	<0.001
Utility value	5.82	0.96	5.32	1.15	4.30	1.39	130.62	2, 817	<0.001
Effort cost	2.51	1.01	4.06[a]	1.23	4.03[a]	1.20	153.77	2, 817	<0.001
Psych. cost	3.02	1.51	5.06	1.33	4.31	1.66	86.52	2, 817	<0.001
Social cost	1.41	0.61	4.38	1.04	1.61	0.81	703.97	2, 817	<0.001
Performance	2.94[a]	1.64	3.65	1.72	3.00[a]	1.55	8.06	2, 850	<0.001
Mastery	5.18[a]	1.40	4.79[ab]	1.45	4.55[b]	1.65	11.89	2, 850	<0.001
Mark-aimed	8.51[a]	1.49	7.85[a]	1.97	6.02	2.68	76.82	2, 849	<0.001
Effort exerted	5.54	1.07	5.04	1.36	4.35	1.41	67.16	2, 836	<0.001
STEM career (subjective)	4.51	1.54	3.89	1.79	2.36	1.49	123.10	2, 836	<0.001
STEM career (O*NET)*									
Biology	44.34[a]	33.66	37.68[a]	33.51	29.98[a]	30.95	2.51	2, 576	0.082
Chemistry	40.83[a]	25.10	33.30[a]	24.65	24.38	20.95	7.13	2, 576	0.001
Physics	34.15[a]	23.48	30.03[ab]	22.11	21.80[b]	18.63	6.16	2, 576	0.002
Mathematics	60.96[a]	17.22	55.25[ab]	19.03	48.00[b]	15.44	12.70	2, 576	<0.001
Engineering	35.36[a]	28.11	33.11[a]	25.48	26.42[a]	20.41	1.89	2, 576	0.152
Depression	0.67	0.60	1.19	0.76	0.94	0.77	25.28	2, 776	<0.001
Anxiety	0.55	0.54	1.01	0.76	0.78	0.72	20.42	2, 776	<0.001
Stress	0.80[a]	0.69	1.19	0.82	0.97[a]	0.76	11.35	2, 776	<0.001

[a,b] *Paired letters within each row denote non-significant differences. All other means within each row significantly differ at* $p < 0.05$.
**O*NET scores were compared within a single MANOVA that included mathematics and science profiles as independent variables (n.s. interaction effect).*

whereas Positively engaged and Disengaged students scored similarly and lower (**Table 10A**). In mathematics, once NAPLAN achievement was included as a covariate to discern unique effects of mathematics profiles, Positively engaged students no longer differed from the Struggling ambitious on mastery environments (**Table 10B**).

Career Aspirations

For subjectively rated mathematics and science career plans, Positively engaged students aspired to careers more highly related to mathematics/science, respectively, followed by Struggling ambitious, and Disengaged students lowest (**Table 10A**). After controlling for NAPLAN achievement, the same pattern in subjectively rated mathematics careers remained (**Table 10B**).

For objectively coded STEM career aspirations, mathematics profiles differed on each of physics, chemistry, mathematics, and engineering careers (no significant differences for biology; **Table 9**). Specifically, Positively engaged and Struggling ambitious profiles were similar and highest for mathematics and chemistry career aspirations. They were also similar and higher on physics and engineering careers, although Struggling ambitious did not significantly differ from Disengaged. When NAPLAN achievement was included as a covariate (**Table 10B**), Positively engaged students had significantly higher mathematics, physics and chemistry-related career aspirations than Disengaged; however, there were no significant differences between profiles on biology or engineering. Struggling ambitious students scored similarly high to Positively engaged students on physics and mathematics career plans, but not significantly higher than Disengaged students. Struggling ambitious students were similarly high to Positively engaged students on chemistry career plans, and significantly higher than Disengaged.

Science profiles differed on each of objectively coded physics, chemistry, biology and mathematics career aspirations (no significant differences for engineering). Positively engaged and Struggling ambitious science profiles were similar and highest for mathematical career aspirations. Positively engaged science students were highest on the other STEM career dimensions, whereas Struggling ambitious and Disengaged students scored similarly and lower.

Psychological Wellbeing

Dimensions of psychological wellbeing were distinguished by profile memberships in each of mathematics and science. In mathematics, all profile differences were significant, with Struggling ambitious students highest on depression, anxiety, and stress, followed by Disengaged, and Positively engaged students lowest (**Table 10A**). With the NAPLAN achievement covariate included (**Table 10B**), Positively engaged and Disengaged profiles no longer differed on stress. In science, Struggling ambitious students were again highest on depression, anxiety, and stress. Positively engaged and Disengaged profiles were similar and lower on anxiety and stress; Disengaged students scored higher on depression (**Table 10A**).

DISCUSSION

This study followed two central aims. First, was to validate the new expectancy-value cost measure among adolescents across the domains of mathematics and science and examine associations with demographics, achievement background, mastery/performance-focused learning environments, achievement striving, STEM-related career aspirations and dimensions of wellbeing. Second, was to discern hypothesised theoretically coherent expectancy-value latent profiles of students within each of mathematics and science, potential antecedents (demographics, experienced learning environments, and achievement background), outcomes (achievement striving, career aspirations, and psychological wellbeing), and degree of domain specificity versus generality.

Expectancy-Value Cost Dimensions

Alongside typically examined positive motivational factors (i.e., perceived talent, intrinsic, and utility values), the purpose-adapted adolescent cost measure (based on Perez et al., 2014) proved well-functioning within each of mathematics and science and was key to identifying the student profiles to address our second aim. Effort, psychological and social costs were moderately positively intercorrelated within each domain (in contrast to extreme correlations between the dimensions by Flake et al., 2015), and were negatively correlated with domain-specific positive motivational factors, except social cost, which was not consistently significantly related. Between domains, the same cost factors were strongly correlated, and different costs were moderately correlated.

Psychological cost was highest rated (more so among girls, consistent with the mathematics anxiety literature), referring to concern about the degree of stress involved. Effort cost referred to the degree of effort required outweighing what students were willing to exert and social cost was least endorsed (intriguingly, higher among boys), tapping concern regarding loss of friendships due to working at mathematics/science. Students from more educationally advantaged backgrounds experienced lower costs. In general, performance-oriented learning environments related to increased effort and social costs in both domains, and mastery-oriented environments related to reduced costs (less so in mathematics). As we had anticipated, costs appeared to undermine achievement striving and STEM-related career aspirations (least so for social cost), and relate to heightened levels of general depression, anxiety and stress.

Motivational Profiles in Mathematics and Science

Regarding the second aim and hypothesis, three profiles were supported in each of mathematics/science: the majority were Positively engaged (high on perceived talent and intrinsic/utility values), approximately one-sixth were Struggling ambitious (high on all), and 30.7%/21.0% (for mathematics/science, respectively) were Disengaged (low perceived talent and intrinsic value, high utility value and costs). The hypothesised Struggling ambitious type resonated with Covington and Omelich's "overstrivers" (1991) and "excessively ambitious" workers in the German occupational health study (Schaarschmidt and Fischer, 1997). Although past studies found more boys to be represented in positive mathematics types (e.g., Chow and Salmela-Aro, 2011; Lazarides et al., 2019), such studies had not included costs in their profiling. With costs included, we found no gender difference in the positive mathematics type (and more girls in the positive science type), but a higher proportion of boys were split into the Struggling ambitious type both in mathematics and science. The Disengaged type enriched the mixed utility focused type identified by Lazarides et al. (2016a,b) in German and Finnish studies. As we had speculated, inconsistent valuing of mathematics/science as being useful, coupled with low interest value and perceived talent, coincided with elevated effort cost and psychological cost. Results were consistent with the research literature showing girls' lower perceived talents and interest in mathematics (e.g., Watt et al., 2012) and higher mathematics anxiety (see Hyde, 2005); more girls fitted this type in mathematics. No gender difference was found in the Disengaged type for science, perhaps because physical and biological sciences were not distinguished, and girls tend to show more interest in biology.

Because mathematics involved a standard national assessment at grade 9, achievement background could be assessed on a common metric, unlike science. The achievement differences between mathematics profiles partly depended on selective school attendance – once accounted for, Positively engaged and Struggling ambitious profiles performed equally high in mathematics, aimed for similarly high marks, and aspired to similarly high objectively coded STEM careers. Struggling ambitious, however, reported most pronounced levels of depression, anxiety and stress. Although Positively engaged and Struggling ambitious experienced similar mastery-oriented learning environments, the Struggling ambitious students experienced more performance-focused environments. Disengaged students had significantly lower background mathematical achievement, exhibited lowest achievement striving, aimed for least STEM-related careers (except biology on which all profiles were similar), and scored in between the other two profiles on psychological wellbeing – equivalently low on stress as the Positively engaged profile once NAPLAN achievement background was controlled. Disengaged mathematics students came from similarly mastery-oriented learning environments as the Struggling ambitious, and similarly performance-oriented environments as the Positively engaged.

In science, Positively engaged students reported highest 'usual' achievement background, exerted highest achievement striving, aspired to highly STEM-related careers in biology, chemistry and physics (similarly high as the Struggling ambitious for mathematics careers; no profile differences for engineering), exhibited lowest levels of depression, and similarly low anxiety and stress as the Disengaged. Positively engaged science students came from the most mastery-oriented learning environments, and equally low performance-oriented environments as the Disengaged.

Overall, the Struggling ambitious profile (more boys) proved most maladaptive. In mathematics, despite equivalent achievement background, achievement striving and STEM-related career aspirations as a consequence of their high positive motivations, they suffered debilitated psychological wellbeing seemingly due to elevated costs, potentially exacerbated by their more performance-oriented learning environments. The Positively engaged profile was the most adaptive, exhibiting similarly high achievement striving and STEM career aspirations as the Struggling ambitious, and the most positive psychological wellbeing. The Disengaged profile (more girls) showed lowest achievement background, striving and career aspirations linked to their lower perceived talent and intrinsic value; and, moderate psychological wellbeing linked to their perceived effort and psychological costs.

In science, it was not possible to compare background achievement on any common metric, to disentangle profile effects for Struggling ambitious students who reported lower 'usual' marks, achievement striving, and some dimensions of STEM career aspirations than the Positively engaged. Struggling ambitious science students (more boys) reported the poorest psychological wellbeing, likely exacerbated by their more performance-oriented learning environments. In contrast, the Positively engaged profile experienced the highest mastery-oriented environments.

The typological approach moved beyond measuring differences in the extent of students' motivations, to examine how they varied qualitatively in combination, enabling a more nuanced examination of complex relationships with potential antecedents and outcomes. Profiles related to past achievement and predicted mathematics/science achievement striving and career intentions, as well as psychological wellbeing. Profiles resonated with those educed using similar methods but different measures within other motivational frameworks and not specific to STEM. For example, our Positively engaged type shared features with the "mastery oriented" group among Finnish adolescents within an achievement goal perspective (Tuominen-Soini et al., 2012) and "engaged" group in another Finnish study among adolescents within a psychological wellbeing framework (highest school value, wellbeing and achievement; Tuominen-Soini and Salmela-Aro, 2014), and the "good motivated strategies for learning" group (positive motivation and achievement, low anxiety) in a study framed by self-determination theory (SDT) of Singaporean adolescents in mathematics and science (Ng et al., 2016).

Our Struggling ambitious type somewhat resembled the "success-oriented" group in the first Finnish study (Tuominen-Soini et al., 2012) who showed high motivation and achievement but emotional distress, the "engaged-exhausted" group in the Finnish psychological wellbeing study (high school value and achievement, but lower self-esteem, preoccupation with failures and depression; Tuominen-Soini and Salmela-Aro, 2014), and the "average motivation, low wellbeing" type in another Finnish study of adolescents in both mathematics and literacy (Parhiala et al., 2018). Similarities were also shown with the "high motivated strategies for learning" group (high motivations, moderate achievement, high anxiety) among Singaporean students (Ng et al., 2016), and "moderate engagement but maladaptive cognitions" in an Australian tertiary study (Elphinstone and Tinker, 2017).

Finally, our Disengaged type resembled the "cynical" group in the Finnish study (lowest school value and achievement, moderate wellbeing; Tuominen-Soini and Salmela-Aro, 2014) and the "poor motivated strategies for learning" group (lowest motivations and achievement, high anxiety) in the Singapore SDT study (Ng et al., 2016).

Implications

The results of this study have theoretical implications. Our study has linked the breadth of expectancy-value constructs including multidimensional costs explicitly to measures of experienced learning environments as conceptualized within achievement goal theory, psychological wellbeing, achievement striving and career aspirations, in relation to two key domains of mathematics and science. Including negative cost values alongside typically measured positive expectancy-value dimensions enabled identification of students who experience particular combinations of motivations and pressures, contrasting with variable-centred approaches where the focus is on normative patterns. Profiles could be conceptualized along dimensions of achievement striving and psychological wellbeing. Similar profiles for mathematics and science, and coherent patterns of antecedents and outcomes, suggest they deserve further investigation. Positively engaged/Struggling ambitious were distinguished by high costs perceived by Struggling ambitious, associated with debilitated psychological wellbeing, but not eroding achievement striving. A greater proportion of boys was in this risk type. Disengaged students reported lowest STEM-related career aspirations, aimed marks, and history of results; in mathematics, a greater proportion of girls was in this risk type. Gender differences may be reflective of the pressures and expectations resulting from entrenched stereotypes that boys should be naturally better in STEM subjects – more boys were Struggling ambitious (in mathematics and science), and that girls are not expected to be good at or interested in mathematics – more girls were Disengaged (in mathematics).

This study provides some practical implications, as well. The fact that there were differences in students' memberships across mathematics/science profiles suggests that while there may be a dispositional or core base, there is much shaped within each learning domain. These differences signal amenability to change by intervention (Crick, 2012). Eccles and her colleagues have demonstrated that girls are engaged by activities they perceive as socially meaningful and important (Watt et al., 2012; Eccles, 2013), but because mathematics is frequently taught in skills-based, abstract, decontextualised ways, it would seem less likely to engage girls. Practical approaches could include making explicit connections between mathematics and its social uses and purposes, to especially help girls who were overrepresented in the Disengaged profile to develop a sense of personal significance and practical value (Su et al., 2009; Eccles and Wang, 2012; Watt et al., 2012). For example, a promising utility-value intervention for parents has shown positive effects on STEM career preparation and pursuit (Rozek et al., 2017). Our findings suggest that the learning environment may also be a potential target of intervention.

Situating students' motivational profiles within the ecology of their experienced learning environments revealed systematic differences across examined mastery- and performance-oriented dimensions. A focus on developing mastery-oriented classrooms could be the starting point to promote students' engagement in STEM through the nature of the motivational environments that teachers create. Distinct profiles of motivation within classrooms suggest that teachers may need to tailor approaches for the different types, keeping in mind that even within subgroups, one size does not fit all (Lawson and Lawson, 2013). Of particular concern is what may happen to the Struggling ambitious students whose experienced learning environments related to heightened experiences of depression, anxiety, and stress. Relatedly, for the "engaged-exhausted" profile identified among Finnish students (Tuominen-Soini and Salmela-Aro, 2014), authors recommended a learning environment that does not focus on performance or social comparison, and cautioned the danger of overlooking these students because of their positive academic motivation and achievement. Covington (1992) similarly expressed concern that the failure avoidance and success of "overstrivers" can come at great cost to their wellbeing and highlighted that risk as a core issue that should be of concern to researchers. In future, it will be important to discover whether Struggling ambitious students adapt their standards and expectations over time and shift to one of the other profiles, or become exhausted, unable to cope and eventually burn out.

Limitations

In interpreting the results, some limitations should be kept in mind. First, the study was limited to grade 10 Australian students from nine metropolitan schools in Melbourne and Sydney, some of which had rather low response rates. It would be very interesting to see if the result patterns can be replicated with older or younger students, and how profile memberships may change over time. Second, this study relied on student-reported achievement assessments in science, which was not comparable across classes or schools. Third, with the exception of background achievement information, data were cross-sectional, and longitudinal or intervention studies would be required to tease apart directionality of effects between profiles and correlates. Finally, this study is limited to the domains of mathematics and science; science in particular is a multifaceted field which could be fruitfully disaggregated into its component disciplines, as especially biology showed different effect patterns in terms of career aspirations.

Conclusion and Outlook

Although similar profiles were educed within each of the cognate domains of mathematics and science, the fact that substantial numbers of students were in different profiles for mathematics versus science supports a domain-specific interpretation. Patterns of correlation for the same constructs across domains strengthens this inference. Student characteristics interacted with features of their learning environments with implications for their achievement striving, career aspirations and psychological wellbeing. This study leads to some intriguing future questions, most importantly, how to shift students to a more adaptive profile? While it was encouraging to observe that most students were characterised as "positively engaged", they may not remain so, especially in unsupportive learning environments. It is also of high concern that Struggling ambitious students suffered elevated costs and debilitated psychological wellbeing, and that sizeable proportions were Disengaged. This implies the need for a dual focus in the design of productive interventions to harness conditions that facilitate and sustain positive motivations, at the same time as providing support for students experiencing anxiety or difficulties. Our study has contributed to the expectancy-value body of literature by furthering our understanding of how dimensions of students' expectancies and values, including costs, combine, and their importance for a range of academic, vocational, and psychological outcomes along the STEM pipeline.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Monash University Human Research Ethics Committee (MUHREC) with written informed consent from all participants in accordance with the MUHREC requirements. The project was approved by MUHREC, the Department of Education and Early Childhood Development (DEECD), NSW Government Education & Communities State Education Research Application Process (SERAP), and the Catholic Education Office Melbourne (CEOM).

AUTHOR CONTRIBUTIONS

All authors listed have made a direct and intellectual contribution to the work, and approved it for publication.

FUNDING

The STEPS Study (Study of Transitions and Education Pathways; www.stepsstudy.org) was supported by Australian Research Council (ARC) DP110100472 and ARF to Watt, 2011–2015, and ARC Future Fellowship FT170100153 to Watt, 2017–2021.

REFERENCES

Ainley, J., Kos, J., and Nicholas, M. (2008). *Participation in Science, Mathematics and Technology in Australian Education*. Melbourne: ACER Research Monographs.

Akaike, H. (1974). A new look at the statistical model identification. *Automat. Control IEEE Trans.* 19, 716–723. doi: 10.1007/978-1-4612-1694-0_16

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *J. Educ. Psychol.* 84, 261–271.

Andersen, L., and Chen, J. (2016). Do high-ability students disidentify with science? A descriptive study of U.S. ninth graders in 2009. *Sci. Educ.* 100, 57–77.

Asante, K. O. (2012). Social support and the psychological wellbeing of people living with HIV/AIDS in Ghana. *Afr. J. Psychiatry* 15, 340–345.

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Struct. Equat. Model.* 12, 411–434. doi: 10.1207/s15328007sem1203_4

Asparouhov, T. (2006). General multi-level modeling with sampling weights. *Commun. Stat. Theory Methods* 35, 439–460. doi: 10.1080/03610920500476598

Atwater, M. M., Wiggins, J., and Gardner, C. M. (1995). A study of urban middle school students with high and low attitudes toward science. *J. Res. Sci. Teach.* 32, 665–677. doi: 10.1002/tea.3660320610

Barron, K. E., and Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: testing multiple goal models. *J. Personal. Soc. Psychol.* 80, 706–722. doi: 10.1037//0022-3514.80.5.706

Battle, A., and Wigfield, A. (2003). College women's value orientations toward family, career, and graduate school. *J. Vocat. Behav.* 62, 56–75.

Bøe, M. V., and Henriksen, E. K. (2013). Love it or leave it: Norwegian students' motivations and expectations for postcompulsory physics. *Sci. Educ.* 97, 550–573. doi: 10.1002/sce.21068

Carmichael, C., Callingham, R., and Watt, H. M. G. (2017). Classroom motivational environment influences on emotional and cognitive dimensions of student interest in mathematics, *ZDM Mathemat. Educ.* 49, 449–460. doi: 10.1007/s11858-016-0831-7

Cheryan, S., Ziegler, S. A., Montaya, A. K., and Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychol. Bull.* 143, 1–35. doi: 10.1037/bul0000052

Chittum, J. R., and Jones, B. D. (2017). Identifying pre-high school students' science class motivation profiles to increase their science identification and persistence. *J. Educ. Psychol.* 109, 1163–1187. doi: 10.1037/edu0000176

Chow, A., Eccles, J. S., and Salmela-Aro, K. (2012). Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Dev. Psychol.* 48, 1612–1628. doi: 10.1037/a0030194

Chow, A., and Salmela-Aro, K. (2011). Task-values across subject domains: a gender comparison using a person-centered approach. *Int. J. Behav. Dev.* 35, 202–209. doi: 10.1177/0165025411398184

Chu, J., and Richdale, A. L. (2009). Sleep quality and psychological wellbeing in mothers of children with developmental disabilities. *Res. Dev. Disabil.* 30, 1512–1522.

Conley, A. M. (2012). Patterns of motivation beliefs: combining achievement goal and expectancy-value perspectives. *J. Educ. Psychol.* 104, 32–47. doi: 10.1037/a0026042

Covington, M. V. (1984). The self-worth theory of achievement motivation: findings and implications. *Elem. Sch. J.* 85, 4–20. doi: 10.1086/461388

Covington, M. V. (1992). *Making the Grade: A Self-Worth Perspective on Motivation and School Reform*. Cambridge: Cambridge University Press.

Covington, M. V., and Omelich, C. L. (1985). Ability and effort valuation among failure-avoiding and failure-accepting students. *J. Educ. Psychol.* 77, 446–459. doi: 10.1037/0022-0663.77.4.446

Covington, M. V., and Omelich, C. L. (1991). "Need achievement revisited: verification of Atkinson's original 2 × 2 model," in *Stress and Emotion*, Vol. 14, eds C. D. Spielberger, I. G. Sarason, Z. Kulcsar, and G. L. Van Heck (New York, NY: Hemisphere).

Crick, R. D. (2012). "Deep engagement as a complex system: Identity, learning power, and authentic enquiry," in *Handbook of Research on Student Engagement*, eds S. L. Christenson, A. L. Reschly, and C. Wylie (New York, NY: Springer), 675–694.

Crouzevialle, M., and Butera, F. (2013). Performance-approach goals deplete working memory and impair cognitive performance. *J. Exp. Psychol. Gen.* 142, 666–678. doi: 10.1037/a0029632

Dobson, I. R. (2012). *Unhealthy Science? University Natural and Physical Sciences 2002 to 2009/10*. Available at: http://www.chiefscientist.gov.au/wp-content/uploads/Unhealthy-Science-Report-Ian-R-Dobson.pdf (accessed January 9, 2019).

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychol.* 41, 1040–1048. doi: 10.1037/0003-066X.41.10.1040

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychol. Women Quart.* 11, 135–172. doi: 10.1111/j.1471-6402.1987.tb00781.x

Eccles, J. S. (2005). "Subjective task value and the Eccles et al. model of achievement-related choices," in *Handbook of Competence and Motivation*, eds A. J. Elliot and C. S. Dweck (New York, NY: Guilford), 105–131.

Eccles, J. S. (2009). "Who am I and what am I going to do with my life?" Personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368

Eccles, J. S. (2013). Gender and STEM: opting in versus dropping out. *Int. J. Gender Sci. Technol.* 5, 2–3.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., et al. (1983). "Expectancies, values and academic behaviors," in *Achievement and Achievement Motives: Psychological and Sociological Approaches*, ed. J. T. Spence (San Francisco, CA: Freeman), 75–146.

Eccles, J. S., Adler, T., and Meece, J. (1984). Sex differences in achievement: a test of alternate theories. *J. Personal. Soc. Psychol.* 46, 26–43.

Eccles, J. S., and Wang, M. (2012). "Part I commentary: so what is student engagement anyway?," in *Handbook of Research on Student Engagement*, eds S. L. Christenson, A. L. Reschly, and C. Wylie (New York, NY: Springer), 133–145.

Eccles, J. S., and Wigfield, A. (1995). In the mind of the actor: the structure of adolescents' achievement task values and expectancy-related beliefs. *Personal. Soc. Psychol. Bull.* 21, 215–225. doi: 10.1177/0146167295213003

Elliot, A. J., and Harackiewicz, J. M. (1996). Approach and avoidance achievement goals and intrinsic motivation: a mediational analysis. *J. Personal. Soc. Psychol.* 70, 461–475. doi: 10.1037/0022-3514.70.3.461

Elphinstone, B., and Tinker, S. (2017). Use of the motivation and engagement scale-university/college as a means of identifying student typologies. *J. College Student Dev.* 58, 457–462. doi: 10.1353/csd.2017.0034

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Ethington, C. A. (1991). A test of a model of achievement behaviors. *Am. Educ. Res. J.* 28, 155–172. doi: 10.3102/00028312028001155

Farmer, H. S., Wardrop, J. L., and Rotella, S. C. (1999). Antecedent factors differentiating women and men in science/non-science careers. *Psychol. Women Quart.* 23, 763–780. doi: 10.1111/j.1471-6402.1999.tb00396.x

Federation of Australian Scientific and Technological Societies [FASTS] (2002). *Australian Science: Investing in the Future*. Canberra: FASTS.

Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., and Welsh, M. E. (2015). Measuring cost: the forgotten component of expectancy-value theory. *Contem. Educ. Psychol.* 41, 232–244. doi: 10.1016/j.cedpsych.2015.03.002

Frenzel, A. C., Goetz, T., Pekrun, R., and Watt, H. M. G. (2010). Development of mathematics interest in adolescence: influences of gender, family, and school context. *J. Res. Adolesc.* 20, 507–537. doi: 10.1111/j.1532-7795.2010.00645.x

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., et al. (2015). More value through greater differentiation: gender differences in value beliefs about math. *J. Educ. Psychol.* 107, 663–677. doi: 10.1037/edu0000003

Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., and Nagengast, B. (2017). Assessing task values in five subjects during secondary school: measurement structure and mean level differences across grade level, gender, and academic subject. *Contem. Educ. Psychol.* 48, 67–84. doi: 10.1016/j.cedpsych.2016.09.003

Giallo, R., Wood, C. E., Jellett, R., and Porter, R. (2013). Fatigue, wellbeing and parental self-efficacy in mothers of children with an autism spectrum disorder. *Autism* 17, 465–480.

Guo, J., Wang, M. T., Ketonen, E. E., Eccles, J. S., and Salmela-Aro, K. (2018). Joint trajectories of task value in multiple subject domains: from both variable- and pattern-centered perspectives. *Contem. Educ. Psychol.* 55, 139–154. doi: 10.1016/j.cedpsych.2018.10.004

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., and Tauer, J. M. (2008). The role of achievement goals in the development of interest: reciprocal relations between achievement goals, interest, and performance. *J. Educ. Psychol.* 100, 105–122. doi: 10.1037/0022-0663.100.1.105

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equat. Model. Multidiscipl. J.* 6, 1–55. doi: 10.1080/10705519909540118

Hyde, J. S. (2005). The gender similarity hypothesis. *Am. Psychol.* 60, 581–592. doi: 10.1037/0003-066X.60.6.581

Jacobs, J. E., and Simpkins, S. D. (2005). Mapping leaks in the math, science, and technology pipeline. *New Dir. Child Adolesc. Dev*. 2005, 3–6. doi: 10.1002/cd.145

Jiang, Y., Rosenzweig, E., and Gaspard, H. (2018). An expectancy-value-cost approach in predicting adolescent students' academic motivation and achievement. *Contem. Educ. Psychol.* 54, 139–152. doi: 10.1016/j.cedpsych.2018.06.005

Kish, L. (1965). *Survey Sampling*. New York, NY: John Wiley.

Larcombe, W., Tumbaga, L., Malkin, I., and Nicholson, P. (2013). Does an improved experience of law school protect students against depression, anxiety and stress: an empirical study of wellbeing and the law school experience of LLB and JD students. *Syd. Law Rev.* 35, 407–432.

Lawson, M. A., and Lawson, H. A. (2013). New conceptual frameworks for student engagement research, policy, and practice. *Rev. Educ. Res.* 83, 432–479. doi: 10.3102/0034654313480891

Lazarides, R., and Watt, H. M. G. (2015). Girls' and boys' perceived mathematics teacher beliefs, classroom learning environments and mathematical career intentions. *Contem. Educ. Psychol.* 41, 51–61. doi: 10.1016/j.cedpsych.2014.11.005

Lazarides, R., Dietrich, J., and Taskinen, P. H. (2019). Stability and change in students' motivational profiles in mathematics classrooms: the role of perceived teaching. *Teach. Teach. Educ.* 79, 164–175. doi: 10.1016/j.tate.2018.12.016

Lazarides, R., Rubach, C., and Ittel, A. (2016a). Motivational profiles in mathematics: what role do gender, age and parents' valuing of mathematics play? *Int. J. Gender Sci. Technol.* 8, 124–143.

Lazarides, R., Viljaranta, J., Aunola, K., Pesu, L., and Nurmi, J. E. (2016b). The role of parental expectations and students' motivational profiles for educational aspirations. *Learn. Ind. Diff.* 51, 29–36. doi: 10.1016/j.lindif.2016.08.024

Lazarides, R., Viljaranta, J., Aunola, K., and Nurmi, J. E. (2018). Teacher ability evaluation and changes in elementary student profiles of motivation and performance in mathematics. *Learn. Ind. Diff.* 67, 245–258. doi: 10.1016/j.lindif.2018.08.010

LeBreton, J. M., and Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organ. Res. Methods* 11, 815–852.

Linnenbrink-Garcia, L., Wormington, S. V., Snyder, K. E., Riggsbee, J., Perez, T., Ben-Eliyahu, A., et al. (2018). Multiple pathways to success: an examination of integrative motivational profiles among upper elementary and college students. *J. Educ. Psychol.* 110, 1026–1048. doi: 10.1037/edu0000245

Lovibond, S. H., and Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales*, 2nd. Edn. Sydney, NSW: Psychology Foundation.

Luttrell, V. R., Callen, B. W., Allen, C. S., Wood, M. D., Deeds, D. G., and Richard, D. C. S. (2010). The mathematics value inventory for general education students: development and initial validation. *Educ. Psychol. Measure.* 70, 142–160. doi: 10.1177/0013164409344526

Maas, C. J. M., and Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology* 1, 86–92. doi: 10.1027/1614-1881.1.3.86

Maas, C. M. J., and Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Stat. Nederlandica* 58, 127–137.

Magnussen, D. (2000). "The individual as the organizing principle in psychological inquiry: a holistic approach," in *Developmental Science and the Holistic Approach*, eds L. R. Bergman, R. B. Cairns, L. Nilsson, and L. Nystedt (Mahwah, NJ: Erlbaum).

Maltese, A. V., and Tai, R. H. (2011). Pipeline persistence: examining the association of educational experiences with earned degrees in STEM among U.S. students. *Sci. Educ.* 95, 877–907. doi: 10.1002/sce.20441

Masyn, K. E. (2013). "Latent class analysis and finite mixture modeling," in *The Oxford Handbook of Quantitative Methods*, 2nd Edn, Vol. 2, ed. T. D. Little (New York, NY: Oxford University Press), 551–611.

Midgley, C., Kaplan, A., and Middleton, M. (2001). Performance approach goals: good for what, for whom, under what circumstances, and at what cost? *J. Educ. Psychol.* 93, 77–86. doi: 10.1037//0022-0663.93.1.77

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the Patterns of Adaptive Learning Scales*. Ann Arbor, MI: University of Michigan.

National Center for O*NET Development (2016). *O*NET OnLine.* Available at: https://www.onetonline.org/ (accessed May 31, 2018).

Ng, B. L., Liu, W. C., and Wang, J. C. (2016). Student motivation and learning in mathematics and science: a cluster analysis. *Int. J. Sci. Mathemat. Educ.* 14, 1359–1376. doi: 10.1007/s10763-015-9654-1

Ng, C. H. C. (2018). High school students' motivation to earn mathematics: the role of multiple goals. *Int. J. Sci. Mathemat. Educ.* 16, 357–375. doi: 10.1007/s10763-016-9780-4

Nicholls, J. G. (1984). Achievement motivation: conceptions of ability, subjective experience, task choice, and performance. *Psychol. Rev.* 91, 328–346.

NSCRC (2013). *Snapshot Report-Degree Attainment*. Available at: https://nscresearchcenter.org/snapshotreport-degreeattainment3/ (accessed January 9, 2019).

Nylund, K. L., Asparouhov, T., and Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct. Equat. Model.* 14, 535–569. doi: 10.1080/10705510701575396

Office of the Chief Scientist [OCS] (2014a). *Science, Technology, Engineering and Mathematics: Australia's Future*. Canberra: Australian Government.

Office of the Chief Scientist [OCS] (2014b). *Benchmarking Australian Science, Technology, Engineering and Mathematics*. Canberra: Australian Government.

Organisation for Economic Co-operation and Development [OECD] (2004). *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development [OECD] (2006). *Evolution of Student Interest in Science and Technology Studies: Policy Report*. Paris: OECD Global Science Forum.

Pantziara, M., and Philippou, G. N. (2015). Students' motivation in the mathematics classroom. Revealing causes and consequences. *Int. J. Sci. Mathemat. Educ.* 13(Suppl. 2), 385–411. doi: 10.1007/s10763-014-9607-0

Parhiala, P., Torppa, M., Vasalampi, K., Eklund, K., Poikkeus, A. M., and Aro, T. (2018). Profiles of school motivation and emotional well-being among adolescents: associations with math and reading performance. *Learn. Ind. Diff.* 61, 196–204. doi: 10.1016/j.lindif.2017.12.003

Perez, T., Cromley, J. G., and Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *J. Educ. Psychol.* 106, 315–309. doi: 10.1037/a0034027

Perez, T., Wormington, S., Barger, M., Schwartz-Bloom, R., Lee, Y., and Linnenbrink-Garcia, L. (2019). Science expectancy, value, and cost profiles and their proximal and distal relations to undergraduate science, technology, engineering, and math persistence. *Sci. Educ.* 103, 264–286. doi: 10.1002/sce.21490

Pintrich, P. R. (2000). Multiple goals, multiple pathways: the role of goal orientation in learning and achievement. *J. Educ. Psychol.* 92, 544–555. doi: 10.1037/0022-0663.92.3.544

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., et al. (2006). *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf Eines Schuljahres*. Münster: Waxmann Publishing Co.

Raudenbush, S. W., and Bryk, A. S. (2002). *Hierarchical Linear Modeling. Applications and Data Analysis Methods*, 2nd Edn. Thousand Oaks, CA: Sage.

Riegle-Crumb, C., and King, B. (2010). Questioning a White male advantage in STEM: examining disparities in college major by gender and race-ethnicity. *Educ. Res.* 39, 656–664. doi: 10.3102/0013189X10391657

Roeser, R. W., Eccles, J. S., and Freedman-Doan, C. (1999). Academic functioning and mental health in adolescence: patterns, progressions, and routes from childhood. *J. Adolesc. Res.* 14, 135–174. doi: 10.1177/0743558499142002

Roeser, R. W., Strobel, K. R., and Quihuis, G. (2002). Studying early adolescents' academic motivation, social-emotional functioning, and engagement in learning: variable-and person-centered approaches. *Anxiety Stress Coping* 15, 345–368. doi: 10.1080/1061580021000056519

Rost, J. (2006). "Latent-class-analyse [Latent class analyses]," in *Handbuch der Psychologischen Diagnostik [Handbook of Psychological Diagnostics]*, eds F. Petermann and M. Eid (Goettingen: Hogrefe), 275–287.

Rozek, C., Svoboda, R., Harackiewicz, J., Hulleman, C., and Hyde, J. (2017). Utility-value intervention with parents increases students' STEM preparation and career pursuit. *Proc. Natl. Acad. Sci. U.S.A.* 114, 909–914. doi: 10.1073/pnas.1607386114

Salmela-Aro, K., and Read, S. (2017). Study engagement and burnout profiles among Finnish higher education students. *Burnout Res.* 7, 21–28. doi: 10.1016/j.burn.2017.11.001

Schaarschmidt, U., and Fischer, A. (1997). AVEM – ein diagnostisches instrument zur differenzierung von typen gesundheitrelevanten verhaltens und erlebens gegenuber der arbei. [AVEM – An instrument for diagnosing different types of work- and health-related behavior and experience]. *Z. Diff. Diagnost. Psychol.* 18, 151–163.

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.*6, 461–464.

Sells, L. W. (1980). Mathematics: the invisible filter. *Eng. Educ.* 70, 340–341.

Sharp, L., O'Leary, E., Kinnear, H., Gavin, A., and Drummond, F. J. (2016). Cancer-related symptoms predict psychological wellbeing among prostate cancer survivors: results from the PiCTure study. *Psycho-Oncol.* 25, 282–291.

Simpkins, S. D., Fredricks, J. A., and Eccles, J. S. (2012). Charting the Eccles' expectancy-value model from mothers' beliefs in childhood to youths' activities in adolescence. *Dev. Psychol.* 48, 1019–1032. doi: 10.1037/a0027468

Spearman, J., and Watt, H. M. G. (2013). Perception shapes experience: the influence of actual and perceived classroom environment dimensions on girls' motivations for science. *Learn. Environ. Res.* 16, 217–238. doi: 10.1007/s10984-013-9129-7

Stevenson, H. W., and Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading. *Child Dev.* 57, 646–659. doi: 10.2307/1130343

Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychol. Bull.* 135, 859–884. doi: 10.1037/a0017364

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., and Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: a latent interaction modeling study. *J. Educ. Psycho.* 104, 763–777. doi: 10.1037/a0027470

Tuominen-Soini, H., and Salmela-Aro, K. (2014). Schoolwork engagement and burnout among Finnish high school students and young adults: profiles, progressions, and educational outcomes. *Dev. Psychol.* 50, 649–662. doi: 10.1037/a0033898

Tuominen-Soini, H., Salmela-Aro, K., and Niemivirta, M. (2011). Stability and change in achievement goal orientations: a person-centered approach. *Contem. Educ. Psychol.* 36, 82–100. doi: 10.1016/j.cedpsych.2010.08.002

Tuominen-Soini, H., Salmela-Aro, K., and Niemivirta, M. (2012). Achievement goal orientations and academic well-being across the transition to upper secondary education. *Learn. Ind. Diff.* 22, 290–305. doi: 10.1016/j.lindif.2012.01.002

Updegraff, K. A., Eccles, J. S., Barber, B. L., and O'Brien, K. M. (1996). Course enrollment as self-regulatory behavior: who takes optional high school math courses? *Learn. Ind. Diff.* 8, 239–259. doi: 10.1016/S1041-6080(96)90016-3

Viljaranta, J., Nurmi, J.-E., Aunola, K., and Salmela-Aro, K. (2009). The role of task values in adolescents' educational tracks: a person-oriented approach. *J. Res. Adolesc.* 19, 786–798. doi: 10.1111/j.1532-7795.2009.00619.x

Virtanen, T. E., Lerkkanen, M.-K., Poikkeus, A.-M., and Kuorelahti, M. (2018). Student engagement and school burnout in Finnish lower-secondary schools: latent profile analysis. *Scand. J. Educ. Res.* 62, 519–537. doi: 10.1080/00313831.2016.1258669

Watt, H. M. G. (2002). *Gendered Achievement-Related Choices and Behaviours in Mathematics and English: The Nature and Influence of Self-, Task- and Value Perceptions*. Unpublished PhD thesis. University of Sydney, Sydney, NSW.

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th- through 11th-grade Australian students. *Child Dev.* 75, 1556–1574. doi: 10.1111/j.1467-8624.2004.00757.x

Watt, H. M. G. (2006). The role of motivation in gendered educational and occupational trajectories related to maths. *Educ. Res. Eval.* 12, 305–322. doi: 10.1080/13803610600765562

Watt, H. M. G. (2010). "Gender and occupational choice," in *Handbook of Gender Research in Psychology*, eds J. C. Chrisler and D. R. McCreary (New York, NY: Springer), 379–400.

Watt, H. M. G. (2016). "Gender and motivation," in *Handbook of Motivation at School*, 2nd Edn, eds K. Wentzel and D. Miele (New York, NY: Routledge, Taylor & Francis), 320–339.

Watt, H. M. G., Carmichael, C., and Callingham, R. (2017a). Students' engagement profiles in mathematics according to learning environment dimensions: developing an evidence base for best practice in mathematics education. *Sch. Psychol. Int.* 38, 166–183. doi: 10.1177/0143034316688373

Watt, H. M. G., Hyde, J. S., Petersen, J., Morris, Z. A., Rozek, C. S., and Harackiewicz, J. M. (2017b). Mathematics - a critical filter for STEM-related career choices? A longitudinal examination among Australian and U.S. adolescents. *Sex Roles* 77, 254–271. doi: 10.1007/s11199-016-0711-1

Watt, H. M. G., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., and Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: a comparison of samples from Australia, Canada, and the United States. *Dev. Psychol.* 48, 1594–1611. doi: 10.1037/a0027838

Wigfield, A., and Cambria, J. (2010). Students' achievement values, goal orientations, and interest: definitions, development, and relations to achievement outcomes. *Dev. Rev.* 30, 1–35. doi: 10.1016/j.dr.2009.12.001

Wigfield, A., and Eccles, J. (1992). The development of achievement task values: a theoretical analysis. *Dev. Rev.* 12, 265–310. doi: 10.1016/0273-2297(92)90011-P

Wolters, C. A. (2004). Advancing achievement goal theory: using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *J. Educ. Psychol.* 96, 236–250. doi: 10.1037/0022-0663.96.2.236

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 28 June 2019
doi: 10.3389/fpsyg.2019.01472

Check for updates

Gendered Development of Motivational Belief Patterns in Mathematics Across a School Year and Career Plans in Math-Related Fields

Julia Dietrich[1] and Rebecca Lazarides[2]*

[1] Institute of Educational Science, University of Jena, Jena, Germany, [2] Department of Education, University of Potsdam, Potsdam, Germany

OPEN ACCESS

Edited by:
*Ana Lucia Pereira,
Universidade Estadual de Ponta Grossa, Brazil*

Reviewed by:
*Sandra Simpkins,
University of California, Irvine, United States*
*Hanna Gaspard,
University of Tübingen, Germany*

****Correspondence:***
*Julia Dietrich
julia.dietrich@uni-jena.de*

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *13 January 2019*
Accepted: *11 June 2019*
Published: *28 June 2019*

Citation:
Dietrich J and Lazarides R (2019) Gendered Development of Motivational Belief Patterns in Mathematics Across a School Year and Career Plans in Math-Related Fields. Front. Psychol. 10:1472. doi: 10.3389/fpsyg.2019.01472

Rooted in Eccles and colleagues' expectancy-value theory, this study aimed to examine how expectancies and different facets of task value combine to diverse profiles of motivational beliefs, how such complex profiles develop across a school year, and how they relate to gender and career plans. Despite abundant research on the association between gender and motivational beliefs, there is a paucity of knowledge regarding the gendered development of student motivational belief profiles in specific domains. Using latent-transition analysis in a sample of $N = 751$ ninth to tenth graders (55.9% girls), we investigated girls' and boys' development of motivational belief profiles (profile paths) in mathematics across a school year. We further analyzed the association between these profile paths and math-related career plans. The results revealed four motivational belief profiles: high motivation (intrinsic and attainment oriented), balanced above average motivation, average motivation (attainment and cost oriented), and low motivation (cost oriented). Girls were less likely than expected by chance to remain in the high motivation profile, while the opposite was true for boys. The math-relatedness of students' career plans was significantly higher in the "stable high motivation" profile path than in all other stable profile paths.

Keywords: motivation in mathematics, latent transition analysis/latent profile analysis, expectancy-value theory, heterogeneity, adolescence

INTRODUCTION

Motivation declines during adolescence, especially in STEM (science, technology, engineering, and math) subjects. This decline regards, for instance, math interest/intrinsic value (Frenzel et al., 2010; Dietrich et al., 2015), the perceived usefulness of math (Watt, 2004), and perceived competence in math (Watt, 2004). Although girls' and boys' motivation similarly declines across adolescence (Jacobs et al., 2002), girls report lower levels of mathematics interest (Frenzel et al., 2010) and competence beliefs (Watt, 2004) than boys. Such gendered motivational beliefs in math and other STEM fields are related to gendered career plans (Lauermann et al., 2017; Lazarides et al., 2017).

Most existing research that focused on the development of girls' and boys' motivational beliefs and the relations of such beliefs with career plans has been variable-centered. One limitation of this approach is the underlying assumption that the associations between gender, motivational beliefs and career plans are similar across the whole continuum of low to high motivation. Furthermore, it remains unknown how expectancies and the different value facets combine to different patterns of motivational beliefs, how such complex patterns develop over time, and how they relate to gender and career plans. The present study aimed to overcome some of these limitations.

We focus on adolescents' motivational beliefs about mathematics as defined in Eccles and colleagues' expectancy-value theory (EVT) of achievement motivation (Wigfield and Eccles, 2000). The EVT seeks to explain achievement-related (e.g., career) choices through expectancies for success and subjective task values. In this study, we operationalized the expectancy component as a student's academic self-concept, defined as the subjective beliefs about one's abilities in mathematics (Marsh and Martin, 2011). In line with EVT, the value component is differentiated into interest (intrinsic value), usefulness for future goals (utility value), personal importance (attainment value) and perceived costs of learning mathematics (cost value) (Wigfield and Eccles, 2000). Expectancies and values are conceptualized as correlated but independently functioning in predicting achievement-related behaviors. Moreover, the theory predicts variation in the relative salience of different value facets (Eccles, 2005).

Indeed, research suggests that the relations between different motivational beliefs are heterogeneous in the student population. Corresponding with findings from the correlational literature, many classification-based studies report patterns of overall high, moderate, and low motivation (Viljaranta et al., 2017; Lazarides et al., 2019). Going further, patterns of mixed motivation have been found, such as high self-concept combined with low interest (Lazarides et al., 2019), and high interest combined with high perceived cost of doing well in math (Conley, 2012).

Only few studies have addressed changes in motivational profiles using longitudinal data (in the following labeled "profile paths," e.g., Marcoulides et al., 2008; Lazarides et al., 2019). These studies suggest that patterns of motivation are relatively stable in adolescence with probabilities around 0.70–0.90 of staying in the same profile. But that also means that some adolescents do exhibit profile changes, even during a very short time span (e.g., Martinent and Decret, 2015). Building on these previous studies we adopted a longitudinal classification-based research strategy. This allows to discern the extent to which the development of students with low motivation profiles differs from the development of students with high motivation profiles and from that of students with mixed motivation profiles. Our first two research questions were:

(1) Which kinds of expectancy-value patterns or profiles of math motivation can be identified at two time points in the school year? We expected to find high, moderate, low, and mixed motivation profiles.
(2) How many students change their motivation profile across the school year? We hypothesized profile stability to be typical with only some students changing their profile of math motivation.

Abundant research has shown that gendered pathways into and away from STEM fields are mediated through motivational beliefs (Eccles and Wang, 2015; Lauermann et al., 2017). Even in the case of equal achievements in mathematics, girls find mathematics less interesting (Frenzel et al., 2010), perceive lower job utility of mathematics (Gaspard et al., 2015), and feel less competent compared to similarly achieving boys (Marsh and Yeung, 1998). The few studies investigating gendered patterns of motivation in STEM fields showed that boys dominated the high-math-and-science profiles and the group with high mathematics self-concept, but low interest, while girls dominated the low-math-and-science profiles (Chow and Salmela-Aro, 2011; Lazarides et al., 2019). Studies are largely missing that would investigate *gendered changes in patterns* (i.e., longitudinal profile paths) of motivational beliefs and relations between such profile paths and plans for math-intensive careers. Such studies would add more detail to the associations of motivation with gender and with career plans. In adolescence, girls and boys face developmental changes such as a greater importance of gender identity development and closer peer relations (Erikson, 1959). Interests in school subjects are a means to communicate self-identity (Kessels, 2005), which might be particularly relevant to motivation in school subjects which are typically stereotyped as "male," such as math and physics (Nosek et al., 2002). Research has shown that physics-oriented girls are less popular than are their peers, because they behave in opposition to the female stereotype (Kessels, 2005). Also in math, girls might feel pressured to behave in gender-typical ways and face a greater likelihood than boys of changing their motivational profile, especially when they are highly motivated.

Our third and fourth research questions were:

(3) In which longitudinal profile paths (stable profile paths and profile changes) are gender disparities especially prevalent? We expected girls (boys) to be under(over)represented in stable high motivation paths and over(under)represented in stable low motivation paths. We moreover expected a greater likelihood for girls to change their motivation profile.
(4) Which longitudinal motivation paths are most strongly related to plans for math-intensive careers? We expected that the career plans of students with stable high motivation would evidence the highest mathematics-relatedness, compared to other motivation paths.

METHOD

Data stem from two waves of the German longitudinal Motivation for Learning Mathematics study (Lazarides and Rubach, (2015-2017)). The Berlin Senate for Education, Youth, and Research approved the study. An ethics approval was not required at the time the study was conducted as per the then applicable institutional and national guidelines and regulations. The participating students and their parents gave written informed consent. For the present analyses, data from 751 ninth to tenth graders (55.9% girls; 71.2% native speaker) in two school types were used (academic track schools, "Gymnasium": 53.8%; integrated secondary school, "Integrierte Sekundarschule": 46.2%). Data were assessed 2 months after the beginning of the school year (Time 1) and again 6 months later (Time 2). **Supplementary Appendix A** shows descriptive statistics for all variables.

Mathematics task values were assessed with four subscales: Intrinsic value (e.g., "I like doing math"), utility value (e.g., "Math content will help me in my life"), and attainment value (e.g., "It is important to me to be good at math") were assessed with three items each (adapted from Steinmayr and Spinath (2010), 1 = *does not apply at all* - 5 = *fully applies*). Cost value was also assessed with a three-item scale comprising effort cost, emotional cost, and opportunity (e.g., "Doing math is exhausting to me", effort cost) based on Gaspard et al. (2015). Cronbach's alpha reliabilities at T1/T2 were 0.92/0.92 for intrinsic value, 0.90/0.92 for attainment value, 0.87/0.89 for utility value, and 0.79/0.78 for cost value.

Self-concept in mathematics was assessed with four items (e.g., "I think I am … in mathematics" from "1 = *not talented* - 5 = *very talented;* (Steinmayr and Spinath, 2010)). Reliabilities at T1/T2 were 0.87/0.88.

Mathematics-related career plans were assessed with the item "What job would you like to have in the future?" Students' open-ended answers were coded for the mathematics-relatedness of the nominated career using the Occupational Information Network (O*NET; National Center for O*NET Development, 2014) to quantify relatedness to "knowledge of arithmetic, algebra, geometry, calculus, statistics, and their applications" on a scale ranging from 0 = *not mathematics-related* to 100 = *completely mathematics-related*.

RESULTS

We used Latent Transition Analysis (LTA; Nylund-Gibson et al., 2014) to examine research questions 1 and 2. Means and variances of the motivation variables (profile indicators) were allowed to vary across latent classes. We imposed measurement invariance with equal means and variances in a given latent class over time. We statistically evaluated the appropriate number of latent classes based on information criteria (BIC, aBIC, AIC). Theoretical interpretation and the number of cases per class were also used for model selection (Berlin et al., 2014).

We identified the four-class model (**Figure 1**) as the best fitting LTA model (see **Supplementary Appendix B**). Pattern 1 was characterized by particularly high intrinsic and attainment value and low perceived cost and was labeled as "*high motivation* (*intrinsic and attainment oriented*)" pattern. Pattern 2, labeled "*balanced above average*" motivation, was similar to pattern 1, but the levels of motivational beliefs were one scale point closer to the mid of the scale. Pattern 3 was characterized by motivation levels around the scale midpoint, with highest levels on attainment and cost values, and was labeled "*average motivation* (*attainment and cost oriented*)." Pattern 4 was characterized by low intrinsic, attainment and utility values, low self-concept, and high costs and was labeled "*low motivation* (*cost oriented*)" pattern. As hypothesized, profiles were highly stable with 77% of students remaining in their motivational pattern across the school year. The most frequent change was from "average" to "balanced above average" motivation (7% of students). Other changes are depicted in **Figure 1**, and the transition probabilities are depicted in **Supplementary Appendix C**.

For further analyses involving gender and career plans (research questions 3 and 4) we saved each individual's most likely latent class as a manifest variable (the classification probabilities ranged 0.70 to 0.97, M = 0.81). In doing so, we used the latent class patterns which refer to each individual's longitudinal profile path (e.g., stable high motivation or moving from "average" to "low" motivation).

We used Configural Frequency Analysis (Heine and Alexandrowicz, 2015) to examine the association between gender and the longitudinal profile paths. Based on a cross-tabulation of profile path by gender, ConFA provides a test for each cell indicating whether this cell contains more or fewer individuals than expected by chance. We used the z-Test with Bonferroni correction (alpha = 0.002) for the cell-specific significance tests. There was a significant overall association between gender and profile path (**Table 1**). Cell-specific tests indicated that girls

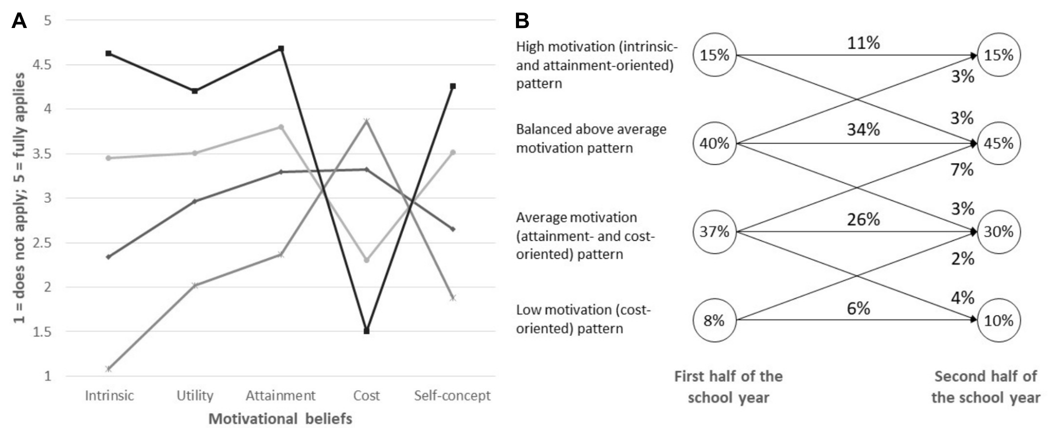


FIGURE 1 | (A) Means across all motivational beliefs for the four-profile solution. **(B)** Profile sizes (proportions for the latent class variables: circles) and profile changes (proportions for the latent class patterns: arrows) based on the estimated LTA model. Circles depict the percentage of the sample in a given profile in the first/the second half of the school year. Arrows depict the percentage of the sample that shifts between the two profiles linked by each arrow.

TABLE 1 | Cell-Wise associations between longitudinal profile paths and student gender (ConFa).

Profile path		Observed	Expected	z	P
Girls					
High	High	28.00	48.35	−2.93	0.002
	Balanced	12.00	14.62	−0.68	0.247
	Average	0.00	0.00	0.00	0.500
	Low	0.00	0.00	0.00	0.500
Balanced	High	8.00	11.24	−0.97	0.167
	Balanced	138.00	150.12	−0.99	0.161
	Average	9.00	7.31	0.63	0.266
	Low	0.00	1.12	−1.06	0.144
Average	High	3.00	1.69	1.01	0.156
	Balanced	36.00	28.11	1.49	0.068
	Average	133.00	109.64	2.23	0.013
	Low	20.00	17.43	0.62	0.269
Low	High	0.00	0.00	0.00	0.500
	Balanced	1.00	1.12	−0.12	0.453
	Average	8.00	4.50	1.65	0.049
	Low	24.00	24.74	−0.15	0.441
Boys					
High	High	58.00	37.65	3.32	0.000
	Balanced	14.00	11.38	0.78	0.219
	Average	0.00	0.00	0.00	0.500
	Low	0.00	0.00	0.00	0.500
Balanced	High	12.00	8.76	1.10	0.136
	Balanced	129.00	116.88	1.12	0.131
	Average	4.00	5.69	−0.71	0.239
	Low	2.00	0.88	1.20	0.115
Average	High	0.00	1.31	−1.15	0.126
	Balanced	14.00	21.89	−1.69	0.046
	Average	62.00	85.36	−2.53	0.006
	Low	11.00	13.57	−0.70	0.243
Low	High	0.00	0.00	0.00	0.500
	Balanced	1.00	0.88	0.13	0.447
	Average	0.00	3.50	−1.87	0.031
	Low	20.00	19.26	0.17	0.433

Note. Overall test of association between gender and profile path: χ^2 (15) = 54.41, $p < 0.001$. Bonferroni corrected alpha for the cell-specific z-tests = 0.002.

were less likely than expected by chance to remain in the 'high motivation profile' across the school year. The opposite was true for boys. Additionally, we tested whether girls were more likely to change their profile (change: 97 girls, 62 boys; stability: 323 girls, 269 boys), but that was not the case, χ^2 (1) = 2.11, $p = 0.146$.

In examining the relation between profile paths and career plans we focused on the stable paths (high – high, n = 86; balanced – balanced, n = 267; average – average, n = 195; low – low, n = 44) due to the relatively large sample sizes in these profile paths compared to profile changes (<1 to 7% of students). We conducted ANOVAs to test for differences in math-related career plans between the four profile path groups, and found significant overall effects both in the first half [F(3,400) = 11.19, $p < 0.001$] and in the second half of the school year [F(3,389) = 11.23, $p < 0.001$]. *Post hoc* comparisons (Tukey HSD test, **Supplementary Appendix D**) indicated that the math-relatedness of students' career plans was significantly higher in the "stable high motivation" path than in all other profile paths. The math-relatedness of the career plans in students with "stable low motivation" did not differ from the math-relatedness in students with "stable balanced" or "stable average" motivation.

DISCUSSION

This study aimed to look beyond main-effects variable-(correlation)-centered models to study the interrelations of gender, career plans, and change and stability in profiles of motivational beliefs. We found profiles reflecting mainly level-differences which correspond to those shown in previous studies (Viljaranta et al., 2017; Lazarides et al., 2019). While these results converge with a variable-centered perspective on motivational beliefs, they additionally show that, for example, students with low motivation are especially low on intrinsic value, compared to their utility and attainment values (see also Eccles, 2005). Contrary to our expectations and to some previous studies (e.g., Conley, 2012) we did not find "mixed motivation" profiles. It might be that such profiles are more prevalent among primary school (Viljaranta et al., 2017) and younger secondary school students (Conley, 2012) than among ninth and tenth graders (Lazarides et al., 2019).

Going beyond previous studies, we were able to show some gender disparities in profile development. Boys were more likely to remain in the "high motivation (intrinsic and attainment oriented)" profile across the school year, while such a stable high motivation path was untypical for girls. These results are relevant as this stable high motivation path was in turn associated with the highest levels of math career plans, with differences in career plans getting even larger across the school year (Hedges' g increasing by 0.12/0.15 for the differences between stable high and stable average/stable balanced motivation). Interestingly, gender differences were not evident in the stable low profile path.

Some of our results are in contrast to a previous study of Lazarides et al. (2019) who found no gender differences for stable high motivation from Year 9 to Year 10, but did find differences for stable low motivation such that girls were more likely to remain in that profile compared to boys. Other studies on the development of motivational profiles in adolescents (i.e., Alexander and Murphy, 1998; Nurmi and Aunola, 2005; Viljaranta et al., 2016) did not consider the role of student gender for such developmental changes. There is thus a need for more studies on these developmental aspects from a holistic, person-oriented perspective, and the special value of the present study is to help understand how gender differences in motivation evolve in different groups of students.

Important limitations of this study pertain to its reliance on self-report data (e.g., no achievement data were assessed) and the short time span studied. It might be that because stability was so high in this study, we did not find more frequent profile changes among girls as expected.

Overall, a research focus on the development of motivational profiles is worthwhile to capture the heterogeneity within and between students: it is very unlikely that every person develops in the same way (Molenaar, 2004). Accordingly, the results of this study suggest that the associations between (the development of) motivational beliefs, gender, and career plans vary across different levels and patterns of motivation.

ETHICS STATEMENT

The Berlin Senate for Education, Youth, and Research approved the study. An ethics approval was not required at the time the study was conducted as per the then applicable institutional and national guidelines and regulations. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

Both authors analyzed the data from a research project (MOVE) designed and conducted by RL and jointly contributed to the writing of the article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01472/full#supplementary-material

REFERENCES

Alexander, P. A., and Murphy, P. K. (1998). Profiling the differences in students' knowledge, interest, and strategic processing. *J. Educ. Psychol.* 90, 435–447. doi: 10.1037/0022-0663.90.3.435

Berlin, K. S., Williams, N. A., and Parra, G. R. (2014). An introduction to latent variable mixture modeling (part 1): overview and cross-sectional latent class and latent profile analyses. *J. Pediatr. Psychol.* 39, 174–187. doi: 10.1093/jpepsy/jst084

Chow, A., and Salmela-Aro, K. (2011). Task-values across subject domains: a gender comparison using a person-centered approach. *Int. J. Behav. Dev.* 35, 202–209. doi: 10.1177/0165025411398184

Conley, A. (2012). Patterns of motivation beliefs: combining achievement goal and expectancy-value perspectives. *J. Educ. Psychol.* 104, 32–47. doi: 10.1037/a0026042

Dietrich, J., Dicke, A.-L., Kracke, B., and Noack, P. (2015). Teacher support and its influence on students' intrinsic value and effort: dimensional comparison effects across subjects. *Learn. Inst.* 39, 45–54. doi: 10.1016/j.learninstruc.2015.05.007

Eccles, J. S. (2005). "Subjective task value and the Eccles et al. model of achievement-related choices," in *Handbook of Competence and Motivation*, eds A. J. Elliot and C. S. Dweck (New York, NY: Guilford Press), 105–131.

Eccles, J. S., and Wang, M.-T. (2015). What motivates females and males to pursue careers in mathematics and science? *Int. J. Behav. Dev.* 40, 100–106. doi: 10.1177/0165025415616201

Erikson, E. H. (1959). *Identity and the Life Cycle*. New York, NY: International Universities Press.

Frenzel, A. C., Goetz, T., Pekrun, R., and Watt, H. M. G. (2010). Development of mathematics interest in adolescence: influences of gender, family, and school context. *J. Res. Adoles.* 20, 507–537. doi: 10.1111/j.1532-7795.2010.00645.x

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., et al. (2015). More value through greater differentiation: gender differences in value beliefs about math. *J. Educ. Psychol.* 107, 663–677. doi: 10.1037/edu0000003

Heine, J. H., and Alexandrowicz, R. W. (2015). *Confreq: Configural Frequencies Analysis Using Log-linear Modeling R package version 1.4*. Available at: https://cran.r-project.org

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., and Wigfield, A. (2002). Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Dev.* 73, 509–527. doi: 10.1111/1467-8624.00421

Kessels, U. (2005). Fitting into the stereotype: how gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *Eur. J. Psychol. Educ.* 20, 309–323. doi: 10.1007/BF03173559

Lauermann, F., Tsai, Y.-M., and Eccles, J. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy–value theory of achievement-related behaviors. *Dev. Psychol.* 53, 1540–1559. doi: 10.1037/dev0000367

Lazarides, R., Dietrich, J., and Taskinen, P. H. (2019). Stability and change in students' motivational profiles in mathematics classrooms: the role of perceived teaching. *Teach. Teach. Educ.* 79, 164–175. doi: 10.1016/j.tate.2018.12.016

Lazarides, R., Rubach, C., and Ittel, A. (2017). Adolescents' perceptions of socializers' beliefs, career-related conversations, and motivation in mathematics. *Dev. Psychol.* 53, 525–539. doi: 10.1037/dev0000270

Lazarides, R., and Rubach, C. (2015-2017). *Motivation and Learning in Mathematics Study*. Potsdam: Retrieved from Unpublished project report.

Marcoulides, G. A., Gottfried, A. E., Gottfried, A. W., and Oliver, P. H. (2008). A latent transition analysis of academic intrinsic motivation from childhood through adolescence. *Educ. Res. Eval.* 14, 411–427. doi: 10.1080/13803610802337665

Marsh, H. W., and Martin, A. J. (2011). Academic self-concept and academic achievement: relations and causal ordering. *Br. J. Educ. Psychol.* 81, 59–77. doi: 10.1348/000709910X503501

Marsh, H. W., and Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: gender differences in the development of math and english constructs. *Am. Educ. Res. J.* 35, 705–738. doi: 10.3102/00028312035004705

Martinent, G., and Decret, J.-C. (2015). Motivational profiles among young table-tennis players in intensive training settings: a latent profile transition analysis. *J. Appl. Sport Psychol.* 27, 268–287, doi: 10.1080/10413200.2014.993485

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement* 2, 201–218. doi: 10.1207/s15366359mea0204_1

Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *J. Pers. Soc. Psychol.* 83, 44–59. doi: 10.1037/0022-3514.83.1.44

Nurmi, J.-E., and Aunola, K. (2005). Task-motivation during the first school years: a person-oriented approach to longitudinal data. *Learn. Inst.* 15, 103–122. doi: 10.1016/j.learninstruc.2005.04.009

Nylund-Gibson, K., Grimm, R., Quirk, M., and Furlong, M. (2014). A latent transition mixture model using the three-step specification. *Struct. Equ. Model. Multidis. J.* 21, 1–16. doi: 10.1080/10705511.2014.915375

Steinmayr, R., and Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW) [Construction and validation of a scale for the assessment of subjective task values (SESSW)]. *Diagnostica* 56, 195–211. doi: 10.1026/0012-1924/a000023

Viljaranta, J., Aunola, K., and Hirvonen, R. (2016). Motivation and academic performance among first-graders: a person-oriented approach. *Learn. Individ. Differ.* 49, 366–372. doi: 10.1016/j.lindif.2016.06.002

Viljaranta, J., Kiuru, N., Lerkkanen, M.-K., Silinskas, G., Poikkeus, A.-M., and Nurmi, J.-E. (2017). Patterns of word reading skill, interest and self-concept of ability. *Educ. Psychol.* 37, 712–732. doi: 10.1080/01443410.2016.1165798

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th-through 11th-grade Australian students. *Child Dev.* 75, 1556–1574. doi: 10.1111/j.1467-8624.2004.00757.x

Wigfield, A., and Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps1999.1015

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 02 July 2019
doi: 10.3389/fpsyg.2019.01449

Check for updates

The Relations of Science Task Values, Self-Concept of Ability, and STEM Aspirations Among Finnish Students From First to Second Grade

Janica Vinni-Laakso[1]*, Jiesi Guo[2], Kalle Juuti[1], Anni Loukomies[3,4], Jari Lavonen[1,4] and Katariina Salmela-Aro[1,5]

[1] Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, [2] Institute for Positive Psychology and Education, Australian Catholic University, North Sydney, NSW, Australia, [3] Viikki Teacher Training School, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland, [4] Department of Childhood Education and Centre for Education Practice Research, University of Johannesburg, Soweto, South Africa, [5] Collegium Helveticum – ETH Zürich, Zurich, Switzerland

OPEN ACCESS

Edited by:
Silke Luttenberger,
University College of Teacher Education Styria, Austria

Reviewed by:
Emily Quinn Rosenzweig,
University of Wisconsin–Madison, United States
Penelope Winifred St John Watson,
The University of Auckland, New Zealand

***Correspondence:**
Janica Vinni-Laakso
janica.vinni-laakso@helsinki.fi

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: 15 January 2019
Accepted: 06 June 2019
Published: 02 July 2019

Citation:
Vinni-Laakso J, Guo J, Juuti K, Loukomies A, Lavonen J and Salmela-Aro K (2019) The Relations of Science Task Values, Self-Concept of Ability, and STEM Aspirations Among Finnish Students From First to Second Grade. Front. Psychol. 10:1449. doi: 10.3389/fpsyg.2019.01449

According to modern expectancy-value theory, students' motivation in school subjects begins to vary at the very beginning of their school careers, showing a task-specific pattern of motivation. However, there is no clear evidence in the literature on how students' value beliefs are formed and interact with each other in early elementary schools. Using the longitudinal structural equation modeling, this study examined relations between science-related task values (i.e., intrinsic value and cost), self-concept of ability, and future occupational aspirations based on first graders and 1-year follow-up from seven schools in Helsinki (N = 332; ages = 7 and 8 years; girls = 51%). Results showed that the students who had a high science-related self-concept of ability and intrinsic value tended to perceive low cost of science learning. Science-related self-concept of ability was the most stable construct, while in intrinsic value and cost, there were significant levels of fluctuation across the first and second grades. A high science-related self-concept of ability in the first grade predicted a lower cost value in the second grade, and a high science-related intrinsic value was a marginally significant predictor of future occupational aspirations in science, technology, engineering, and mathematics (STEM). Mean-level differences revealed that the girls' science-related self-concept of ability, intrinsic value, and cost remained the same in both grades, while the boys' self-concept of ability decreased. The girls' mean levels in science-related intrinsic value were higher than those of the boys, while students' self-concept of ability and cost were similar across gender in both grades. A cross-lagged panel model revealed that the girls reported more STEM occupational aspirations than the boys in the second grade, while controlling for the motivational beliefs. In summary, the results indicate that a high-level of science interest in young students predicts STEM occupational aspirations; high girls' intrinsic value in early science education does not steer them away from STEM occupations; boys' task motivation might be at greater risk of decline during early science education.

Keywords: expectancy-value theory, intrinsic value, cost, self-concept of ability, STEM occupational aspirations, gender differences, elementary students

INTRODUCTION

During the last decade, increased attention has been given to low student interest and engagement in science learning and science-related careers. For example, Horizon 2020, a European Union Research and Innovation program, emphasized the need for a science education supporting students in developing positive attitudes toward science and nurturing their curiosity and cognitive resources (Ryan, 2015). Research on students' science motivation and adolescents' educational and occupational aspirations has also increased in recent years [see Tytler (2014) for a review]. Trends consistently show a decrease over time in students' positive attitudes and motivation toward science and their pursuit of science-related careers. Recent international large-scale assessments, such as the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMMS), have showed that Finnish students' science-related achievement and motivation have been declining (Martin et al., 2016; OECD, 2016). According to the PISA 2015 report (OECD, 2016), Finnish ninth graders' interest in and enjoyment of science, technology, engineering, and mathematics (STEM) are below the average for OECD countries, and their occupational expectations in the STEM field are among the lowest in OECD countries (OECD, 2016).

Gendered pathways into STEM have also received growing attention in the literature [see Watt (2016) for a review]. Women are clearly underrepresented in math-intensive science education and STEM careers (OECD, 2015), stimulating research into possible explanations. Some findings suggest that stereotypes might play a role in students' future occupational aspirations and that women have not stereotypically been seen as scientists (Chambers et al., 2018; Miller et al., 2018). Female students' self-concept of ability in science is lower than that of male students, despite the fact that girls often outperform boys in school science (Watt, 2016). There are also gender differences in attitudes to science among Finnish students, with boys showing higher enjoyment and girls holding higher instrumental value in science learning (OECD, 2016). Still, the gender gap favoring boys in bachelor's degrees obtained in science remains significant in Finland (OECD, 2015). High academic motivation in science has been positively linked to deeper engagement, persistent learning, better knowledge acquisition, and higher aspirations in that domain, which prepares individuals to pursue further education and careers in STEM fields [see Wang and Degol (2013) for a review].

Until now, studies of science motivation and STEM aspirations and their development have mainly focused on middle and high school students. In order to form a more comprehensive picture of the phenomenon, and to understand middle and high school students' science-related motivation and aspirations regarding STEM, we need to examine the point at which decline in students' science motivation begins to emerge, how stable their motivational beliefs are, and what factors relate to this development. Thus, in order to investigate the trajectories of STEM motivation from the beginning of elementary students' school careers, we need to address their science-related motivation. In addition, most previous studies of young students' STEM motivation have been conducted in the United States; therefore, more research is needed in other educational contexts.

Task Motivation in Early Science Learning

In this study, we draw on the modern expectancy-value theory (EVT) framework (Eccles, 2009). EVT (Eccles et al., 1983; Eccles, 2009) posits that achievement-related performance and choices are most directly influenced by students' expectations of success on achievement-related tasks and their subjective assessments of the relative value of different achievement-related tasks. Eccles' EVT of achievement-related choices is a major theoretical framework for studying achievement motivation. It has been widely used to tackle both individual and gender differences in educational and career choices [see Wang and Degol (2013, 2017) and Watt (2016) for reviews]. Within the academic domain, Eccles et al. (1983) operationally defined *expectancies of success* as children's beliefs of how well they will do on an upcoming task. Expectancies of success are children's evaluations of their current abilities as well as how they think they compare to other students (Wigfield et al., 2016) in a given task. Thus, we use the term task-specific self-concept of ability. Wigfield and Eccles (1992) and Eccles and Wigfield (2002) also distinguished between multiple components of subjective task values: *intrinsic value* (enjoyment or liking), *utility value* (the usefulness of a task for helping to fulfill personal goals), *attainment value* (relevance of a task to one's sense of self, identity, and core personal values), and *costs* (perceived negative aspects of making a specific choice).

Intrinsic value refers to the extent to which an individual gains enjoyment from performing an activity (Eccles, 2009, 2011). *Cost* refers to the things that students perceive they are investing or giving up in order to engage in a task (Flake et al., 2015), including the degree of potential loss of time; effort demands; the loss of valued alternatives, such as spending time with friends; and additional negative experiences, such as stress. According to EVT (Eccles et al., 1983; Eccles, 2009), people are most likely to select those tasks for which they hold the highest expectations for success and the highest levels of subjective task value.

Students' achievement-related beliefs and attitudes play an important role in academic environments by directing their behavior and effort in learning situations (Eccles, 2009; Marsh and Martin, 2011). Students who have a positive self-concept of ability and intrinsic value in specific academic subjects are likely to perform better and be more engaged in school than those who have a less positive self-concept of ability in a given subject. Previous studies have shown that students hold high beliefs in their abilities at the beginning of elementary school and that they are highly optimistic about their competences in different areas and domains (Stipek and Mac Iver, 1989; Wigfield et al., 2016). The trajectories of students' self-concept of ability decline from the elementary to the middle and high school years, although some domain-specific features can be identified in these emerging trajectories (Jacobs et al., 2002).

Studies have provided several reasons for the decline in students' self-concept of ability, including children's developmental changes, increasing social comparisons among students, and environmental changes in the school context, for example, increasing numbers of subjects and teachers and greater emphasis on grades (Stipek and Mac Iver, 1989; Marsh et al., 1995; Harter, 2012).

Intrinsic motivation is well known to be associated with ability beliefs, i.e., those with a higher self-concept of ability are more willing to engage in learning processes and enjoy learning. Thus, it is no wonder that, upon entering school, elementary students have high academic interest (e.g., Viljaranta et al., 2016). However, this interest soon starts to decline across domains (Wigfield et al., 1997; Gottfried et al., 2001; Jacobs et al., 2002). Furthermore, previous EVT studies have tended to show moderate to large correlations between academic self-concept of ability and the value components [see Wigfield and Eccles (2002) and Wigfield et al. (2009, 2016) for reviews]. Self-concept of ability is more highly correlated with intrinsic value than with the other value components within a specific domain (Wigfield et al., 2009, 2016), and the positive correlation between task motivation and self-concept of ability has been found to strengthen with age (Wigfield et al., 1997; Fredricks and Eccles, 2002; Jacobs et al., 2002).

A general downward shift in students' interest from elementary to middle school is also evident in the science domain (Gottfried et al., 2001). In general, academic interest is rather stable, and this stability increases with age across subject domains (Gottfried et al., 2001). This kind of development poses a challenge for children who begin their school careers with low science motivation. The trend of declining motivation affects students' achievement and increases the risk of dropout from science education later on and, as a consequence, from STEM careers.

In addition to low interest, students might also experience costs that affect their motivation as a negative valence of a task (Barron and Hulleman, 2015). Cost is the least studied component of EVT, and it is distinct from other components of the EVT model (see Flake et al., 2015). In addition, students have been found to report different types of cost, some of which include the requirement of too much effort, emotional/psychological demands, and loss of other valuable opportunities. It has been shown that perceived cost can detract students from engaging in a task or activity and that it can be a powerful predictor of career and education-related outcomes in middle school, high school, and college students (e.g., Battle and Wigfield, 2003; Wigfield and Cambria, 2010; Perez et al., 2014; Flake et al., 2015; Jiang et al., 2018). However, research on elementary students' EVT motivation including perceived cost is lacking.

To sum up, in general, self-concept of ability and intrinsic motivation decrease as students proceed from the first grade onward. Alongside this development, interest may play a role in the formation of self-concept of ability. However, little is known about elementary students' experience of cost, particularly in the context of science. First graders would probably engage in a task that is interesting and fun rather than a task that they evaluate as useful for their future or important for their personal selves (Wigfield and Eccles, 1992; Eccles and Wigfield, 2001). In a similar vein, students would most probably disengage from tasks they perceive as overly demanding or emotionally exhausting. In this study, our primary aim is to explore the stability and associations among first graders' science-related self-concept of ability, intrinsic value, and perceived cost during a 1-year follow-up.

Task Motivation and Academic Aspirations

Research based on EVT has demonstrated that self-concept of ability and value beliefs represent the most proximal precursors of academic achievement, effort, school engagement, and educational aspirations (e.g., Marsh et al., 1995; Eccles, 2009; Watt et al., 2012; Wang et al., 2013; Guo et al., 2015a). Several studies showed that interest in a certain subject domain is related to academic achievement in that domain (e.g., Harackiewicz and Hulleman, 2010; Guo et al., 2015b, 2017). Positive associations of intrinsic value and self-concept of ability with achievement have also been found among elementary school students (Denissen et al., 2007; Viljaranta et al., 2014) and in the context of science (Guo et al., 2018b). According to EVT, the cost component of task values is assumed to dampen students' motivation, and it is strongly and negatively related to expectancy and moderately and negatively related to value, long-term interest, course grades, and overall motivation (Flake et al., 2015). Thus, it is important to differentiate and consider cost, self-concept of ability, and intrinsic value among elementary school students to further disentangle the relationships between self-concept of ability and positive and negative value beliefs in achievement-related outcomes of future occupational aspirations.

The recent study on elementary students' career aspirations demonstrates that children's aspirations are shaped from a young age (Chambers et al., 2018). Students' attitudes shape their interests and later behavior. Thus, student motivation determines the choices students make about their educational pathways. Highly motivated students are more likely to choose courses and aspire to careers that correspond with the subjects in which they are motivated (Simpkins et al., 2006; Chow et al., 2012; Wang, 2012). Students' intrinsic value and academic self-concept in mathematics and science have been found to predict their STEM aspirations in middle and high school (Wang and Degol, 2013; Guo et al., 2015b). Moreover, longitudinal tracking has showed that students who do not express STEM-related aspirations at the age of 10 years are unlikely to develop STEM aspirations by the age of 14 years, and consequently are less likely to pursue science subjects (Archer et al., 2013). In contrast, perceived cost is a negative predictor of interest and performance outcomes (Perez et al., 2014; Barron and Hulleman, 2015; Flake et al., 2015; Jiang et al., 2018). These findings further underline the importance of student motivation for long-term academic and career success. However, whether these motivational beliefs are related to elementary students' future occupational aspirations has not been tested. In fact, we know very little about the factors that influence early career aspirations, despite the fundamental role of aspirations in individuals' career choices and development throughout the lifespan.

Of particular relevance is that no previous study has integrated science-related self-concept of ability, intrinsic value, and cost to determine the extent to which these beliefs and emotions relate to future occupational aspirations among early elementary students. Thus, we have chosen to focus on self-concept of ability and positive and negative aspects of task value – namely, intrinsic value and cost – to examine the associations of these constructs among first and second graders in science learning, and how these constructs are related to students' occupational aspirations.

Gendered Science Motivation and Occupational Aspirations

Previous findings show that students' self-concept of ability and intrinsic value become gendered, especially in relation to mathematics and literacy (see Eccles et al., 1993; Wigfield et al., 1997; Jacobs et al., 2002). Girls' self-concept of ability in mathematics is found to be lower than that of boys, but girls show a lower decline over time (Fredricks and Eccles, 2002), indicating that the gender gap decreases over time. Meanwhile, boys' self-concept of ability in language and arts is lower and declines more than that of girls (Jacobs et al., 2002). According to a recent meta-analysis (Miller et al., 2018), the last five decades have witnessed a developmental change in children's gender-science stereotypes. In a draw-a-scientist study, children in the 1960s almost exclusively depicted scientists as males; in 2000, significantly more children depicted female scientists than their 1960s counterparts. In another study, Bian et al. (2017) observed that not only gender-science stereotypes were still prevalent, but they also started to emerge early. Bian et al. (2017) also found that children perceived males as more intellectual than females, which had a clear influence on their interests in selecting tasks that are described to be easy or difficult, even at the age of 6 years. However, no gender differences in science-related self-concept were found in preschool and early elementary school-aged children (Leibham et al., 2013).

Boys have been shown to hold higher interest in science in early education, although high science interest in preschool predicted higher self-concept and achievement for 8-year-old girls (Leibham et al., 2013). Particularly in the case of the early elementary school years, it appears that interest helps build a higher self-concept of ability. Various studies have found that, as early as elementary school, boys hold higher intrinsic values in mathematics, while girls hold higher intrinsic values in language (Eccles et al., 1983; Jacobs et al., 2002). These gendered value beliefs also feature among secondary school students (e.g., Gaspard et al., 2015).

As noted, cost is salient in student motivation and is linked to several educational outcomes (Flake et al., 2015). Watt (2016) investigated adolescents' gender differences in science and found different types of cost to be differently gendered. For example, girls experienced greater psychological cost (e.g., "It frightens me that math/science courses are harder than other courses"), while boys experienced more social cost (e.g., "I'm concerned that working hard in math/science classes might mean that I lose some of my close friends") in their science learning in Grade 10. In terms of effort-related cost, no gender differences were found. In this study, we are interested in studying whether the experience of cost emerges from the first school years and whether this experience is gendered. This might provide additional insights as to why girls and boys end up valuing different subjects and choosing different career paths, in spite of their equal competences.

Regardless of the predictive power of cost on educational outcomes, to our knowledge, no previous study has examined the cost component of elementary students' science learning. Moreover, no study has examined the possible gendered patterns in science motivation at such an early age (for an exception, see Oppermann et al., 2018). It is crucial to examine these aspects of science motivation in early education in order to understand why students, especially girls, are opting out of science education and careers. In addition to our primary aim in the present study, we examine if there are gender differences in young students' science-related motivational beliefs and aspirations.

THE CURRENT STUDY

In this study, we draw on the framework of modern EVT (Eccles, 2009) to analyze a large sample of first-grade students (aged 7 years) in Finland, who were studied twice, 1 year apart. We examined science-related self-concept of ability, intrinsic value, and cost; the stability of these factors; and their unique contributions to science motivation development in students. In addition, the study investigates the extent to which science-related self-concept of ability, intrinsic value, and cost predict students' future STEM occupational aspirations. Finally, we address gender differences in students' self-concept of ability, task values, and STEM aspirations. Of central importance, the present study captures the positive and negative valence of science task values to explore the unique power of first graders' science-related self-concept of ability and task motivation on their future STEM occupational aspirations in the second grade.

Research Question 1: What Are the Autoregressive and Cross-Lagged Effects Between Science-Related Self-Concept of Ability, Intrinsic Value, and Cost Across the First and Second Grades?

Our first aim is to examine the mean-level stability and rank-order stability of science-related self-concept of ability, intrinsic value, and cost from Grades 1 to 2. Young children tend to be optimistic about their abilities across different academic subjects, and they place high subjective task values on different school subjects (Viljaranta et al., 2016). However, as they gather more experience with different academic subjects, gain more cognitive skills, and experience a wider range of school environments, such optimism changes to pronounced realism and even pessimism for many children (Stipek and Mac Iver, 1989; Wigfield et al., 2016). Based on these results, we expect first graders to have high self-concept of ability and intrinsic value beliefs at the beginning of their school career, and that their self-concept might decrease from Grade 1 to Grade 2. Moreover, based on prior literature on the development of task values, we assume that students' motivational beliefs will not be very stable at the age of 7–8 years (Wigfield and Eccles, 1992; Eccles and Wigfield, 2001). In the science domain, cost has not been previously studied in students of this age cohort. Thus, we are unable to hypothesize the stability of perceived cost or whether first graders perceive science learning as exhausting and demanding.

We also aim to examine the cross-lagged relations of science-related self-concept of ability, intrinsic value, and cost across the first and second grades. In line with the literature (Eccles et al., 1993; Wigfield and Eccles, 2002; Wigfield et al., 2009), we expect self-concept of ability to be positively related to intrinsic value and cost to be negatively related to self-concept of ability and intrinsic value (Barron and Hulleman, 2015).

Research Question 2: Do First Graders' Science-Related Task Values Predict Their Future STEM Occupational Aspirations 1 Year Later?

In middle and high school, students' science motivation has been found to predict educational and occupational aspirations (Wang and Degol, 2013; Guo et al., 2018a). Based on EVT, we hypothesize that students' high intrinsic value and self-concept of ability in science are positively associated with their STEM aspirations and that perceived cost in science is negatively associated with STEM aspirations a year later.

Research Question 3: Are There Gender Differences in Students' Science-Related Task Values and STEM Occupational Aspirations in the First and Second Grades?

It has been shown that boys hold higher self-beliefs and intrinsic value in science in early education (Leibham et al., 2013). However, recent findings indicate differences in girls' and boys' value beliefs in the physical (e.g., physics) and life (e.g., biology) sciences, and that there are increasing gender differences in physics and biology in middle school (Gaspard et al., 2017; Guo et al., 2018b). In line with previous findings, we expect boys to have a higher self-concept of ability and intrinsic value toward science at the beginning of elementary school. Gender equality is strongly promoted in Finnish society and emphasized in school; however, despite these efforts, gendered trajectories persist in education and occupations. Therefore, we are unable to formulate a hypothesis about the effect of gender on students' future STEM occupational aspirations.

MATERIALS AND METHODS

Participants and Procedure

The study sample consisted of 332 students, who underwent two rounds of testing: in the first grade and, 1 year later, in the second grade (Time 1: median age = 7 years, SD = 0.319, 188 girls, 144 boys; Time 2: median age = 8 years, SD = 0.389, 188 girls, 144 boys). The data were collected in 2016 and 2017 during the spring semester. The students were from 7 schools and 20 classes (two to five classes per school) located in the eastern suburbs of Helsinki, characterized by mixed levels of socio-economic status. There were one to two researchers per class, instructing and guiding the data collection. First, the students were introduced to the principles of answering a questionnaire and what the scales meant. It was emphasized that the most important thing was to answer honestly, that each opinion was valuable, that the responses would not be used for classroom evaluation purposes, and that their teachers would not see the responses. The students answered the questionnaires as part of a guided activity; the researcher read each item aloud, explaining unfamiliar concepts as needed. Special emphasis was placed on explaining the reversed items and how the scale should be interpreted with respect to those items. Students who had difficulties with the Finnish language or with following these procedures were assisted. The data were collected at the beginning of the spring semester in the student's first year of school to ensure that they had acquired basic reading skills and could more easily follow the questionnaire. At that time, they had half a year of experience studying science, or environmental studies, as it is called in the curriculum. Thus, it can be expected that the students were familiar with the context of the questions and understood the questions when they provided their answers. After the group completed each page of the questionnaire, they took a short break. The questionnaire was completed during one lesson (about 45 min).

The research project follows the strict national ethical guidelines of scientific studies of human subjects set by the Finnish Advisory Board of Research Integrity (TENK[1]), which are in line with the European Code of Conduct for Research Integrity of All European Academies (ALLEA) and the General Data Protection Regulation recently issued by the European Commission. The University of Helsinki Ethical Review Board in the Humanities and Social and Behavioural Sciences sanctions these national guidelines (TENK) and provides six descriptions of research designs that need to be handed for ethical reviews[2]. According to these guidelines, this study did not require ethical review, and therefore, no ethics application was made. Furthermore, to follow good scientific practice, the research plan was pre-examined and approved by the Education Division of the city of Helsinki. Since the participants of the study were elementary school-aged children, the study description and the permission forms for participation were sent to the students' parents beforehand.

[1] https://www.tenk.fi/en

[2] https://www.helsinki.fi/sites/default/files/atoms/files/when_are_ethical_reviews_required.pdf (accessed May 2019).

Parental consent was sought, and parents' were asked to either give or decline permission to take part in the study. Written active parental consent was obtained from all the student participants. Data collection was integrated in students' normal classroom activities. The headmasters and teachers of the participating schools were informed about and agreed to the data collection schedule. The class teacher organized separate activities for those students who did not have permission to participate in the study.

The Finnish Science Education Context

In Finland, students start school in the year they turn 7. Before school begins, children attend preschool for 1 year. The concept of science refers to school science, or environmental studies, as defined in the Finnish National Core Curriculum for Basic Education (NCCBE, 2014). According to the NCCBE (2014), environmental studies are an integrated subject, which comprises the knowledge fields of biology, geography, physics, chemistry, and health education. Its key objective is to guide students to understand the impact of the choices made by humans on life and the environment. The multidisciplinary nature of the subject requires that students learn to acquire, process, produce, present, evaluate, and appraise information in different situations (NCCBE, 2014). The viewpoints of scientific information and critical thinking are emphasized. In the first and second grades, the teaching and learning of environmental studies is structured into units in which the students' own environment, the students themselves, and their actions are examined. The students' curiosity and interest in phenomena in their surroundings are stimulated through problem-solving and inquiry assignments based on play. Students practice analyzing and naming elements in their surroundings and examine issues related to their own well-being and safety. The objectives of the subject in the first and second grades emphasize the development of environmental awareness, attitudes, values; developing research and working skills; and understanding the meanings of basic concepts, such as processes and structures in nature, the environment, and energy (NCCBE, 2014).

Measures

Task Motivation in Science

Students' science-related self-concept of ability, intrinsic value, and cost were examined using a task-value instrument based on EVT (Eccles et al., 1983; Eccles, 2009). The scale included self-concept of ability in science (i.e., "I am good at science," "I am good at schoolwork on this subject," and "Schoolwork on this subject is easy for me"; Time 1 $\alpha = 0.66$, Time 2 $\alpha = 0.63$), science intrinsic value (i.e., "I find science fun," "I like to do schoolwork on this subject," and "I just like this subject"; Time 1 $\alpha = 0.89$, Time 2 $\alpha = 0.85$), and science cost (i.e., "I am tired after doing schoolwork on this subject," "Studying this subject takes a lot of energy," and "I don't have time to do the thing I want, if I want to be good at this subject"; Time 1 $\alpha = 0.59$, Time 2 $\alpha = 0.66$). We used Likert-type visual scales from 1 = "Totally disagree" to 5 = "Totally agree," in which 1 was indicated with the smallest star and 5 with the biggest star, etc.

Occupational Aspirations

Information on the students' future occupational aspirations was sought using an open-ended question about their dream jobs. In the second grade, 61% of the students were able to name an occupation depicted as their dream job. The answers were classified by occupation level: support occupation (e.g., hairdresser) and professional occupation (e.g., medical doctor), with the most frequent answers being police officer, medical doctor, teacher, and professional football player. The occupations were further classified according to whether they fit in the STEM field (e.g., medical doctor, astronaut, game inventor, veterinarian) or not (e.g., sales person, football player, hairdresser, teacher). Using a coding scheme in which STEM included both physical sciences and life sciences, the answers were coded as 0 = support level and 1 = professional level and as 0 = non-STEM and 1 = STEM.

Background Information

Background information collected in the questionnaire included gender (0 = girl, 1 = boy) and age (i.e., date of birth).

Analytical Strategy

All analyses were conducted using longitudinal structural equation modeling (SEM) (Kline, 2005), as estimated by Mplus 8.0 (Muthén and Muthén, 2017). The models were estimated using the robust maximum-likelihood (MLR) estimator, which is robust against the non-normality of the observed variables and further considers the treatment of responses on a five-point Likert-type scale as the continuous variables (Beauducel and Herzberg, 2006; Hox et al., 2010; Muthén and Muthén, 2017). The MLR estimator was used in conjunction with full information maximum-likelihood (FIML) estimation in order to cope with a reasonable number of missing responses in the data. Only 30 students (9%) dropped out of the follow-up because they were absent from school on the day of the data collection due to illness or because their families had moved to another area. For similar reasons, 39 new students (12%) joined the study in the second grade. The intra-class correlations of the students' self-concept of ability, intrinsic value, and cost were calculated at the classroom level and at the student level. The purpose was to explore whether it would be critical to analyze the model as multilevel. The class-level variances and intra-class correlations between the classes were low (ranging from 0.02 to 0.10) at both time points, indicating that the students' motivation was mainly explained at the student level, which meant that the multilevel model was not necessary.

Model fit was evaluated by considering a wide range of descriptive goodness-of-fit indices (e.g., Marsh et al., 2004), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR), which are reported with the traditional Chi-square statistics and the corresponding degrees of freedom. For the CFI and TLI, values above 0.90 and 0.95, respectively, represent an adequate and good model fit (Hu and Bentler, 1999). SRMR and RMSEA values below 0.06 and 0.08, respectively, reflect a good and acceptable fit to the data (Browne and Cudeck, 1993; Hu and Bentler, 1999).

In order to make gender comparisons, we had to ensure that the constructs were measured similarly for boys and girls and that they remained the same across time points. Thus, the group and longitudinal invariance of the factor loadings and intercepts were tested. To compare models and evaluate invariance, we examined the changes in the descriptive goodness-of-fit indices. According to the guidelines proposed by Cheung and Rensvold (2002), two models can be seen as equivalent, and invariance can be assumed as long as the change in the CFI is not more than 0.01 and the RMSEA increases by less than 0.015 for a more parsimonious model. Given the various goodness-of-fit indices and their controversial cut-off criteria for model fit evaluation, researchers are recommended to simultaneously take different goodness-of-fit indices into account and treat the respective cut-off criteria as guidelines instead of golden rules.

RESULTS

Descriptive Statistics and Correlations

According to the present study's descriptive statistics (**Table 1**), the students had a high science-related self-concept of ability, i.e., they felt science was interesting, and they did not perceive a high cost in science learning.

To examine the factor structure of science-related self-concept of ability, intrinsic value, and cost, confirmatory factor analysis (CFA) was employed. Correlations between the latent factors are given in **Table 2** (see **Supplementary Table S1**, for factor loadings and effect sizes of measured items in science task value scale). The estimated latent correlations were drawn from the strong measurement model with equal intercepts. All latent correlations among self-concept, intrinsic value, and cost within the time points were statistically significant. Moreover, self-concept of ability in the first grade correlated positively with intrinsic value and cost a year later, and intrinsic value in the first grade correlated negatively with cost a year later. However, the correlations between the first graders' science-related intrinsic value and cost and their later self-concept of ability were not significant.

Three-factor measurement models were specified for both boys and girls as well as separately for the first and second graders. The fit indices of the models were considered to be good (**Table 3**). After satisfactory measurement models were found separately for gender and each grade, we tested the measurement invariance of the CFA models across time and gender. The configurally invariant CFA models, where no constraints were placed on any of the parameter estimates, fit the data well (**Table 3**). Testing for weak measurement invariance involved constraining each corresponding factor loading to be equal across gender and time, while strong measurement invariance also involved equalizing the corresponding intercepts across gender and time. The change in model fit between the configural and weak model as well as the change between the weak and strong model were modest and considered acceptable (**Table 3**). Since the multiple-group models were invariant, we decided to collapse the covariance information across the groups and specify the full SEM as a single-group model, while setting gender as a covariate (Little et al., 2007).

Results for Research Question 1: What Are the Autoregressive and Cross-Lagged Effects Between Science-Related Self-Concept of Ability, Intrinsic Value, and Cost Across the First and Second Grades?

After establishing measurement invariance, we investigated the autoregressive and cross-lagged effects of self-concept of ability, intrinsic value, and cost across the first and second grades, using a cross-lagged panel model. The cross-lagged panel model is used to examine reciprocal relationships or directional influences between variables across time (Kearney, 2017). Autoregressive effects describe the stability of the construct between measurement points, while cross-lagged effects indicate the association between two variables across time. The results showed that students' science-related self-concept of ability was somewhat stable ($\beta = 0.384$, $p = 0.004$), while there was a great deal of fluctuation in their science-related intrinsic value ($\beta = 0.216$, $p = 0.039$) and cost ($\beta = 0.225$, $p = 0.047$). Large regression coefficients indicate greater stability, while small regression coefficients indicate more variance in the construct, i.e., less stability across time (Kearney, 2017). To examine the interrelations of the students' task values, cross-lagged effects between self-concept of ability, intrinsic value, and cost were investigated. The only significant cross-lagged effect was found between self-concept of ability and cost: A higher self-concept of ability in science in the first grade predicted lower cost in science in the second grade (**Figure 1**).

Results for Research Question 2: Do First Graders' Science-Related Task Values Predict Their Future STEM Occupational Aspirations 1 Year Later?

Students' future STEM occupational aspirations and levels of occupational aspirations (hence, educational aspirations) were analyzed by regressing occupational aspirations on first-grade science motivation and gender. Since the variables for occupational and educational aspirations were dichotomous, logistic regression was used to model the relationship between these aspirations and science motivation. The results show that first-grade science intrinsic value was a marginally significant predictor of future STEM occupational aspirations in the second grade ($\beta = 0.668$, $p = 0.056$).

Results for Research Question 3: Are There Gender Differences in Students' Science-Related Task Values and Future STEM Occupational Aspirations in the First and Second Grades?

Finally, results indicate that girls were more interested in science than boys in both the first and the second grade (T1: $Z = 0.364$, $SE = 0.141$, $p = 0.10$; T2: $Z = 0.241$, $SE = 0.118$, $p = 0.041$). The mean levels of the girls' science task values remained stable across the first and second grades, while the mean levels of the boys'

TABLE 1 | Descriptive statistics and correlations.

	First grade (Time 1)									Second grade (Time 2)								
Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
Time 1																		
1. I am good at science	–																	
2. I am good at schoolwork on this subject	0.53**	–																
3. Schoolwork on this subject is easy for me	0.26**	0.43**	–															
4. I think this subject is fun	0.44**	0.39**	0.25**	–														
5. I like to do schoolwork on this subject	0.51**	0.50**	0.19**	0.79**	–													
6. I just like this subject	0.39**	0.39**	0.26**	0.69**	0.69**	–												
7. I am tired after doing schoolwork on this subject	–0.09	–0.20**	–0.08	–0.17**	–0.17**	–0.17**	–											
8. Studying this subject takes a lot of energy	–0.01	–0.08	0.00	–0.05	–0.01	–0.09	0.39**	–										
9. I don't have time to do the thing I want, if I want to be good in this subject	–0.02	–0.06	0.06	–0.17**	–0.20**	–0.16**	0.34**	0.24**	–									
Time 2																		
10. I am good at science	0.19**	0.12	–0.03	0.01	0.00	0.01	–0.01	–0.04	–0.01	–								
11. I am good at schoolwork on this subject	0.23**	0.23**	0.07	0.15*	0.11	0.13*	0.00	–0.10	–0.01	0.35**	–							
12. Schoolwork on this subject is easy for me	0.10	0.19**	0.24**	0.05	0.04	0.05	–0.04	–0.05	–0.09	0.32**	0.40**	–						
13. I think this subject is fun	0.19**	0.03	0.00	0.26**	0.19**	0.28**	–0.09	–0.08	–0.16*	0.29**	0.47**	0.24**	–					
14. I like to do schoolwork on this subject	0.20**	0.16**	0.04	0.20**	0.21**	0.25**	–0.05	–0.05	–0.17**	0.32**	0.40**	0.32**	0.67**	–				
15. I just like this subject	0.27**	0.15*	0.00	0.28**	0.29**	0.33**	–0.07	–0.08	–0.11	0.24**	0.40**	0.27**	0.60**	0.68**	-			

(Continued)

TABLE 1 | Continued

Variables	First grade (Time 1)									Second grade (Time 2)								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
16. I am tired after doing schoolwork on this subject	−0.18**	−0.16*	−0.14*	−0.14*	−0.14*	−0.17**	0.20**	0.17**	0.04	−0.20**	−0.26**	−0.20**	−0.20**	−0.24**	−0.25**	–		
17. Studying this subject takes a lot of energy	−0.15*	−0.23**	−0.15*	−0.13*	−0.16*	−0.14*	0.12	0.20**	0.01	−0.10	−0.28**	−0.29**	−0.20**	−0.20**	−0.19**	0.58**	–	
18. I don't have time to do the thing I want, if I want to be good in this subject	−0.06	−0.09	0.00	−0.12	−0.09	−0.19**	0.11	0.18**	0.21**	−0.16**	−0.16**	−0.13*	−0.19**	−0.22**	−0.14*	0.34**	0.25**	–
M	4.06	4.19	3.88	3.78	3.79	3.71	2.7	2.64	2.58	3.82	3.96	3.86	3.83	3.84	3.87	2.64	2.39	2.27
SD	1.102	1.105	1.271	1.467	1.401	1.543	1.708	1.62	1.641	1.009	1.048	1.163	1.318	1.25	1.212	1.51	1.485	1.471
Range	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

**Correlation is significant at the 0.05 level (two-tailed). **Correlation is significant at the 0.01 level (two-tailed).*

self-concept of ability decreased over the study's 1-year period ($Z = 0.330$, $SE = 0.128$, $p = 0.010$) (**Table 2**). The cross-lagged model revealed that the boys were more likely to experience higher cost in science learning in the first grade, and the girls were more likely to have a higher self-concept of ability in science in the second grade while controlling for the first grade motivation variables. In addition, the girls reported more future STEM occupational aspirations than the boys, when both genders had a similar level of EVT motivation. The regression coefficients for the gender effects are presented in **Figure 1**.

DISCUSSION

Our aim in the present study was to examine the stability and interrelations of early elementary students' self-concept of ability, intrinsic value, and cost in science learning within a 1 year time period. We also examined whether these motivational beliefs were associated with students' occupational aspirations and whether these beliefs differed between genders at the beginning of their school careers. The study had four major findings, expanding the existing literature on young students' science motivation. First, we found that students' science-related self-concept of ability and intrinsic value were high and that they perceived a low cost in science learning in the first and second grades. Some stability in students' self-concept of ability and positive and negative science motivation were found, but there was also a great deal of fluctuation in the rank-order of intrinsic value and cost across the first and second grades. Second, only one cross-lagged effect between the motivational beliefs was significant: high self-concept of ability was linked to low cost a year later. Third, the students' high intrinsic value in science in Grade 1 marginally significantly predicted their STEM aspirations in Grade 2. Fourth, we found that compared to boys, girls had higher science motivation at both time points, and higher self-concept of ability in the second grade. Boys perceived higher cost in science learning in the first grade compared to girls. Moreover, girls reported more STEM aspirations in the second grade. These findings suggest that girls are initially motivated in science and that it is worthwhile to investigate gendered trajectories in STEM as early as possible.

Stability and Interrelations of Science-Related Self-Concept of Ability, Intrinsic Value, and Cost Across the First and Second Grades

First, we found that the students' self-concept of ability and intrinsic value were high and that they perceived a low cost in science learning in the first and second grades. The different science-related EVT components were already separable among the first graders, supporting earlier studies on mathematics and languages (Eccles et al., 1993; Viljaranta et al., 2016). Students, especially the girls, were motivated in science and perceived themselves as skillful at the beginning of their school careers. This result confirms our hypothesis and is in accordance with

TABLE 2 | Estimated correlation matrix for the latent variables.

	First grade (T1)			Second grade (T2)					
Variable	1	2	3	4	5	6	7	8	9
1. Self-concept (T1)	–								
2. Intrinsic value (T1)	0.67***	–							
3. Cost (T1)	−0.2**	−0.25***	–						
4. Self-concept (T2)	0.33***	0.15	−0.11	–					
5. Intrinsic value (T2)	0.24**	0.33***	−0.16	0.7***	–				
6. Cost (T2)	−0.32***	−0.23***	0.27**	−0.49***	−0.37***	–			
7. Occupation level (T2)	0.17	0	−0.06	0	−0.3	−0.16*	–		
8. STEM (T2)	−0.02	0.16*	−0.18*	−0.06	0.04	−0.04	0.06	–	
9. Gender	0.03	−0.16**	0.15*	−0.16*	−0.11	0.06	0.12	−0.22***	–
M girl	0.00	0.00	0.00	−0.14	0.04	−0.17			
M boy	−0.03	−0.36*	0.20	−0.36*	−0.20	0.02			
Std.Error girl	0.00	0.00	0.00	0.11	0.10	0.11			
Std.Error boy	0.14	0.14	0.14	0.14	0.13	0.14			

*Latent means by gender are compared to girls' fixed mean in T1. **Correlation is significant at the 0.01 level (two-tailed). *Correlation is significant at the 0.05 level (two-tailed).*

TABLE 3 | Model fit statistics for the longitudinal confirmatory factor analysis (CFA) and structural equation modeling (SEM) models.

Models	*df*	χ^2	Scaling correction factor	*p*	RMSEA	CFI	TLI	SRMR
CFA models separately for boys and girls								
CFA boy time 1	24	25.00	0.99	0.406	0.018	0.997	0.996	0.042
CFA boy time 2	24	23.58	1.14	0.486	0	1	1.003	0.043
CFA girl time 1	24	39.29	1.20	0.025	0.062	0.954	0.932	0.047
CFA girl time 2	24	19.03	1.24	0.751	0	1	1.028	0.034
Longitudinal multiple-group models	222	265.43	1.06	0.024	0.034	0.969	0.958	0.059
Configural across gender and time								
Factor loading invariance across gender and time	240	287.99	1.07	0.018	0.035	0.966	0.957	0.07
Intercept invariance across gender and time	258	312.26	1.06	0.012	0.036	0.962	0.955	0.071
Single-group CFA models								
CFA time 1	24	50.24	1.11	0.001	0.06	0.961	0.941	0.041
CFA time 2	24	28.07	1.24	0.257	0.024	0.992	0.988	0.031
Configural across time	111	147.45	1.11	0.012	0.031	0.974	0.964	0.043
Factor loading invariance across time	117	155.59	1.12	0.010	0.031	0.972	0.964	0.049
Intercept invariance across time	123	168.26	1.12	0.004	0.033	0.967	0.959	0.049
SEM models								
Cross-lagged panel model with gender as a covariate	138	185.46	1.10	0.004	0.032	0.966	0.958	0.049
CLPM with logistic regression (STEM aspirations as outcome)	162	206.59	1.09	0.010	0.029	0.97	0.961	0.048

those of previous studies, which reported that at the beginning of their school careers, children are typically optimistic about their abilities (e.g., Stipek and Mac Iver, 1989) and show high levels of intrinsic value toward various school subjects (e.g., Gottfried et al., 2001; Viljaranta et al., 2016). These results vary considerably from the results usually found for students at the end of compulsory education. According to the PISA 2015 (OECD, 2016) affective measurements, Finnish students reported the fourth lowest enjoyment in science learning among OECD countries. Moreover, students' self-concept of ability and positive and negative motivation were somewhat stable; however, there was also a great deal of fluctuation in these constructs across Grades 1 and 2. The results show that in terms of the science-related self-concept of ability, there was some within-student (rank order) stability between the time points, whereas intrinsic value and cost fluctuated. This finding is in line with our hypotheses, and confirms existing literature indicating that the development of students' motivational beliefs is not very stable from the age of 7–8 years (Wigfield and Eccles, 1992; Eccles and Wigfield, 2001). The literature also suggests that science-related self-concept of ability, interest, and cost are still malleable at the beginning of elementary school, making it possible to influence the formation of young students' science motivation.

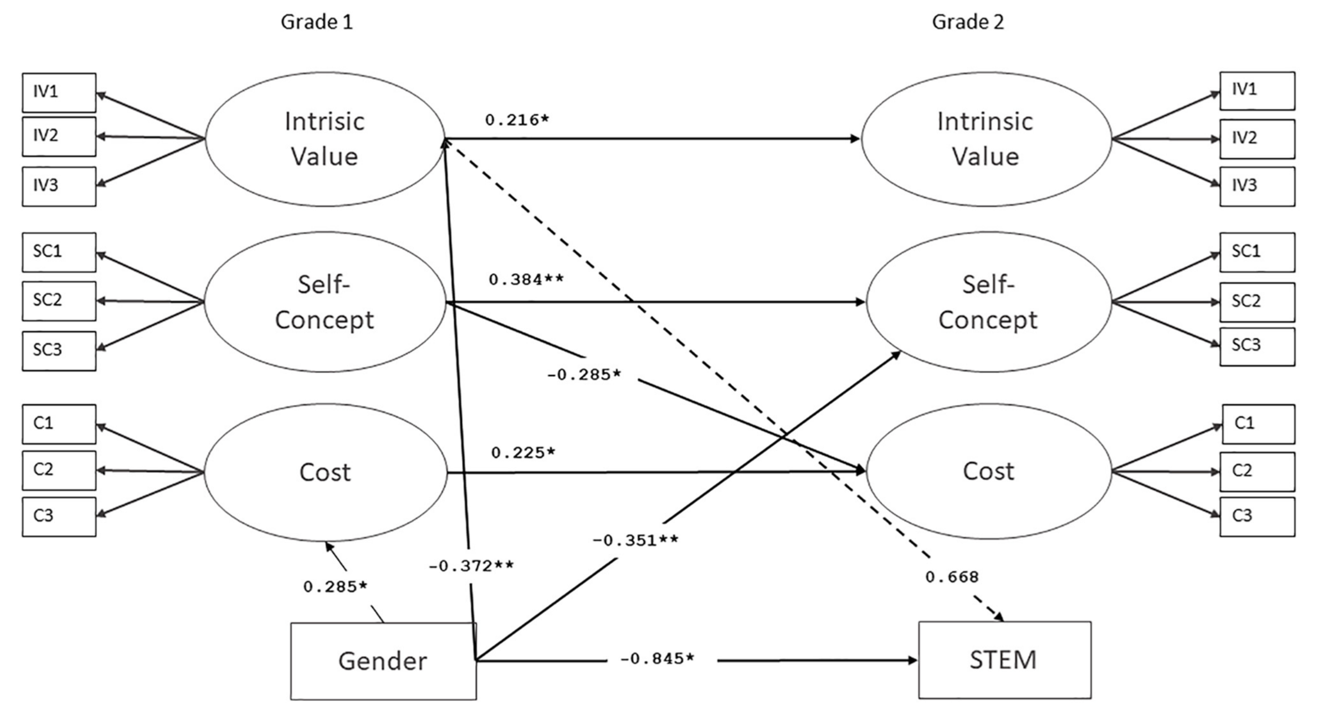


FIGURE 1 | Cross-lagged panel model for students' task motivation and occupational aspirations. IV, intrinsic value; SC, self-concept; C, cost. *$p < 0.015$, **$p < 0.01$, ***$p < 0.001$.

Second, the present study reveals a negative link between self-concept of ability and cost after 1 year. This finding is in line with our hypothesis and the EVT (Eccles et al., 1983; Eccles and Wigfield, 2001) and suggests that science-related exhaustion might result from a lack of self-evaluated abilities. The finding gives depth to existing literature on young students' science motivation, especially in relation to cost, which has not been studied before. As an educational implication, we need to emphasize intrinsic value while planning science lessons, and teachers need to ensure that students are able to do the science tasks. Strengthening students' ability beliefs by providing supportive teaching practices, and emphasizing formative and encouraging assessment in the early elementary years would be crucial. In addition, by organizing interesting science activities and avoiding ranking-oriented summative assessment, we might be able to engage students in science learning and encourage them to develop STEM aspirations.

First Graders' Science Interest and Future STEM Occupational Aspirations

Our results showed that students' high intrinsic value in science in first grade predicted their STEM aspirations marginally significantly in the second grade. Although the relation between science intrinsic value and future STEM occupational aspirations was only marginally significant, the effect size was rather large, suggesting a link between interest and STEM aspirations in young students. This tentative finding, which should be interpreted with caution, is in line with existing literature on students' interest and related STEM aspirations in middle and high school (Guo et al., 2015b, 2017). The poor probability value might be due to missing data, since one-third of the students in the second grade could not name a dream job. Still, a high intrinsic value in science seems to evoke or perhaps create possibilities for imagining future occupational aspirations at the beginning of elementary education. This raises the question of whether intrinsic value in school science starts to direct students' career choices as early as age 7. At the very least, it does seem that the links regarding science motivation and STEM aspirations emerge very early. As an educational implication, since one-third of the students did not mention their dream occupation, STEM occupations should be introduced in elementary school to increase students' awareness (Miller et al., 2018) and to connect STEM occupations to situations where intrinsic value is emphasized in school.

Gender Differences in Students' Science Task Values and Future STEM Occupational Aspirations

In this study, girls had higher science motivation and self-concept of ability than boys at the beginning of elementary school. The boys perceived greater cost in their science learning than the girls, and their science-related self-concept of ability decreased from the first to the second grade. The results revealed the importance of gender differences in school science learning; the girls were more interested in science in the first grade and had a higher self-concept of ability and more future STEM occupational aspirations in the second grade than the boys. It seems that gender differences in valuing STEM start to develop early. However, gender effects in the present study partly contradict earlier findings in science interest (e.g., Leibham et al., 2013) and self-concept of ability (e.g., Guo et al., 2018b), which showed that girls have a higher science-related task motivation and self-concept of ability than boys. Previous research has also showed that although globally boys tend to present higher science self-concept than girls, in some countries the gender gap was wider and in other countries narrower (Wilkins, 2004).

Discrepancy in the findings might reflect developmental changes in students' science motivation. It has been shown that boys have higher self-concept of ability in mathematics and higher value in mathematics learning in elementary school (Eccles et al., 1983), and that girls' mathematics interest is significantly lower than that of boys' in middle school (Gaspard et al., 2015; Guo et al., 2015b, 2017). The changing character of science and the increasing intensity of mathematics in physics and chemistry in middle school might influence the later decline in girls' science motivation. It is also possible that there are some unique elements in the Finnish education system that explain the present results on gender differences in early science motivation. For example, boys' underachievement and general lack of motivation in school, and the strong promotion of gender equality in schools, which the Pisa 2015 data highlight: Finnish girls' are outperforming boys and the majority of the students in other OECD countries (OECD, 2016). In addition, girls might be more mature and ready for school demands than boys, who experience greater cost in science learning. However, as the Pisa 2015 results also showed, boys have more positive attitudes to science than girls at the end of compulsory education, which indicates that girls' science interest also declines in Finland (OECD, 2016).

In the current study, girls reported more STEM aspirations than boys in Grade 2. In prior studies, boys have been found to have more STEM aspirations than girls (Eccles, 2011; Wang and Degol, 2013), but these studies were conducted among middle and high school students. It has been suggested that girls' low science interest and lack of STEM-related career aspirations result from gender socialization (Watt et al., 2012). Prior studies show that stereotypical beliefs on gender roles are well-developed before the start of formal education influencing children's interest (see Bian et al., 2017) although there might be an ongoing generational shift in children gender stereotypes in science (Miller et al., 2018). The media, students' peers, parents, and teachers might promote gender stereotypes in which girls are not expected to be interested or achieve success in mathematics and science, which might prevent gender-atypical behaviors (Watt, 2016). Thus, the results of the current study might vary with age. Girls' higher STEM aspirations in this study could be related to interest in life sciences, which in our coding were included in the STEM occupation category. Thus, those girls whose dream job was to be a medical doctor or a veterinarian were included as having occupational STEM aspirations. It is important to follow the development of science interest in these students to investigate possible changes that might occur when they transition to middle school.

CONCLUSION

In conclusion, young students' science motivation was rather high at the beginning of elementary school, but the mean levels of intrinsic value declined during the first year. Science-related self-concept was more stable compared to intrinsic value and cost in science learning, and high self-concept of ability seemed to buffer against perceived cost. Students' high motivation was related to STEM occupational aspirations. Gendered differences in science motivation were found at the age of 7 years, and favored girls.

Limitations and Further Research

The SEM model in the current study was not analyzed as a multilevel model, which would have been appropriate as the students were nested within classrooms and schools. However, as the class-level variances and the intra-class correlations between the classes were small, and the number of classrooms was insufficient to take the hierarchical structure adequately into account, we decided not to use the multilevel model. In future studies, it would be important to analyze data collected from students in different classrooms and schools as multilevel.

We acknowledge that predicting the occupational aspirations of students in the first years of elementary school is far-fetched, since their knowledge of possible careers is limited, influenced by their parents' occupations and occupations visible in media, especially TV animations, being perhaps rather traditional or fictional. Unfortunately, students' awareness of their parents' occupations was too limited for us to use as an indicator of their socio-economic status. Now that the students are older and their awareness of their parents' jobs might be clearer, their parents' occupations will be the subject of further enquiry in a follow-up data collection effort. Moreover, more sensitive career coding in science (e.g., physical sciences and life sciences) is crucial to draw conclusions on possible changes in gendered pathways in STEM.

In order to understand students' low science motivation, it would be worthwhile to investigate the development of their motivational trajectories across their elementary, middle, and high school education to try to identify *how* science motivation develops across school years, *when* changes occur, and *why* these changes occur. It would also be necessary to compare students' views of science to their views of mathematics and language to examine more closely the relationships between school subjects in students' motivational beliefs. In the future, we plan to investigate whether these gender effects remain unchanged or whether later on in their schooling boys develop higher science motivation levels than girls and if so, why and when. Moreover, the link between science interest and STEM aspirations in young students calls for further research into the tentative finding of the current study.

ETHICS STATEMENT

The research project follows strict national ethical guidelines regarding scientific studies of human subjects set by the Finnish Advisory Board of Research Integrity (TENK, https://www.tenk.fi/en), which are in line with the European Code of Conduct for Research Integrity of ALLEA and the General Data Protection Regulation recently issued by the European Commission. The University of Helsinki Ethical Review Board in the Humanities and Social and Behavioural Sciences sanctions these national guidelines (TENK) and provides six descriptions of research designs that require ethical review (see https://www.helsinki.fi/sites/default/files/atoms/files/when_are_ethical_reviews_required.pdf). According to these guidelines, this study did not require ethical review, and therefore, no application was made for an ethics approval. Furthermore, to follow good scientific practice, the research plan was pre-examined and approved by the Education Division of the city of Helsinki. Since the participants of the study were elementary school-aged children, the description of the study and permission forms for participation were sent to the students' parents beforehand. Parental consent was sought, and parents had the opportunity to decline their child's participation in the study. Written informed parental consent was obtained for all the student participants. The data collection was integrated in the students' normal classroom activities. The headmasters and teachers of the participating schools were informed about and agreed to the data collection schedule. The class teacher organized separate activities for the students who did not have permission to participate in the study.

AUTHOR CONTRIBUTIONS

JV-L and JG performed the analytic calculations. JV-L wrote the manuscript, with the help of JG and KS-A. KS-A conceived the original idea. KJ and AL collected the data. JL and KS-A helped supervise the project. All authors discussed the results and contributed to the final manuscript.

FUNDING

This study was supported by the Academy of Finland grant no. 298323 to KS-A.

ACKNOWLEDGMENTS

We thank Helen Watt and Filomena Parada for their useful discussions and helpful advice. We are grateful to the teachers in the participating schools in the Helsinki area, who made the data collection possible.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01449/full#supplementary-material

REFERENCES

Archer, L., Osbourne, J., DeWitt, J., Dillon, J., and Wong, B. (2013). *ASPIRES: Young People's Science and Career Aspirations, Age 10-14*. London: King's College.

Barron, K. E., and Hulleman, C. S. (2015). Expectancy-value-cost model of motivation. *Psychology* 84, 261–271. doi: 10.1111/medu.13617

Battle, A., and Wigfield, A. (2003). College women's value orientations toward family, career, and graduate school. *J. Vocat. Behav.* 62, 56–75. doi: 10.1016/s0001-8791(02)00037-4

Beauducel, A., and Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Struct. Equ. Model.* 13, 186–203. doi: 10.1207/s15328007sem1302_2

Bian, L., Leslie, S. J., and Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science* 355, 389–391. doi: 10.1126/science.aah6524

Browne, M. W., and Cudeck, R. (1993). Alternative ways of assessing model fit. *Sociol. Methods Res.* 154, 136–136.

Chambers, N., Kashefpakdel, E. T., Rehill, J., and Percy, C. (2018). *Drawing the Future: Exploring the Career Aspirations of Primary School Children From Around the World*. London: Education and Employers.

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equa. Model.* 9, 233–255. doi: 10.1097/NNR.0b013e3182544750

Chow, A., Eccles, J. S., and Salmela-Aro, K. (2012). Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Dev. Psychol.* 48, 1612–1628. doi: 10.1037/a0030194

Denissen, J. J., Zarrett, N. R., and Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Dev.* 78, 430–447. doi: 10.1111/j.1467-8624.2007.01007.x

Eccles, J. S. (2009). Who am I and what am I going to do with my life? personal and collective identities as motivators of action. *Educ. Psychol.* 44, 78–89. doi: 10.1080/00461520902832368

Eccles, J. S. (2011). Gendered educational and occupational choices: applying the eccles et al. model of achievement-related choices. *Int. J. Behav. Dev.* 35, 195–201. doi: 10.1177/0165025411398185

Eccles, J. S., and Wigfield, A. (2001). "Academic achievement motivation, development of," in *International Encyclopedia of the Social and Behavioral Sciences*, eds N. J. Smelser and P. B. Baltes (New York, NY: Elsevier), 14–20. doi: 10.1016/B978-0-08-097086-8.26001-7

Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Ann. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153

Eccles, J. S., Wigfield, A., Harold, R. D., and Blumenfield, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Dev.* 64, 830–847. doi: 10.1111/j.1467-8624.1993.tb02946.x

Eccles, P. J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). "Expectations, values and academic behaviors," in *Achievement and Achievement Motivation*, ed. J. T. Spence (San Francisco, CA: W. H. Freeman), 75–146.

Finnish Advisory Board of Research Integrity (2019). *Finnish Advisory Board of Research Integrity*. Available at: https://www.tenk.fi/en (accessed January 2019).

Finnish National Board of Education (2014). *National Core Curriculum for Basic Education*. Helsinki: Finnish National Board of Education, Publications.

Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., and Welsh, M. E. (2015). Measuring cost: the forgotten component of expectancy-value theory. *Contemp. Educ. Psychol.* 41, 232–244. doi: 10.1016/j.cedpsych.2015.03.002

Fredricks, J. A., and Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: growth trajectories in two male-sex-typed domains. *Dev. Psychol.* 38, 519–533. doi: 10.1037//0012-1649.38.4.519

Gaspard, H., Dicke, A. L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., et al. (2015). More value through greater differentiation: gender differences in value beliefs about math. *J. Educ. Psychol.* 107, 663–677. doi: 10.1037/edu0000003

Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., and Nagengast, B. (2017). Assessing task values in five subjects during secondary school: measurement structure and mean level differences across grade level, gender, and academic subject. *Contemp. Educ. Psychol.* 48, 67–84. doi: 10.1016/j.cedpsych.2016.09.003

Gottfried, A. E., Fleming, J. S., and Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: a longitudinal study. *J. Educ. Psychol.* 93:3. doi: 10.1037//0022-0663.93.1.3

Guo, J., Eccles, J. S., Sortheix, F. M., and Salmela-Aro, K. (2018a). Gendered pathways toward STEM careers: the incremental roles of work value profiles above academic task values. *Front. Psychol.* 9:1111. doi: 10.3389/fpsyg.2018.01111

Guo, J., Marsh, H. W., Morin, A. J., Parker, P. D., and Kaur, G. (2015a). Directionality of the associations of high school expectancy-value, aspirations, and attainment: a longitudinal study. *Am. Educ. Res. J.* 52, 371–402. doi: 10.3102/0002831214565786

Guo, J., Marsh, H. W., Parker, P. D., and Dicke, T. (2018b). Cross-cultural generalizability of social and dimensional comparison effects on reading, math, and science self-concepts for primary school students using the combined PIRLS and TIMSS data. *Learn. Instruct.* 58, 210–219. doi: 10.1016/j.learninstruc.2018.07.007

Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J., and Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: dimensional comparison processes and expectancy-by-value interactions. *Learn. Instruct.* 49, 81–91. doi: 10.1016/j.learninstruc.2016.12.007

Guo, J., Parker, P. D., Marsh, H. W., and Morin, A. J. (2015b). Achievement, motivation, and educational choices: a longitudinal study of expectancy and value using a multiplicative perspective. *Dev. Psychol.* 51, 1163–1176. doi: 10.1037/a0039440

Harackiewicz, J. M., and Hulleman, C. S. (2010). The importance of interest: The role of achievement goals and task values in promoting the development of interest. *Soc. Personal. Psychol. Compass* 4, 42–52. doi: 10.1111/j.1751-9004.2009.00207.x

Harter, S. (2012). *The Construction of the Self: Developmental and Sociocultural Foundations*, Vol. 2. New York, NY: Guilford Publications.

Hox, J. J., Maas, C. J., and Brinkhuis, M. J. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Stat. Neerland.* 64, 157–170. doi: 10.1111/j.1467-9574.2009.00445.x

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equa. Model. Multidiscipl. J.* 6, 1–55. doi: 10.1080/10705519909540118

Jacobs, J., Lanza, E. S., Osgood, D. W., Eccles, J. S., and Wigfield, A. (2002). Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Dev.* 73, 509–527. doi: 10.1111/1467-8624.00421

Jiang, Y., Rosenzweig, E. Q., and Gaspard, H. (2018). An expectancy-value-cost approach in predicting adolescent students' academic motivation and achievement. *Contemp. Educ. Psychol.* 54, 139–152. doi: 10.1016/j.cedpsych.2018.06.005

Kearney, M. W. (2017). *Cross Lagged Panel Analysis. The SAGE Encyclopedia of Communication Research Methods,* 16. Los Angeles: Sage.

Kline, T. (2005). *Psychological Testing: A Practical Approach to Design and Evaluation*. London: Sage.

Leibham, M. B., Alexander, J. M., and Johnson, K. E. (2013). Science interests in preschool boys and girls: relations to later self-concept and science achievement. *Sci. Educ.* 97, 574–593. doi: 10.1002/sce.21066

Little, T. D., Preacher, K. J., Selig, J. P., and Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *Int. J. Behav. Dev.* 31, 357–365. doi: 10.1177/0165025407077757

Marsh, H. W., Chessor, D., Craven, R., and Roche, L. (1995). The effects of gifted and talented programs on academic self-concept: the big fish strikes again. *Am. Educ. Res. J.* 32, 285–319. doi: 10.3102/00028312032002285

Marsh, H. W., Hau, K. T., and Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Struct. Equa. Model.* 11, 320–341. doi: 10.1207/s15328007sem1103_2

Marsh, H. W., and Martin, A. J. (2011). Academic self-concept and academic achievement: relations and causal ordering. *Br. J. Educ. Psychol.* 81, 59–77. doi: 10.1348/000709910X503501

Martin, M. O., Mullis, I. V. S., Foy, P., and Hooper, M. (2016). *TIMSS 2015 International Results in Science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Miller, D. I., Nolla, K. M., Eagly, A. H., and Uttal, D. H. (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of US draw-a-scientist studies. *Child Dev.* 89, 1943–1955. doi: 10.1111/cdev.13039

Muthén, L., and Muthén, B. (2017). *Mplus User's Guide*. Available at: http://www.statmodel.com/download/usersguide/Mplus%20user%20guide%20Ver_7_r6_web (accessed September 2018).

NCCBE (2014). *National Core Curriculum for Basic Education. Opetushallitus. [Finnish National Agency of Education]*. Helsinki: Next Print Oy.

OECD (2015). *Education at a Glance 2015: OECD Indicators*. Paris: OECD Publishing.

OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA. Paris: OECD Publishing.

Oppermann, E., Brunner, M., Eccles, J. S., and Anders, Y. (2018). Uncovering young children's motivational beliefs about learning science. *J. Res. Sci. Teach.* 55, 399–421. doi: 10.1002/tea.21424

Perez, T., Cromley, J. G., and Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *J. Educ. Psychol.* 106, 315–329. doi: 10.1037/a0034027

Ryan, C. (2015). *Science Education for Responsible Citizenship*. Brussels: European Commission.

Simpkins, S. D., Davis-Kean, P. E., and Eccles, J. S. (2006). Math and science motivation: a longitudinal examination of the links between choices and beliefs. *Dev. Psychol.* 42, 70–83. doi: 10.1037/0012-1649.42.1.70

Stipek, D., and Mac Iver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Dev.* 60, 521–538. doi: 10.1111/j.1467-8624.1989.tb02734.x

Tytler, R. (2014). "Attitudes, identity, and aspirations toward science," in *Handbook of Research on Science Education*, eds N. G. Lederman and S. K. Abell (New York, NY: Routledge), 82–103.

Viljaranta, J., Aunola, K., and Hirvonen, R. (2016). Motivation and academic performance among first-graders: a person-oriented approach. *Learn. Individ. Diff.* 49, 366–372. doi: 10.1016/j.lindif.2016.06.002

Viljaranta, J., Tolvanen, A., Aunola, K., and Nurmi, J. E. (2014). The developmental dynamics between interest, self-concept of ability, and academic performance. *Scand. J. Educ. Res.* 58, 734–756. doi: 10.1080/00313831.2014.904419

Wang, M. T. (2012). Educational and career interests in math: a longitudinal examination of the links between perceived classroom environment, motivational beliefs, and interests. *Dev. Psychol.* 48, 1643–1657. doi: 10.1037/a0027247

Wang, M. T., and Degol, J. (2013). Motivational pathways to STEM career choices: using expectancy-value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* 33, 304–340. doi: 10.1016/j.dr.2013.08.001

Wang, M. T., and Degol, J. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): current knowledge, implications for practice, policy, and future directions. *Educ. Psychol. Rev.* 29, 119–140. doi: 10.1007/s10648-015-9355-x

Wang, M. T., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice: individual and gender differences in STEM career choice. *Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937

Watt, H. M. (2016). "Gender and motivation," in *Handbook of Motivation at School*, eds K. R. Wentzel and D. B. Miele (New York, NY: Routledge), 320–339.

Watt, H. M., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., and Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: a comparison of samples from Australia, Canada, and the United States. *Dev. Psychol.* 48, 1594–1611. doi: 10.1037/a0027838

Wigfield, A., and Cambria, J. (2010). Students' achievement values, goal orientations, and interest: definitions, development, and relations to achievement outcomes. *Dev. Rev.* 30, 1–35. doi: 10.1016/j.dr.2009.12.001

Wigfield, A., and Eccles, J. S. (1992). The development of achievement task values: a theoretical analysis. *Dev. Rev.* 12, 265–310. doi: 10.1016/0273-2297(92)90011-p

Wigfield, A., and Eccles, J. S. (2002). "The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence," in *Development of Achievement Motivation*, eds A. Wigfield and J. S. Eccles (San Diego, CA: Academic Press), 91–120. doi: 10.1016/b978-012750053-9/50006-1

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J., Freedman-Doan, C., et al. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *J. Educ. Psychol.* 89, 451–469. doi: 10.1037/0022-0663.89.3.451

Wigfield, A., Tonks, S., and Klauda, S. L. (2009). "Expectancy-value theory," in *Handbook of Motivation at School*, eds K. R. Wentzel and D. B. Miele (New York, NY: Routledge), 69–90.

Wigfield, A., Tonks, S., and Klauda, S. L. (2016). "Expectancy-value theory," in *Handbook of Motivation at School*, eds K. R. Wentzel and D. B. Miele (New York, NY: Routledge), 55–74.

Wilkins, J. L. M. (2004). Mathematics and science self-concept: an international investigation. *J. Exp. Educ.* 72, 331–346. doi: 10.3200/jexe.72.4.331-346

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 10 July 2019
doi: 10.3389/feduc.2019.00060

Check for updates

The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations

Elena Makarova[1]*, *Belinda Aeschlimann*[2] *and Walter Herzog*[3]

[1] *Institute for Educational Sciences, University of Basel, Basel, Switzerland,* [2] *Swiss Federal Institute for Vocational Education and Training SFIVET, Bern, Switzerland,* [3] *Institute of Educational Science, University of Bern, Bern, Switzerland*

Studies have repeatedly reported that math and science are perceived as male domains, and scientists as predominantly male. However, the impact of the gender image of school science subjects on young people's career choice has not yet been analyzed. This paper investigates the impact of the masculinity image of three school subjects—chemistry, mathematics, and physics—on secondary students' career aspirations in STEM fields. The data originated from a cross-sectional study among 1'364 Swiss secondary school students who were close to obtaining their matriculation diploma. By means of a standardized survey, data on students' perception of masculinity of science school subjects were collected using semantic differentials. The results indicate that for both sexes, math has the strongest masculinity attribution, followed by physics as second, and, finally, chemistry with the lowest masculinity attribution. With respect to gender differences, our findings have shown that among female students, the attribution of masculinity to the three school subjects does not differ significantly, meaning that female students rated all subjects similarly strongly as masculine. Within the group of male students however, the attribution of masculinity to math compared to chemistry and physics differs significantly, whereas the attribution of masculinity to chemistry and physics does not. Our findings also suggest that gender-science stereotypes of math and science can potentially influence young women's and men's aspirations to enroll in a STEM major at university by showing that a less pronounced masculine image of science has the potential to increase the likelihood of STEM career aspirations. Finally, the paper discusses ways of changing the image of math and science in the context of secondary education in order to overcome the disparities between females and males in STEM.

Keywords: gender, career aspirations, science, mathematics, secondary school students

OPEN ACCESS

Edited by:
Bernhard Ertl, Universität der Bundeswehr München, Germany

Reviewed by:
Bettina Hannover, Freie Universität Berlin, Germany
Jessica Lauren Degol, Pennsylvania State University, United States

***Correspondence:**
Elena Makarova
elena.makarova@unibas.ch

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *28 December 2018*
Accepted: *11 June 2019*
Published: *10 July 2019*

Citation:
Makarova E, Aeschlimann B and Herzog W (2019) The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations. Front. Educ. 4:60. doi: 10.3389/feduc.2019.00060

INTRODUCTION

Gender segregation in the vocational orientation of adolescents has been well documented for decades in most OECD countries (OECD, 2006, 2012). The persistence of gendered paths in career choices has recently been reflected in the current Global Gender Gap Report of the World Economic Forum (WEF), which states that on average men are underrepresented in the fields of education, health and welfare whereas women are underrepresented in the STEM fields (WEF, 2017, p. 31). Moreover, on the basis of the occupational aspirations of 15-year-old adolescents, the prognosis for change in gender-based disparities in occupational and academic choices suggests that gender segregation in the education and labor market will remain persistent (OECD, 2017).

The persistence of horizontal gender segregation in educational and occupational fields contributes decisively to the spread of gender-stereotypic beliefs about a natural fit of women in careers in more expressive and human-centered fields and men in technical and math-intensive fields (Charles and Bradley, 2009). Gender stereotypes are part of a broader belief system that includes attitudes toward female and male family roles, female and male occupations, and gender-associated perceptions of the self. As bipolar constructs, gender stereotypes imply that what is masculine is not feminine and vice versa (Deaux and LaFrance, 1998; Worell, 2001; Renfrow and Howard, 2013). The social role theory (Eagly and Wood, 2012) suggests that gender roles and their occupants are highly visible in everyday contexts and that gender stereotypes emerge in response to the observation of women and men in different social roles and in role-linked activities related to occupational choices (Koenig and Eagly, 2014). This theoretical assumption was confirmed in a study by Miller et al. (2015), which analyzed how women's enrollment in science courses relates to the gender-science stereotype. Based on a survey of about 3,50,000 participants in 66 nations, this study concluded that explicit and implicit national gender-science stereotypes were weaker in countries with a higher female enrollment in tertiary science education. This study also demonstrated that stereotypes about science were strongly gendered, even in countries with high overall gender equity. In addition, a meta-analysis of two major international data sets—"Trends in International Mathematics and Science Study" (TIMMS) and the "Programme for International Student Assessment" (PISA)—has confirmed that gender equity in education is important not only for girls' math achievement but also for girls' self-confidence and valuing of mathematics (Else-Quest et al., 2010). Furthermore, a cross-national data analysis has indicated that gender differences in math are closely related to cultural variations in opportunity structures for girls and women, in particular to gender equity in school enrollment, women's share of research jobs, and women's parliamentary representation (ibid., p. 103). Accordingly, the low proportion of women in STEM leads to the spread of a gender stereotypical image of math and science as a male domain and beliefs about male supremacy in technical and math-intensive fields. In turn, such beliefs affect young people's career choices, leading to a mutual reinforcement of gender stereotypes, and gender gaps in career related interests and choices (Nosek et al., 2009, p. 10,596).

In Switzerland gender segregation is also persistent and is especially noticeable in the STEM field (FSO, 2013). In educational tracks at the universities of applied science, with only 21.3% of women enrolled in STEM courses in academic year 2017–2018. However, some STEM fields are more strongly gender segregated than others. The lowest proportion of women is in the fields of informatics (10.4%) and technology (8.5%), whereas in the fields of chemistry and life-sciences the proportion of women is considerably higher (43.7%) (FSO, 2019a). In secondary education, gender is almost balanced in chemistry and biology (girls 18.4% and boys 20.5%) as a subject of specialization, whereas considerably more boys (18.4%) than girls (4.4%) decided to specialize in the subjects math and physics (FSO, 2019b). It is, thus, important to distinguish between different STEM disciplines and subjects when addressing the gender gap in the STEM field (Rosser, 2012; Ertl et al., 2017).

Following this notion, our study aimed to analyze the *gender stereotype of school science subjects among female and male students* and *the impact of gender-science stereotypes on the career aspirations* of young people. The ultimate goal of our study is to provide a more comprehensive understanding of gender equity in STEM.

THE GENDER STEREOTYPE OF MATH AND SCIENCE

The gender stereotype of math and science has been analyzed via a variety of quantitative and qualitative methods (review in Makarova and Herzog, 2015). Among those are the *Draw-A-Scientist Test* (DAST) (e.g., Chambers, 1983; Finson, 2002; Scherz and Oren, 2006), the *Implicit Association Test* (IAT) (e.g., Greenwald et al., 1998; Nosek et al., 2002, 2009), *explicit stereotype assessments* using attitude questionnaires (e.g., Kessels, 2005), *semantic differential assessments* (e.g., Herzog et al., 1998; Makarova and Herzog, 2015), and *individual or group interviews* (e.g., Archer et al., 2010).

Studies that applied the DAST method reported that students from kindergarten to high school perceive a *scientist* as a male person. The children's drawings contained very few portrayals of female scientists and these few drawings were mostly drawn by female students. For example, in a study among students from kindergarten through fifth grade there were only 28 pictures of a female scientist out of 4,807, and all of these 28 drawings were drawn by girls (Chambers, 1983); in a study surveying students in grades 2–12 only 135 pictures out of 1,600 displayed female scientists and only six out of 135 pictures of a female scientist were drawn by male students (Fort and Varney, 1989); in a study among students of 9–12 years of age, there were only 72 pictures of a female scientist out of 223, and of those 72, only 13 pictures were drawn by male students (Huber and Burton, 1995). The precise way in which a scientist was pictured by middle school students was reported in a study by Scherz and Oren (2006, p. 977): "The common image was that of a scientist as a bespectacled male with unkempt hair in a white lab-coat." Moreover, the following quote from a study by Mead and Metraux (1957) on high-school students' image of a scientist highlights how persistent the scientist-stereotype remains over decades. The image of a scientist is depicted in students' essays as "a man who wears a white coat and works in a laboratory. He is elderly or middle aged and wears glasses … He may wear a beard, may be unshaven and unkempt" (Mead and Metraux, 1957, p. 386). Finally, the most recent meta-analysis of five decades of U.S. DAST studies based on 78 studies ($N =$ 20,860) among children grades K-12, shows a growth in children's depictions of female scientists in later decades. However, the more female scientist appeared only in drawings by young children, but science was still associated with men among older children (Miller et al., 2018). The authors conclude that despite the increase of women's representation in science over the last decades, children still observe more male than female scientists in their social environments (Miller et al., 2018, p. 1,943).

Furthermore, research on gender stereotypes has revealed that science is not only associated with a male person, but that *masculine traits* are also attributed to it. A study by Archer et al. (2010) suggested that although young children do not have profound knowledge about science subjects, they attribute masculine traits to science at an early age. In the same vein, a study by Cvencek et al. (2011) reported that as early as second grade children perceive that math is a male domain, demonstrating the American cultural stereotype. In addition, a study among high school students reported that better performance in STEM subjects was attributed to boys, and masculine traits to a person who works as a scientist (Hand et al., 2017). Another study among school children and university students by Weinreich-Haste (1981) assessed the gender image of different academic subjects using ratings on a six-point masculine-feminine scale. The study reported that math, physics and chemistry had the strongest connotation as masculine academic subjects. Moreover, it showed that science subjects were not only rated as masculine but also associated with a set of attributes commonly associated with masculinity such as being hard, complex, based on thinking rather than on feelings (Weinreich-Haste, 1981, p. 220f.). In contrast, a study on gender perception of school subjects among students aged 11–12 years, which applied a seven-point masculine-feminine scale, reported that while physics was rated as significantly more masculine, chemistry and mathematics were rated as neither masculine nor feminine (Archer and MacRae, 1991).

To summarize, we can state that the male stereotype of science and of a scientist is persistent and appears as early as in kindergarten age, while the association of science with men is especially persistent among older children. Research has also shown that students predominantly perceive science subjects (math, physics, and chemistry) as a male domain, although findings do not provide a clear picture as to which of these subjects is more strongly associated with male gender. The reason is the very broad age-range of students (K-12) across reported studies, lack of comparison of gender stereotypes of different school subjects within one study, different methodology (explicit and implicit assessment) used to assess gender stereotypes of science, as well as the time span between findings of different studies. Thus, further research on the perception of masculinity of chemistry, math, and physics among school students is needed to gain deeper insight into the impact of the gender stereotypes of science subjects on STEM-career aspirations.

GENDER DIFFERENCES IN THE PERCEPTION OF GENDER-SCIENCE STEREOTYPES

Research on gender-science stereotypes has illustrated differences between female and male youth with respect to the *endorsement of stereotypic beliefs about STEM*. A study among primary school students illustrated that stereotypical beliefs that STEM school subjects are more suitable for boys than for girls were more strongly endorsed by boys than by girls. Moreover, this study has shown that students with stereotype-consistent interest in STEM-related school subjects were particularly likely to endorse gender-science stereotypes. Consequently, especially boys who were highly interested and girls who were relatively uninterested in STEM-related school subjects were more likely to believe that STEM school subjects constitute a male domain (BlaŽev et al., 2017). In line with this, a study among high school students has shown that girls reported lower self-efficacy in math and science compared to boys (Hand et al., 2017). Finally, a study among first-year university students indicated that negative stereotypes of women's engineering and mathematical ability were more strongly endorsed among male students, whereas female students were more likely to report higher perceptions of their engineering abilities (Jones et al., 2013).

With respect to the *perception of different STEM disciplines,* studies among adolescent youth have shown that female students show a more pronounced gender stereotype for math compared to male students, who are less likely to exhibit implicit gender-stereotypic associations (Steffens et al., 2010). In line with these findings, a study by Nosek et al. (2002, p. 44) reported that even women who had selected math-intensive majors had difficulties in associating math with themselves because they associated math with the male gender. Also, studies that analyzed the gender stereotype of physics found that, among high school students, being interested in physics was associated with the male gender (Kessels, 2005; Kessels et al., 2006) and that, among girls, being interested in physics endangered their self-identification with the female gender (Kessels et al., 2006). Furthermore, a typical teacher of mathematics and physics was imagined to be a man (Kessels and Taconis, 2012). Finally, a study among secondary school students in Switzerland showed that, among female students, the semantic profile of math and physics correlated negatively with the semantic profile of the female gender, whereas the semantic attributes of chemistry were significantly related neither to the male nor to the female gender. From the male students' point of view the semantic profile of math correlated negatively with the semantic profile of the female gender, whereas the semantic attributes of chemistry and physics were positively related to the semantic profile of the male gender. Whereas, the female gender was strongly associated with traits such as soft, playful, soulful, dreamy, lenient, frail, and flexible, among the semantic traits associated with math and physics were attributes such as hard, serious, distant, sober, strict, robust, and rigid. Overall, this study has shown that among the three school subjects analyzed in the study, math and physics were either negatively associated with female or positively associated with male gender. In contrast, chemistry was the least gender stereotyped because among female students there were no significant associations of the term chemistry with either gender term and among male students no negative association with the term woman (Makarova and Herzog, 2015). These findings are interesting in light of students' preference for their subject of specialization in secondary schools in Switzerland (FSO, 2019b) showing that chemistry is chosen almost equally often by boys and girls, whereas math and physics are largely avoided by girls as subjects of specialization. Accordingly, students' gender-related perception of different science subjects may differently impact their preferences of STEM subjects at school and vice versa.

To summarize, we can state that female and male students indicate different patterns of gender-science stereotype. It seems that male participants show more endorsement of the gender-science stereotype by regarding STEM subjects as more suitable for boys and attributing less abilities in the STEM disciplines to the female gender compared to the male gender. At the same time, female participants are more likely to associate math and science more strongly with the male gender and masculine traits than with the female gender and feminine traits. Finally, previous research has shown that school science subjects differ with respect to their gender-related connotation, and indicating that chemistry has the least pronounced masculine image among secondary school students.

GENDER-SCIENCE STEREOTYPE AND CAREER ASPIRATIONS IN STEM

The impact of the gender-science stereotype on students' interest in STEM subjects and their aspirations to pursue a career in STEM fields has been addressed from different perspectives.

Based on Eccles' expectancy-value model, which highlights the impact of culturally based stereotypes and identity-related constructs on educational and occupational choices (Eccles, 1994; Eccles and Wigfield, 2002), a number of studies have shown that academic *self-concept* and *subject interests* are among the most relevant determinants in students' selection of secondary school majors (Nagy et al., 2008). Similar mechanisms seem to be crucial for career choice or choice of a major in higher education (Nagy et al., 2006). A recent study among female students in STEM subjects with a low proportion of females revealed that gender stereotypes have a negative impact on students' STEM-specific self-concept even among students with good grades in STEM (Ertl et al., 2017).

According to the theoretical framework of Gottfredson (2002, 2005), occupational aspirations are incorporated in the *individual self-image* developed during socialization from early childhood through adolescence. The process of developing occupational aspirations is embedded in the comparison of one's self-image with the image of an occupation and one's judgment about the match between the two. In this process, the gender image of an occupation is especially crucial for career choice, because the "wrong" sex type of an occupation is more fundamental to self-concept than the prestige of an occupation or individual interests. Applying Gottfredson's theory, the significant impact of the gender image of an occupation on the process of career choice was confirmed in a number of studies (Ratschinski, 2009; Bubany and Hansen, 2011). Moreover, research suggests that girls are more likely to narrow their occupational choices because they perceive particular occupations as inappropriate for their gender. Accordingly, girls tend to shift their occupational aspirations to gender-typical occupational expectations more strongly than do boys. At the same time, boys' perceptions of occupations appear to be more gender-stereotypical (Hartung et al., 2005).

Research focusing on *self-to-prototype similarity* suggests that the lack of similarity between the self and an academic subject is linked to a lower probability of liking this subject or choosing this academic subject as a major (Kessels, 2005; Kessels et al., 2006; Taconis and Kessels, 2009). Moreover, the perceived closeness between the self and a school subject was predictive for youths' career choice intentions (Hannover and Kessels, 2004; Kessels et al., 2006). In the same vein, a study among ninth and tenth-grade students by Neuhaus and Borowski (2018) investigated whether the greater self-to-prototype similarity impacts students' interest in coding courses. This study revealed that, under the condition that course descriptions were related to communal goals, girls showed greater interest in learning to code compare to the agentic-goal condition of the course description (Neuhaus and Borowski, 2018, p. 233).

Likewise, a study among students and faculty reported that agentic traits are more strongly associated with success in science than communal traits, discouraging women from pursuing a science career (Ramsey, 2017). Another study among first-year undergraduate students illustrated that implicit stereotypes of science completely accounted for a gap in male and female students' interests to pursue science. Especially the academic aspirations of women who strongly identified as female were affected by the gender stereotypic image of science (Lane et al., 2012). In line with this, a study among first-year women engineering students reported that engineering identification was a significant predictor of persistence in engineering, and that this relationship was stronger for women than men (Jones et al., 2013). Finally, a study among undergraduate science majors demonstrated that a stronger gender-science stereotype has a diminishing effect on identification with science and science career aspirations among women, whereas, among men, a stronger gender-science stereotype boosts their identification with science and their career aspirations in science fields (Cundiff et al., 2013).

To summarize, we can state that gender-science stereotyping has been shown to hinder the self-identification of young women with STEM academic subjects and fields and also to negatively affect their self-concept and their subject interests. These, in turn, hinder female students from opting for a science major and pursuing a career in science. For male students, gender-science stereotyping seems to have the opposite effect and, thus, boosts their career aspirations in STEM.

FOCUS OF THE STUDY

Given that previous research has often focused on gender-science stereotypes of science in general or on stereotypical beliefs about single STEM disciplines, our study contributes to previous research by simultaneously analyzing the gender stereotype of different school science subjects—chemistry, math, and physics—among female and male students. These three science subjects were chosen because females are strongly underrepresented in math and physics within the educational sector and career fields, whereas chemistry has a more balanced gender ratio. This allows us to investigate the impact of gender-science stereotypes of different science subjects on students' aspirations to study STEM. In view of the theoretical and empirical framework of the study, we define the gender stereotype of three school subjects as the extent of association of each school subject with masculine traits (see section Measurements; *masculinity index*).

In terms of hypotheses, we firstly expected differences with respect to the degree of masculinity which students attribute to chemistry, math, and physics. We hypothesized that chemistry would be ascribed the lowest degree of masculinity compared to math and physics.

Secondly, we expected gender differences among secondary school students in the association of chemistry, math, and physics with male gender. We hypothesized that this association of the three science subjects with masculine traits would be stronger among female students.

Thirdly, we expected that the gender stereotype of math and science would affect female and male secondary school students' aspirations to enroll in a STEM major at university. We hypothesized that to the extent students conceive of STEM-school subjects as masculine they would be less inclined to aspire to enroll in a STEM major at university. We further hypothesized that stereotyping science subjects as masculine would have a greater negative impact on the STEM aspirations of female than male students.

METHODS

Participants

The study presented was part of the research project *Gender atypical career choices of young women*, a project embedded in the Swiss National Science Foundation's Research Program on "Gender Equality" (NRP 60). The study is based on quantitative data which originated from a standardized survey of 1,364 students in Swiss-German-speaking secondary schools. The study was carried out following the ethical principles and codes of the Faculty of Humanities at the University of Bern, which are based on international ethics codes (e.g., of the American Sociological Association and of the American Psychological Association). Accordingly, approval by an ethics authority was not required. Students were informed about the research project and participated in the survey voluntarily. Participants? Informed consent was implied through survey completion; therefore, they were not required to provide written consent to participate. Written parental consent was not necessary either, because all students had reached legal adulthood and could decide for themselves. After the survey all data were anonymized.

The surveyed students were close to obtaining their matriculation diploma (i.e., school leaving certificate), which in Switzerland permits entry into tertiary education. The participants were on average 19 years old ($SD = 1.0$). With regard to sex, the percentage of female students (54.1%) was somewhat higher than that of male students (45.9%).

Measurements

Masculinity Index

Data on students' perception of the gender image of the school subjects chemistry, math, and physics were collected using semantic differentials (Makarova and Herzog, 2015). The semantic differential is one of the most popular techniques of explicit attitude assessment (Millon et al., 2003). An explicit measurement of the gender stereotype of science subjects was chosen over an implicit stereotype test, because the study focuses on the salient gender stereotypes of those subjects (Millon et al., 2003, p. 356). The semantic differential uses bipolar scales with contrasting adjectives at each end to measure people's reactions to stimulus words and concepts (Heise, 1970, p. 235). The methodological advantage of the semantic differential scale is that it highly adaptable in assessing respondents' connotative association with any concept (Osgood et al., 1957; Heise, 1970). The basic assumption of the semantic differential is that attitudes toward two associated concepts tend to converge and toward two dissociated (contrasted) constructs tend to diverge (Heise, 1970, p. 249). In our study attitudes toward gender and science were measured using semantic differentials consisting of 25 pairs of adjectives with semantically opposite meanings (e.g., hard—soft, strong—weak, robust—frail) to assess the connotations of the four terms man, chemistry, math, and physics on a seven-point scale (1 = greatly, 2 = fairly, 3 = somewhat, 4 = neither, 5 = somewhat, 6 = fairly, 7 = greatly). This instrument is based on the original scale (Osgood et al., 1957) which was initially adapted to the German language by Hofstätter (1973) and then validated in Switzerland in two studies on the gender stereotype of school subjects (Herzog et al., 1998; Makarova and Herzog, 2015).

The student sample was divided into groups, with each group completing the semantic differential for one subject term and for the man term: chemistry and man ($n = 406$), math and man ($n = 512$) and physics and man ($n = 446$). In order to avoid response bias, the semantic differential of the subject was introduced at the beginning of the questionnaire and the semantic differential of the term man at the end of the questionnaire. On the basis of these data we calculated a masculinity index by subtracting the 25 items of the man profile from the corresponding items of each subject profile and summing them up to a sum score for each student. At the end of this procedure one value for each student was calculated. For easier interpretation, this value was reversed; a negative value was transformed into a positive value and a positive value into a negative value. Accordingly, the masculinity index expresses the differentiation between high masculinity (low discrepancy between the profiles man and subject; max. = +6) and low masculinity (high discrepancy between the profiles man and subject; min. = −6). For example, a score of 5 on the masculinity index, indicates that the semantic profile of the respective subject (chemistry, math, or physics) and the semantic profile of the term man are very similar, meaning that the discrepancy between the two semantic profiles is low. **Figure 1** illustrates our calculation. Moreover, the masculinity index is approximately normally distributed (Kurtosis = 2.09, $SE = 0.13$; Skewness = 0.47, $SD = 0.07$) (George and Mallery, 2016).

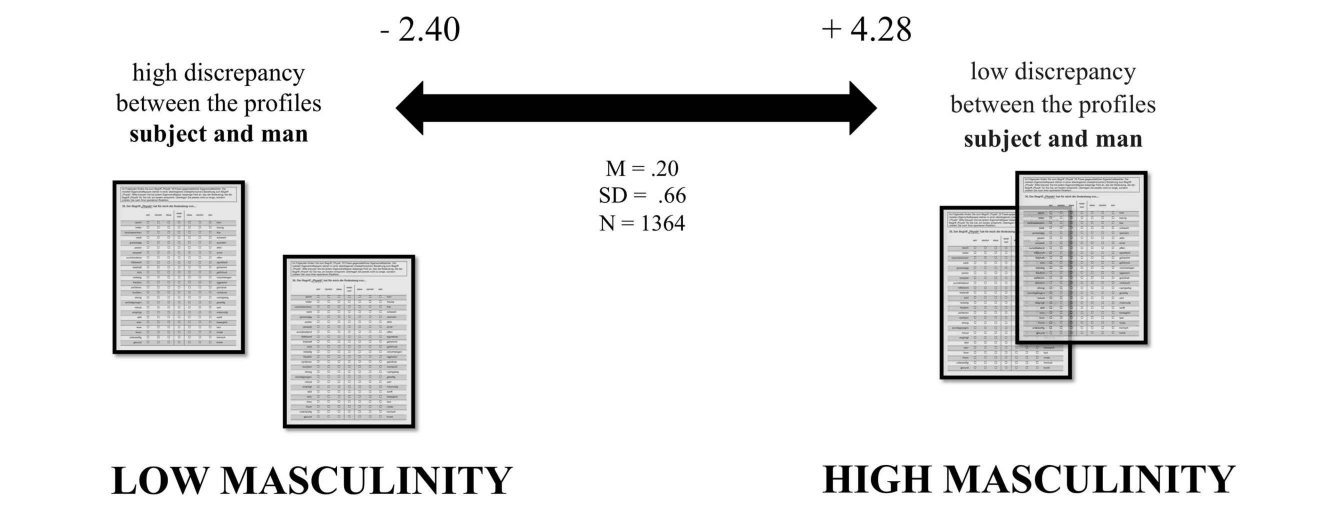


FIGURE 1 | Masculinity index of chemistry, math, and physics.

STEM Field Study Choice

To assess the field of study choice, the secondary school students were asked about their subject preference for study at a university or at a university of applied sciences after the successful completion of secondary school. The answers were coded by the gender-type of the field of study, based on the gender distribution of master's degrees obtained at Swiss universities in the year 2010 (FSO, 2012). A field of study was labeled as female-atypical (male-typical) when the proportion of women who received a master's degree in that field was below 30 per cent. In our sample, Mathematics, Statistics, IT, the Natural Sciences and Engineering fall into this category. Since all listed fields of study can be assigned to the STEM area, the category is henceforth labeled STEM study choice. All other fields of study were assigned to the category "non-STEM study choice." The multivariate analyses were conducted with the dichotomous variable STEM field study choice (STEM field study choice = category 1; non-STEM study choice = reference category 0).

TABLE 1 | Descriptive statistics.

	All	Female	Male
Masculinity of chemistry (min: −2.4, max: 4.28)	$n = 406$ $M = 0.13$ $SD = 0.71$	$n = 240$ $M = 0.27$ $SD = 0.66$	$n = 166$ $M = -0.09$ $SD = 0.72$
Masculinity of math (min: −1.76, max: 2.36)	$n = 512$ $M = 0.29$ $SD = 0.64$	$n = 242$ $M = 0.38$ $SD = 0.62$	$n = 267$ $M = 0.20$ $SD = 0.64$
Masculinity of physics (min: −2.08, max: 3.08)	$n = 446$ $M = 0.18$ $SD = 0.63$	$n = 257$ $M = 0.31$ $SD = 0.66$	$n = 189$ $M = 0.01$ $SD = 0.55$

M = mean, SD = standard deviation.

RESULTS

Attribution of Masculinity to Chemistry, Math, and Physics Among Secondary School Students

The attribution of masculinity to the three science subjects among female and male students was subjected to a two-way ANOVA (school subject and students' sex). The overall model yielded an F ratio of $F_{(5,\ 1,355)} = 15.83$, $p \leq 0.001$. With respect to the degree of masculinity attributed to the three science subjects, our analysis of variance indicated significant differences $F_{(2,\ 1,355)} = 10.76$, $p \leq 0.001$. *Post-hoc* comparisons (Bonferroni) has shown that the attribution of masculinity differs significantly between math and chemistry ($p \leq 0.001$) and between math and physics ($p \leq 0.05$). There were no significant differences in the attribution of masculinity to chemistry and physics. The mean values indicated that math has the strongest attribution of masculinity, followed by physics as second, and finally chemistry with the lowest attribution of masculinity (see **Table 1**). With regard to the sex differences in the attribution of masculinity, our analysis of variance yielded significant differences between female and male students $F_{(1,\ 1,355)} = 63.20$, $p \leq 0.001$. The ascription of masculinity to the three science subjects turned out to be stronger among female than among male students (see **Table 1**).

The interaction effect between two factors school subject and students' sex was non-significant $F_{(2,\ 1,355)} = 2.34$, $p = ns$. Nevertheless, to explore the interaction term in more detail we analyzed the attribution of masculinity to the three science subjects within the group of female and that of male students. For this purpose, the confidence intervals for the three science subjects were calculated. Within the group of female students, the attribution of masculinity to the three school subjects does not differ significantly, meaning that female students rated all subjects similarly as strongly masculine [95% CIs: chemistry [0.19, 0.36], math [0.30, 0.46], and physics [0.23, 0.39]]. Within the group of male students, however, the attribution of masculinity to math and chemistry [95% CIs [0.12, 0.27], [−0.20, 0.02]] as well as to math and physics [[0.12, 0.27], [−0.07, 0.09]]

TABLE 2 | Study choice.

	All	Female	Male	Interaction of gender × study choice
Study choice	$N = 1{,}618$	$n = 873$	$n = 742$	$x^2 = 58.26^{***}$
STEM choice	16.6%	10.1%	24.3%	
NON-STEM choice	83.4%	89.9%	75.7%	

The interaction of gender × study choice is significant at the $^{***}p \le 0.001$ *level,* x^2 = x^2*-value (chi-square-test).*

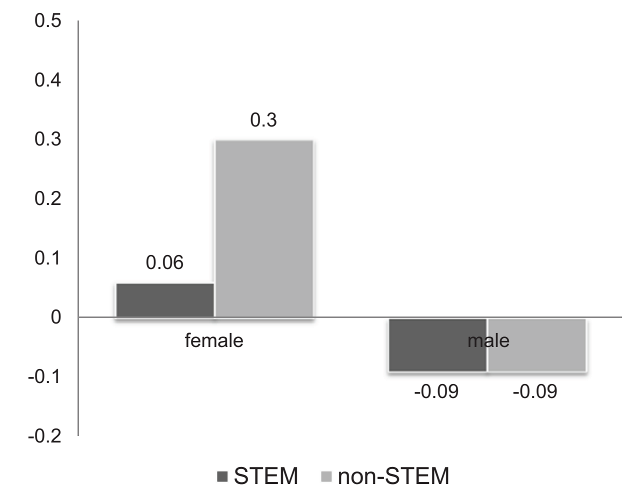


FIGURE 2 | Masculinity index of chemistry and career aspirations.

differs significantly, whereas the attribution of masculinity to chemistry and physics does not [[−0.20, 0.02], [−0.07, 0.09]].

Gender Stereotype of Chemistry, Math and Physics and Students' Study Aspirations

First, we analyzed career aspirations among the secondary school students by carrying out x^2-test (*chi-square test*) for the binomial dependent variable STEM study choice (see **Table 2**). Overall, one sixth of all students aspired to having a STEM major (16.6%). However, aspirations to study STEM subjects were not equally distributed between men and women. While among men every fourth student (24.3%) planned to study STEM, among women only every tenth student (10.1%) was interested in STEM studies.

Second, we analyzed the attribution of masculinity to school subjects (chemistry, physics, and math) among secondary school students who had chosen a STEM compared to those students who had chosen a non-STEM major (**Figures 2–4**).

Our analysis reveals the following findings for each subject:

- *Chemistry* (**Figure 2**): With respect to career aspirations of young women, our results show that female students who opt for a non-STEM study major connotated chemistry significantly strongly as masculine compared to young women with a STEM career choice ($p \le 0.01$). Among young men there were no significant differences in the attribution of masculinity to the subject chemistry between students who had chosen STEM and those who had chosen another study field.

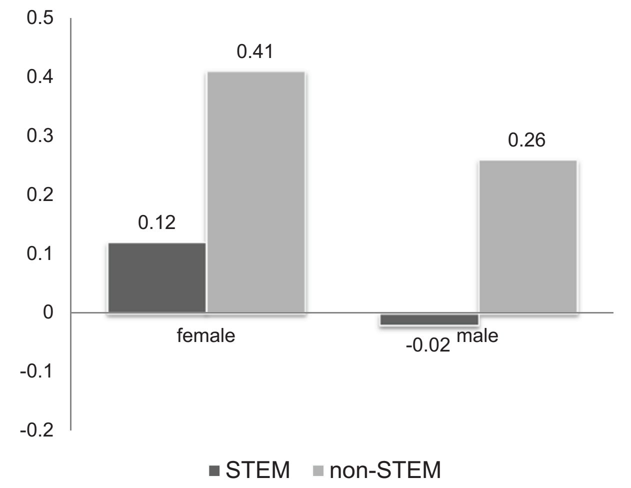


FIGURE 3 | Masculinity index of math and career aspirations.

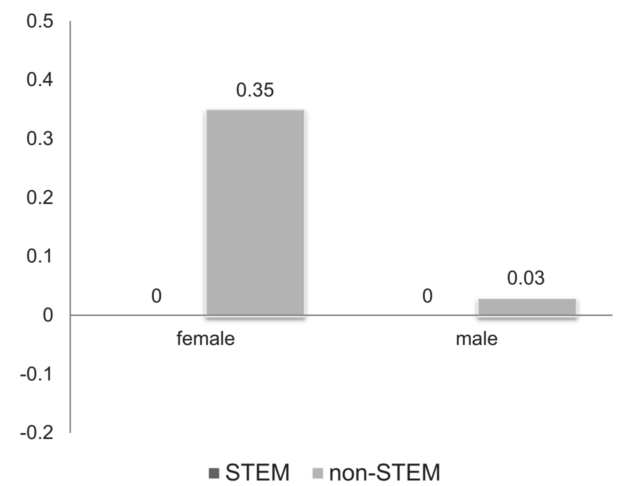


FIGURE 4 | Masculinity index of physics and career aspirations.

- *Math* (**Figure 3**): Our results show that among female and male students who had potentially chosen a non-STEM major, the attribution of masculinity to math was significantly higher compared to youth with a STEM career choice (female: $p \le 0.05$; male: $p \le 0.001$).
- *Physics* (**Figure 4**): Considering female students who had potentially chosen a non-STEM study major, physics was significantly more highly stereotyped as a masculine subject compared to young women with a STEM career choice ($p \le 0.001$). Among young men there were no significant differences in the attribution of masculinity to the subject physics between male students who had chosen STEM and those who had chosen another study field.

To sum up, young women who aspire to study a STEM major stereotype the three subjects as less strongly masculine compared to young women who aspire to study non-STEM subjects. Among young men, only math was rated as highly masculine among those students who had chosen a non-STEM study program. Thus, for young women as well as for young men with a non-STEM career choice, math has a highly masculine image. What is interesting is that even young women who opt for a STEM field rate the subjects—except physics—as masculine, though only slightly.

Finally, Generalized Linear Models (GzLM) were estimated (McCullagh and Nelder, 1989) to shed light on the impact of the gender image of math and science on the likelihood that female and male students aspire for a STEM field of study. The procedure modeled the choice of a STEM study major as the response category, with all other study fields as the reference category (non-STEM). We aggregated the masculinity index for math and the two science subjects for the model of female students, because separate models showed nearly the same effect for each individual subject, and therefore we could increase the power of the model in terms of cases. The model for male students included only the masculinity index of math as a predictor, since there was no significant effect for science subjects between young men who had chosen STEM and non-STEM ones (see also **Figures 2**, **4**). We report the Exp(β), which indicates the likelihood of an occurrence of the tested effect. If the value is below 1, the likelihood decreases; if it is above 1, the likelihood increases.

Table 3 shows the first model estimated for female students [Likelihood Ratio $x^2_{(1,739)} = 17.09$, $p \leq 0.001$, Pearson-Chi-Square 60.95 (88, 739) = 0.69]. The findings reveal that a strong masculine image of math and science decreases the likelihood of young women choosing a STEM study (Exp(β) = 0.44; $p \leq 0.001$). In other words, if young women do not perceive math and science as predominantly masculine, they opt significantly more often for studying a STEM major.

The second model was estimated for male students [Likelihood Ratio $x^2_{(1,267)} = 9.22$, $p \leq 0.01$, Pearson-Chi-Square 73.90 (66, 267) = 1.12]. The results show that the masculinity of math is also a predictor of young men's career aspirations. The higher the masculinity image, the lower the likelihood of a STEM study choice (Exp(β) = 0.48; $p \leq 0.01$).

To conclude, both models show that the image of chemistry, math and physics has an impact on students' career intentions. If the image of the three subjects has strong masculine connotations, career choice is unlikely to be within the STEM field.

TABLE 3 | Impact of the masculine image of math and science on secondary students' career aspirations.

Parameter	*β*	*SE*	*Wald-Chi-Square*	Exp(*β*)
Math and science model for female students				
(Intercept)	−1.98***	0.12	253.81	0.14
Masculinity of math and science	−0.82***	0.17	24.30	0.44
(Scale)	1[a]			
Math model for male students				
(Intercept)	−1.17***	0.15	59.88	0.31
Masculinity of math	−0.73**	−0.23	9.91	0.48
(Scale)	1[a]			

*Generalized Linear Model (binomial/logit). Dependent variable: STEM career (response) vs. Non-STEM Career (reference); [a] Fixed at the displayed value; β = regression coefficient; SE = standard error; ***p ≤ 0.001, **p ≤ 0.01.*

DISCUSSION

This study contributes to the line of research on the gender stereotype of science by analyzing the gender-related image of three *school subjects*. It provides, moreover, more refined knowledge on the impact of gender stereotypical perception of math and science on female and male secondary school students' choice to enroll in a STEM university degree program.

In line with the findings of a study by Weinreich-Haste (1981), our results reveal that students not only perceive chemistry, math and physics as masculine, but also that there is a considerable difference in the *strength of the association* of each subject with the male gender. According to our findings, math is most strongly perceived as a masculine subject among female and male secondary school students, followed by physics and then chemistry, which has the weakest masculine connotations. The weak masculine connotations of chemistry have also been reported by other studies (Archer and MacRae, 1991; Makarova and Herzog, 2015). Consequently, we could confirm the first hypothesis stating that chemistry is accorded the lowest degree of masculinity compared to math and physics.

With respect to differences between female and male students in the gender-stereotypical connotations of science, our findings illustrate that *female secondary school students* perceive all three subjects considerably more strongly as a male domain than do male students. These findings are consonant with findings of previous studies on strong associations of math and physics with the male gender among female adolescents (Nosek et al., 2002; Kessels, 2005; Kessels et al., 2006; Steffens et al., 2010). In addition, our results illustrate that male students regard only math as strongly masculine, whereas physics and chemistry have a comparably low score on the masculinity index. Thus, our findings confirm our second hypothesis by showing that the association of the three science subjects with masculine traits are stronger among female students.

With regard to the *impact of the masculinity image of math and science on secondary students' career aspirations*, the findings of our study show that young women who potentially chose STEM as a field of study at university perceived all three school subjects—math, physics, and chemistry—as less masculine than did those young women who chose other majors. Moreover, our results suggest that among female students a strong masculine image of math and science decreases the likelihood of choosing a STEM major at university. These findings propose that masculine traits associated with science subjects at school constitute a major obstacle, particularly for young women's self-identification with science (Nosek et al., 2002; Cundiff et al., 2013) and for their aspirations to become researchers (Šorgo et al., 2018). Regarding the *career aspirations of young women,* our study supports the notion that stereotypical beliefs about math and science prevent young women from entering a STEM career (Lane et al., 2012; Ramsey, 2017).

Finally, our results on the *career aspirations of young men* in relation to the stereotypical gender connotations of school subjects show that young men with non-STEM career aspirations perceived only math but not science subjects as significantly more strongly masculine than did young men who chose a STEM major. Furthermore, a strong association of math with masculine traits negatively affected male students' STEM career aspirations. These findings suggest that young men who opted for non-STEM majors do not fit the masculinity stereotype and therefore the strong masculine connotations of math may have an inhibiting impact on their career aspirations similar to that on the STEM career aspirations of young women. A possible interpretation of these findings is that, among young women *as well as* among young men, the lack of similarity between their self-image and the image of an academic subject not only affects their choice of specialization in secondary school (Kessels, 2005; Kessels et al., 2006; Taconis and Kessels, 2009) but also leads to a lower probability of choosing those subjects in their further educational career.

Overall, the findings of our study confirm our third hypothesis by illustrating that the higher the extent of association of STEM-school subjects with masculine traits, the lower is the likelihood to enroll in a STEM major at university—both for female and male students. However, our findings also suggest that gender-science stereotypes have a stronger negative impact on the STEM aspirations of female than male students because a strong masculine image of math *and* science significantly decrease the likelihood of choosing a STEM major among female students, whereas only a strong masculine image of math significantly decrease the likelihood of enrollment in a STEM major among male students.

Our findings have some *implications* for overcoming the gender disparities in STEM. As the gender-related image of an academic discipline has a considerable effect on young people's career aspirations, a critical evaluation of the school subjects' image might be one way to break through the gender-image-driven limitations of the career horizons of female *and* male students. For example, a study in Computer Science has shown that women's interest in studying Computer Science can be increased through a change of *image of this academic discipline* (Cheryan et al., 2013). The image of a school subject can, for example, be depicted in school textbooks. An empirical analysis of science textbooks in secondary education not only illustrated the overrepresentation of male protagonists but also revealed stereotypical portrayals of science and scientists (Makarova et al., 2016a). Since stereotypic representations in textbooks have an effect on male and female secondary school students' understanding of and anxiety about science (Good et al., 2010), an effort needs to be made to overcome stereotypical gender representations in textbooks at all educational levels. Especially since decisions to enroll in a field of study or choose a field of work in vocational education are made relatively late, and since gender images of school subjects have most likely by then been internalized and settled, reflections about gender stereotypical images of math and science subjects should preferably be encouraged *in early childhood*. For example, a study by Archer et al. (2010) suggested that although young children do not have profound knowledge about science subjects, they attribute masculine traits to science at an early age. Moreover, gender stereotypical beliefs should be also tackled among *teachers and other gatekeepers* who are potentially involved in the development of vocational interests among children and secondary students. As the study of Thomas (2017) showed, a teacher's implicit science-is-male stereotype can contribute to gender differences in female students' motivational beliefs and probably also their gendered educational choices. Finally, Else-Quest et al. (2010) suggest that proximal factors such as quality of teaching mediate the effect of gender inequality on math achievement. Thus, rise in gender equity in education can also promote boys' academic development.

Our study is subject to a few *limitations*. *Firstly*, our study has a cross sectional design and is, therefore, limited to suggesting a causal relationship between the masculinity image of science and youth career aspirations. *Secondly*, our study assesses the career aspirations of secondary school students and not their actual enrollment in particular majors at the university. Although this operationalization of career choice has been applied by other studies (Nagy et al., 2006; Watt, 2006), it does not exclude the possibility that the anticipated choice of a study major does not necessarily lead to the actual choice of the same major after enrollment at university. *Thirdly*, we should note that our study applies an explicit assessment of masculinity connotations of school subjects by using a semantic differential with 25 opposite semantic meanings. Thus, we cannot rule out that an open-ended questionnaire on masculinity image would yield different results on the semantic connotations and the strength of masculinity of the target school subjects. Moreover, we calculated the masculinity index based on the similarity of the semantic profiles of the term man and the corresponding subject term. As the present study does not include measures of the semantic ratings of the term woman we cannot compare the attribution of the feminine traits to chemistry, math and physics and its impact on the STEM study choice. *Finally*, the gender-related image of school subjects and their implications are one of several determinants that affect the career aspirations of male and female secondary school students. Since we did not control for other potential determinants in the explanatory models (e.g., self-image of students, their abilities, or interest in science), our results are limited to the investigation of the impact of gender-science stereotype on students' aspirations. It has been demonstrated that further school-related factors, such as the instructional design of science classes (Aeschlimann et al., 2016), teachers' support and encouragement (Aeschlimann et al., 2015) as well as family-related factors, and also peers can considerably influence the career-choice decisions of young people (Makarova et al., 2016b).

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

The authors gratefully acknowledge the Swiss National Science Foundation for financial support of the study Gender atypical careers of young women (Grant no. 4060-129279).

REFERENCES

Aeschlimann, B., Herzog, W., and Makarova, E. (2015). Frauen in MINT-Berufen: Retrospektive Wahrnehmung des mathematisch-naturwissenschaftlichen Unterrichts auf der Sekundarstufe I [Women in STEM professions: retrospective perception of mathematics and science in secondary school education]. *Zeitschr. Bildungsfor.* 5, 37–49. doi: 10.1007/s35834-014-0111-y

Aeschlimann, B., Herzog, W., and Makarova, E. (2016). How to foster students' motivation in mathematics and science classes and promote students' STEM career choice. A study in Swiss high schools. *Int. J. Educ. Res.* 79, 31–41. doi: 10.1016/j.ijer.2016.06.004

Archer, J., and MacRae, M. (1991). Gender perceptions of school subjects among 10–11-year-olds. *Br. J. Educ. Psychol.* 61, 99–103. doi: 10.1111/j.2044-8279.1991.tb00965.x

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., and Wong, B. (2010). "Doing" science versus "being" a scientist: examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Sci. Educ.* 94, 617–639. doi: 10.1002/sce.20399

BlaŽev, M., Karabegović, M., Burušić, J., and Selimbegović, L. (2017). Predicting gender-STEM stereotyped beliefs among boys and girls from prior school achievement and interest in STEM school subjects. *Soc. Psychol. Educ.* 20, 831–847. doi: 10.1007/s11218-017-9397-7

Bubany, S. T., and Hansen, J. I. C. (2011). Birth cohort change in the vocational interests of female and male college students. *J. Vocat. Behav.* 78, 59–67. doi: 10.1016/j.jvb.2010.08.002

Chambers, D. W. (1983). Stereotypic images of the scientist: the draw-a-scientist test. *Sci. Educ.* 67, 255–265. doi: 10.1002/sce.3730670213

Charles, M., and Bradley, K. (2009). Indulging our gendered selves? Sex segregation by field of study in 44 countries. *Am. J. Sociol.* 114, 924–976. doi: 10.1086/595942

Cheryan, S., Plaut, V. C., Handron, C., and Hudson, L. (2013). The stereotypical computer scientist: gendered media representations as a barrier to inclusion for women. *Sex Roles* 69, 58–71. doi: 10.1007/s11199-013-0296-x

Cundiff, J. L., Vescio, T. K., Loken, E., and Lo, L. (2013). Do gender–science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Soc. Psychol. Educ.* 16, 541–554. doi: 10.1007/s11218-013-9232-8

Cvencek, D., Meltzoff, A. N., and Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x

Deaux, K., and LaFrance, M. (1998). "Gender," in *The Handbook of Social Psychology*, eds D.T. Gilbert, S.T. Fiske, and G. Lindzey (New York, NY: Mc Graw-Hill), 788–827.

Eagly, A. H., and Wood, W. (2012). "Social role theory," in *Handbook of Theories of Social Psychology*, eds P. van Lange, A. Kruglanski, and E. T. Higgins (Thousand Oaks, CA: Sage), 458–476. doi: 10.4135/9781446249222.n49

Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychol. Women Q.* 18, 585–609. doi: 10.1111/j.1471-6402.1994.tb01049.x

Eccles, J. S., and Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in stem subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Finson, K. D. (2002). Drawing a scientist: what we do and do not know after fifty years of drawings. *Sch. Sci. Math.* 102, 335–345. doi: 10.1111/j.1949-8594.2002.tb18217.x

Fort, D. C., and Varney, H. L. (1989). How students see scientists: mostly male, mostly white, and mostly benevolent. *Sci. Child.* 26, 8–13.

FSO (2012). *Abschlüsse der universitären Hochschulen und Fachhochschulen: Basistabellen [University Diploma and University of Applied sciences Diploma: Tables]*. Available online at: https://www.bfs.admin.ch/bfs/de/home/statistiken/bildung-wissenschaft/bildungsabschluesse/tertiaerstufe-hochschulen/universitaere.assetdetail.5626778.html (accessed February 14, 2016).

FSO (2013). *On the Way to Gender Equality. Current Situation and Developments*. Neuchâtel: FSO. Available online at: https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/publikationen.assetdetail.349122.html (accessed October 10, 2018).

FSO (2019a). *Students of Applied Universities (Basis Table)*. Available online at: www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/tabellen.assetdetail.4762125.html (accessed November 11, 2018).

FSO (2019b). *High School Diplomas According to Main Subject, School Canton, Canton of Residence and Gender (Interactive Tables)*. Available online at: https://www.pxweb.bfs.admin.ch/pxweb/de/px-x-1503020200_102/px-x-1503020200_102/px-x-1503020200_102.px (accessed December 1, 2018).

George, D., and Mallery, P. (2016). *IBM SPSS Statistics 23 Step by Step: A Simple Guide and Reference*. New York, NY: Routledge. doi: 10.4324/9781315545899

Good, J. J., Woodzicka, J. A., and Wingfield, L. C. (2010). The effects of gender stereotypic and counter-stereotypic textbook images on science performance. *J. Soc. Psychol.* 150, 132–147. doi: 10.1080/00224540903366552

Gottfredson, L. S. (2002). Gottfredson's theory of circumscription, compromise, and self-creation. *Career Choice Dev.* 4, 85–148.

Gottfredson, L. S. (2005). "Applying Gottfredson's theory of circumscription and compromise in career guidance and counseling," in *Career Development and Counselling: Putting Theory and Research to Work*, eds S.D. Brown, and R.W. Lent (Hoboken, NJ: John Wiley and Sons), 71–100.

Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464

Hand, S., Rice, L., and Greenlee, E. (2017). Exploring teachers' and students' gender role bias and students' confidence in STEM fields. *Soc. Soc. Psychol. Educ.* 20, 929–945. doi: 10.1007/s11218-017-9408-8

Hannover, B., and Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learn. Instr.* 14, 51–67. doi: 10.1016/j.learninstruc.2003.10.002

Hartung, P. J., Porfeli, E. J., and Vondracek, F. W. (2005). Child vocational development: a review and reconsideration. *J. Vocat. Behav.* 66, 385–419. doi: 10.1016/j.jvb.2004.05.006

Heise, D. R. (1970). "The semantic differential and attitude research," in *Attitude Measurement*, ed. G.F. Summers (Chicago: Rand McNally), 235–253.

Herzog, W., Labudde, P., Neuenschwander, M. P., Violi, E., and Gerber, C. (1998). *Koedukation im physikunterricht. Schlussbericht [Coeducation in Physics Class. Final report]*. Universität Bern: Abteilung für Pädagogik/Abteilung für das Höhere Lehramt, Bern.

Hofstätter, P. R. (1973). *Einführung in die Sozialpsychologie [An Introduction in Social Psychology]*. Stuttgart: Kröner.

Huber, R. A., and Burton, G. M. (1995). What do students think scientists look like? *Sch. Sci. Math.* 95, 371–376. doi: 10.1111/j.1949-8594.1995.tb15804.x

Jones, B., Ruff, C., and Paretti, M. (2013). The impact of engineering identification and stereotypes on undergraduate women's achievement and persistence in engineering. *Soc. Soc. Psychol. Educ.* 16, 471–493. doi: 10.1007/s11218-013-9222-x

Kessels, U. (2005). Fitting into the stereotype: how gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *Eur. J. Psychol. Educ.* 20, 309–323. doi: 10.1007/BF03173559

Kessels, U., Rau, M., and Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *Br. J. Educ. Psychol.* 76, 761–780. doi: 10.1348/000709905X59961

Kessels, U., and Taconis, R. (2012). Alien or alike? How the perceived similarity between the typical science teacher and a student's self-image correlates with choosing science at school. *Res. Sci. Educ.* 42, 1049–1071. doi: 10.1007/s11165-011-9230-9

Koenig, A. M., and Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: observations of groups' roles shape stereotypes. *J. Personal. Soc. Psychol.* 107:371. doi: 10.1037/a0037215

Lane, K. A., Goh, J. X., and Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles* 66, 220–234. doi: 10.1007/s11199-011-0036-z

Makarova, E., Aeschlimann, B., and Herzog, W. (2016a). "Wenn Frauen in MINT-Studiengängen fehlen: mathematisch-naturwissenschaftlicher Unterricht und die Studienwahl junger Frauen [When women lack from STEM courses: teaching mathematics and Natural sciences and Young Women's Study Choices]," in *Berufsorientierung und Geschlecht [Vocational orientation and gender],* ed. H. Faulstich-Wieland (Weinheim: Juventa-Verlag), 39–57.

Makarova, E., Aeschlimann, B., and Herzog, W. (2016b). Ich tat es ihm gleich – Vorbilder junger Frauen mit naturwissenschaftlich-technischer Berufswahl ["I did it like he did" – role models of young women with science and technical vocational choices]. Berufs- und Wirtschaftspädagogik online. Spezial 12, Berufsorientierung im Lebenslauf – theoretische Standortbestimmungen und empirische Analysen [Career orientation in the life course - theoretical approaches and empirical analyses], 1–19.

Makarova, E., and Herzog, W. (2015). Trapped in the gender stereotype? The image of science among secondary school students and teachers. *Equal. Diver. Inclu. Int. J.* 34, 106–123. doi: 10.1108/EDI-11-2013-0097

McCullagh, P., and Nelder, J. (1989). *Generalized Linear Models, 2nd* Edn. Boca Ration, FL: CRC Press. doi: 10.1007/978-1-4899-3242-6

Mead, M., and Metraux, R. (1957). Image of the scientist among high-school students. *Science* 126, 384–390. doi: 10.1126/science.126.3270.384

Miller, D. I., Eagly, A. H., and Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: evidence from 66 nations. *J. Educ. Psychol.* 107, 631–644. doi: 10.1037/edu0000005

Miller, D. I., Nolla, K. M., Eagly, A. H., and Uttal, D. H. (2018). The development of children's gender-science stereotypes: a meta-analysis of 5 decades of US draw-a-scientist studies. *Child Dev.* 89, 1943–1955. doi: 10.1111/cdev.13039

Millon, T., Lerner, M. J., and Weiner, I. B. (2003). *Handbook of Psychology: Personality and Social Psychology*, Vol. 5. Mishawaka: Wiley and Sons.

Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., and Eccles, J. S. (2008). "Gendered high school course selection as a precursor of gendered careers: the mediating role of self-concept and intrinsic value," in *Gender and Occupational Outcomes*, eds H. M. G. Watt and J. S. Eccles (Washington, DC: American Psychological Association), 115–143. doi: 10.1037/11706-004

Nagy, G., Trautwein, U., Baumert, J., Köller, O., and Garrett, J. (2006). Gender and course selection in upper secondary education: effects of academic self-concept and intrinsic value. *Educ. Res. Eval.* 12, 323–345. doi: 10.1080/13803610600765687

Neuhaus, J., and Borowski, A. (2018). Self-to-prototype similarity as a mediator between gender and students' interest in learning to code. *Int. J. Gen. Sci. Technol.* 10, 233–252.

Nosek, B. A., Banaji, M. R., and Greenwald, A. G. (2002). Math= male, me= female, therefore math≠ me. *J. Pers. Soc. Psychol.* 83:44. doi: 10.1037/0022-3514.83.1.44

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. U. S. A.*106, 10593–10597. doi: 10.1073/pnas.0809921106

OECD (2006). *Evolution of Student Interest in science and Technology Studies. Policy Report*. Available online at: www.oecd.org/dataoecd/16/30/36645825.pdf (accessed October 10, 2018).

OECD (2012). *Closing* the *Gap. Act Now*. Paris: OECD Publishing. doi: 10.1787/9789264179370-en

OECD (2017). *The Pursuit of Gender Equality: An Uphill Battle*. Paris: OECD Publishing. doi: 10.1787/9789264281318-en

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.

Ramsey, L. R. (2017). Agentic traits are associated with success in science more than communal traits. *Pers. Individ. Dif.* 106, 6–9. doi: 10.1016/j.paid.2016.10.017

Ratschinski, G. (2009). *Selbstkonzept und Berufswahl. Eine Überprüfung der Berufswahltheorie von Gottfredson an Sekundarschülern [Self-concept and career choice. A review of Gottfredson's career choice theory among secondary school students]*. Münster: Waxmann.

Renfrow, D. G., and Howard, J. A. (2013). "Social psychology of gender and race," in *Handbook of Social Psychology*, eds. J. DeLamater, and A. Ward (Dordrecht: Springer), 491–531. doi: 10.1007/978-94-007-6772-0_17

Rosser, S. V. (2012). *Breaking Into the Lab: Engineering Progress for Women in Science*. New York, NY: NYU Press. doi: 10.18574/nyu/9780814776452.001.0001

Scherz, Z., and Oren, M. (2006). How to change students' images of science and technology. *Sci. Educ.* 90, 965–985. doi: 10.1002/sce.20159

Šorgo, A., Dojer, B., Golob, N., Repnik, R., Repolusk, S., Pesek, I., et al. (2018). Opinions about STEM content and classroom experiences as predictors of upper secondary school students' career aspirations to become researcher or teachers. *J. Res. Sci. Teach.* 55, 1–21. doi: 10.1002/tea.21462

Steffens, M. C., Jelenec, P., and Noack, P. (2010). On the leaky math pipeline: comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *J. Educ. Psychol.* 102:947. doi: 10.1037/a0019920

Taconis, R., and Kessels, U. (2009). How choosing science depends on students' individual fit to the "science culture." *Int. J. Sci. Educ.* 31, 1115–1132. doi: 10.1080/09500690802050876

Thomas, A. E. (2017). Gender differences in students' physical science motivation: are teachers' implicit cognitions another piece of the puzzle? *Am. Educ. Res. J.* 54, 35–58. doi: 10.3102/0002831216682223

Watt, H. M. G. (2006). The role of motivation in gendered educational and occupational trajectories related to maths. *Educ. Res. Eval.* 12, 305–322. doi: 10.1080/13803610600765562

WEF (2017). *The Global Gender Gap Report*. WEF. Available online at: https:www3.weforum.org/docs/WEF_GGGR_2017.pdf (accessed November 11, 2018).

Weinreich-Haste, H. (1981). "The image of science," in *The Missing Half: Girls and Science Education*, ed A. Kelly (Manchester, UK: Manchester University Press), 216–229.

Worell, J. (2001). *Encyclopedia of Women and Gender: Sex Similarities and Differences and the Impact of Society on Gender*. San Diego, CA: Academic Press.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 13 August 2019
doi: 10.3389/fpsyg.2019.01740

Sources of Male and Female Students' Belonging Uncertainty in the Computer Sciences

Elisabeth Höhne and Lysann Zander*

Division of Empirical Educational Research, Institute of Education, Leibniz Universität Hannover, Hanover, Germany

Belonging uncertainty, defined as the general concern about the quality of one's social relationships in an academic setting, has been found to be an important determinant of academic achievement, and persistence. However, to date, only little research investigated the sources of belonging uncertainty. To address this research gap, we examined three potential sources of belonging uncertainty in a sample of undergraduate computer science students in Germany ($N = 449$) and focused on (a) perceived affective and academic exclusion by fellow students, (b) domain-specific academic self-efficacy beliefs, and (c) perception of one's individual performance potential compared to that of fellow students in the field. Perceived affective and academic exclusion by fellow students and domain-specific academic self-efficacy beliefs were significant predictors of female students' uncertainty about belonging in computer science. The perception of one's individual performance potential in comparison to that of fellow students, however, was a relevant predictor of both male and female students' belonging uncertainty in computer science. Our findings imply an expanded view of the theoretical concept of belonging uncertainty that goes beyond mere concerns of social connectedness.

Keywords: belonging uncertainty, ability-related stereotypes, social identity, minority students, higher education, STEM, computer science, gender

Edited by:
Bernhard Ertl,
Universität der Bundeswehr München, Germany

Reviewed by:
Madeleine Bieg,
Universität Konstanz, Germany
Florian Berding,
University of Oldenburg, Germany

****Correspondence:***
Elisabeth Höhne
elisabeth.hoehne@iew.uni-hannover.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *09 January 2019*
Accepted: *12 July 2019*
Published: *13 August 2019*

Citation:
Höhne E and Zander L (2019) Sources of Male and Female Students' Belonging Uncertainty in the Computer Sciences. Front. Psychol. 10:1740. doi: 10.3389/fpsyg.2019.01740

INTRODUCTION

"I remember walking into one of the classes at Stanford and just deciding not to take the class because I was one of only three women there, and I just felt so intimidated." The experience that former co-president of Women in Computer Science at Stanford University Catherina Xu publicly expressed in 2017 is a feeling that students from stigmatized and underrepresented social groups frequently experience in academic settings. When minority group members question their fit in an educational environment, a state of belonging uncertainty can emerge and manifest, which has been found to adversely affect academic domain identification, achievement, persistence, and career aspirations (Cundiff et al., 2013; Woodcock et al., 2013; Walton et al., 2015; Höhne and Zander, in press). Although there is a growing body of research on the consequences of belonging uncertainty, to date, only little research has empirically investigated the sources of the feeling that "people like me" do not belong (Walton and Cohen, 2007). The current research addresses this gap and examines the perceived exclusion by fellow students, domain-specific academic self-efficacy beliefs, and the perception of one's individual performance potential in the domain compared to that of fellow students as sources of belonging uncertainty.

In the university context, one group that is likely to experience belonging uncertainty are female students in STEM (science, technology, engineering, and mathematics) domains, who, in Germany as well as in many other Western industrial nations, constitute the numerical minority in these stereotypically male-connoted domains (OECD, 2017). We chose the domain of computer science, because with a percentage of 18.35% it is currently one of the subjects with the lowest rate of female students in Germany (status winter term 2017/2018; Federal Statistical Office of Germany, 2018). Furthermore, in contrast to other STEM fields, computer science recently underwent a significant decrease in the number of female first-year students (−8.8% female students vs. −2.7% male students in winter term 2017/2018 as compared to the previous academic year; Federal Statistical Office of Germany, 2017).

Identifying and understanding the sources of belonging uncertainty is not only relevant to provide important cues for creating a stimulating and encouraging learning environment for females studying a traditionally male-dominated subject. In a broader sense, it is crucial to explore possible barriers and its antecedents to IT professions to prevent one of the fastest growing economic sectors from an intensification of skills shortage (European Institute for Gender Equality, 2017).

THEORETICAL FRAMEWORK

Belonging Uncertainty and Its Consequences in Academic Environments

The need to form and maintain positive and stable interpersonal relationships, i.e., the need to belong and to feel socially connected, is a fundamental human motivation (Baumeister and Leary, 1995). In educational environments, students' experienced sense of belonging has been empirically shown to positively affect achievement motivation, performance, well-being, and retention (Walton and Cohen, 2011; Good et al., 2012; Walton and Carr, 2012).

Negative competence-related stereotypes, such as the belief that women lack ability in quantitative fields (Smeding, 2012), can convey the message that people of certain social groups are less qualified, accepted, and valued. As a consequence, negatively stereotyped students who constitute a numerical minority in their respective academic domain, might doubt their belongingness, and experience a state of belonging uncertainty (Walton and Cohen, 2007, 2011; Good et al., 2012; Smith et al., 2013). Conceptually, this mental and emotional state takes the form of a hypothesis that guides individuals' perception and interpretation of various social contexts and is not restricted to specific social groups, settings or contexts (Walton and Cohen, 2007).

Noteworthy, belonging uncertainty differs from the phenomenon of stereotype threat, which describes the psychological predicament of being at risk to confirm a negative stereotype about one's own social group (Steele, 1997). Stereotype threat is evoked in specific, high-stake performance situations. Here, negatively stereotyped students underperform because they anticipate others' low expectations and negative performance-related feedback based on their group membership, ultimately confirming stereotypes they sought to avoid (for two meta-analyses see Nguyen and Ryan, 2008; Appel et al., 2015). To doubt one's belongingness to an academic domain, individuals do not need to experience a situation of evaluation, to anticipate or actually receive negative feedback on a specific task (Mallett et al., 2011). Instead, belonging uncertainty can manifest itself in the absence of a concrete performance situation and describes a more general concern that can give rise to the feeling that "people like me do not belong here" (Walton and Cohen, 2007, p. 83).

Uncertainty about one's belongingness in an academic environment can have a number of negative consequences. Experiences of rejection based on membership in a devalued group can lead individuals to anxiously expect future rejections, which, in turn, can lower their sense of well-being and negatively influence their relationships with peers, and professors (Mendoza-Denton et al., 2002). Female students were also shown to have greater anticipatory self-doubts about their abilities and expectations of unfairness when being preoccupied with concerns and expectations about rejection based on the social category of their gender (London et al., 2012). Continuously doubting one's belongingness in academic contexts can further contribute to actual decrements in intellectual performance (Mendoza-Denton et al., 2002) and lead stigmatized individuals to lower their identification with the scientific discipline (Spencer et al., 1999; Davies et al., 2002; Deemer et al., 2016). By disidentifying from the threatening domain, i.e., by removing the domain as a basis of self-evaluation as adaption strategy, negatively stereotyped students can uphold and maintain their feelings of self-worth (Steele, 1997; Spencer et al., 1999). Because the identification with a domain is an important predictor of career motivation (Schuster and Martiny, 2017), it can have detrimental effects on both personal and societal level when students of marginalized or stigmatized social groups disidentify with certain academic domains and, as a long-term consequence, abstain from scientific careers (Cundiff et al., 2013; Woodcock et al., 2013).

Potential Sources of Belonging Uncertainty

Although previous findings have demonstrated that belonging uncertainty is an important determinant of academic achievement and persistence, only little research has been dedicated to the investigation of its sources.

The theoretical conceptualization of belonging uncertainty as concern about the quality of one's social ties and connectedness (Walton and Cohen, 2007) draws on the hypothesis that human beings have a pervasive need to belong that is reflected in their desire to have positive relationships with others (Baumeister and Leary, 1995). As social beings we rely on interdependencies with conspecifics, wherefore the perception of not being socially accepted and integrated is experienced as aversive. Within social pain theory, the aversive emotional state of social exclusion is even described as unpleasant as the experience of suffering physical pain, because it signals the probability of being socially excluded or isolated, which constitutes an evolutionary disadvantage (Eisenberger et al., 2003; MacDonald and Leary, 2005). Among the negative consequences of social exclusion are cognitive impairments as well as significant mental and physical detriments, including higher rates of morbidity and mortality (Berkman and Syme, 1979; Baumeister et al., 2002; Uchino, 2006; Cohen and Janicki-Deverts, 2009).

In challenging, achievement-oriented and competitive environments, people are suggested to be sensitive to the quality of their social bonds, and members of negatively stereotyped social groups are found to be even more susceptible to feelings of social belonging uncertainty (Walton and Cohen, 2007). The results of an experimental study by Walton and Cohen (2007) substantiate this assumption: a manipulation leading individuals to believe they would only have few friends in the domain of computer science decreased Black but not White students' sense of belonging to that domain. Thus, only individuals afflicted with a threatened social identity, i.e., the part of the self-concept that is based upon social group membership (Tajfel and Turner, 1986), seem to be vulnerable to subtle situational cues that signal a lack of social connectedness (Murphy et al., 2007; Murphy and Taylor, 2012). We therefore expected students who perceived that they were excluded from social activities and academic exchange with fellow students to be more uncertain about their belonging within the domain of computer science – but that this relationship would be stronger for female students.

A slightly modified manipulation targeting academic ability in a quantitative domain, however, negatively affected female students' sense of belonging. Here, students were asked to list either two skills or eight skills they had in the domain of computer science, with the result that female but not male students rated their social fit to that domain lower in the eight skills than in the two skills condition (Walton and Cohen, 2007). To the extent that the experienced difficulty of the task was presumed to increase with the number of skills that had to be generated (see also Schwarz et al., 1991; Hermann et al., 2002), the findings indicate that female students were more sensitive to information relevant to their insecurity, i.e., information about their quantitative ability, eventually leading to larger decrements in their perceived sense of belonging. Belonging uncertainty may therefore not only grow with doubts about one's social connectedness, but also with doubts about one's abilities, and competencies in a discipline. Experimental research on induced feelings of belonging, either in a social or in an academic domain related to university, further supports the idea that there is a difference between social and academic belonging. A study by Skourletos et al. (2013) experimentally manipulated feedback on a measure students completed and found that minority students' performance on an IQ test was significantly higher when they were told they had the academic potential and ability to do well scholastically than when they were told that they had social potential. Because the results indicated that the performance deficit in minority students caused by negative stereotypes could not be remedied by an attributed social potential and by telling students that people with similar scores were involved in various social organizations at university, Skourletos et al. (2013) suggested a difference between academic and social belonging. This also corresponds to the results of a study by Lewis and Hodges (2015), who found a negative correlational relationship between social belonging and a measure of ability uncertainty. Therefore, it seems to take both positive social interactions within a domain and a sense of relative fit regarding one's academic competencies in order to feel one belongs.

To approach minority students' uncertainty about their academic abilities, we focus on academic self-efficacy beliefs, which have been linked to a wide range of desirable scholastic outcomes, such as students' achievement and college retention (Robbins et al., 2004; Valentine et al., 2004; Parker et al., 2014). Within social cognitive theory, self-efficacy beliefs refer to a person's confidence in his or her ability to accomplish certain tasks (Bandura, 1997), and several career-related decisions are influenced by our judgments of self-efficacy (Lent et al., 1994). While people tend to pursue tasks and approach domains in which they feel capable, they avoid those which they believe exceed their abilities (Bandura, 1997). With regard to our target group, much research has shown that females tend to have lower levels of self-efficacy in quantitative fields even when they perform equally well or when they outperform their male counterparts (for a meta-analysis see Huang, 2013). Further, previous experimental studies have indicated the adverse influence of stereotype activation on self-efficacy beliefs (Hoyt and Blascovich, 2007; Lerdpornkulrat et al., 2012; Wang et al., 2017). Because female students have to contend with a negative stereotype about their quantitative ability, we argue that the discrepancy between females' confidence in their ability, and their actual performance can be attributed to ability-related stereotypes that are firmly grounded in society. It is therefore plausible that low academic self-efficacy beliefs are a source of female students' feelings of belonging uncertainty, which we expected in the present study.

Furthermore, research has shown that students' competence beliefs are influenced by the frames of reference that they use to evaluate themselves. According to the "Big-Fish-Little-Pond-Effect" (BFLPE; Marsh and Parker, 1984), students compare their academic ability with that of their classmates – a process that affects the development of students' academic self-concept (Marsh et al., 2007). Thus, a student among low-achieving classmates would show a higher academic self-concept (a "big fish in a little pond") than an equally able student among high-achieving classmates (a "little fish in a big pond") – an effect that has been replicated in different academic domains and across a large number of culturally and economically diverse countries (Seaton et al., 2009).

Even though the concepts of academic self-efficacy and academic self-concept both describe perceptions of the self in academic contexts, they differ in that self-efficacy refers to an individual's convictions to be able to succeed in specific academic tasks (Bandura, 1997), whereas self-concept refers to the perception of one's general academic abilities and skills in a domain (Marsh, 1987; Trautwein et al., 2009). For example, the expectation that one can succeed in a study-related task (e.g., to pass an exam in computer science) is an efficacy judgment, however, it is not a judgment of whether one is competent in this domain in general (e.g., to be a successful computer scientist or a "computer science person"). Whereas self-concept is formed through experiences with and interpretations of one's social environment (Shavelson et al., 1976), frame-of-reference effects are not central to self-efficacy beliefs, and should be largely eliminated in students' responses to self-efficacy measures (Marsh et al., 2018).

Although academic self-concept has empirically been found to have consistent reciprocal effects with achievement and educational attainment (Marsh and O'Mara, 2006), it does not necessarily accurately reflect actual achievement. With a view to females in STEM subjects, previous research found that female students, on average, have lower academic self-concepts in mathematics- and science-related domains – even when they perform on the same level as their male peers (Ludwig, 2010; OECD, 2015). A study by Ertl et al. (2017) could further demonstrate that stereotypes about females' interests, abilities, and need for conformance in STEM directly affect the academic self-concept of female STEM students. Moreover, while upward comparisons in the context of attainable achievement can serve as inspiration (Lockwood and Kunda, 1997) and benefit performance (Blanton et al., 1999), social comparisons in the presence of negative stereotypes seem to work via different psychological mechanisms. Here, upward comparisons with in-group members are enhancing, because the superior other's performance is challenging the negative ability-related stereotype about one's own social group, whereas the opposite holds in case of upward comparisons with out-group members, who constitute most of the salient targets of comparison in historically homogeneous environments (Blanton et al., 2000). Given the negative ability-related stereotype about females and their numerical underrepresentation in the male-dominated STEM domains, we argue that females' comparisons of their academic ability with that of their – predominantly male – classmates can cause feelings of belonging uncertainty in terms of academic fit.

The Present Research

Comprehensively understanding the factors that explain female students' lower academic domain identification and retention in STEM-related subjects is crucial to remedying growing gender disparities in computer science. To gain a better understanding of the situation of female students in computer science and the concept of belonging uncertainty, the present study examined potential sources of male, and female students' uncertainty about belonging in the domain of computer science. In light of relevant previous findings, we expected that (a) the perceived affective and academic exclusion by fellow students, (b) domain-specific academic self-efficacy beliefs, and (c) the perceived individual performance potential in comparison to that of fellow students in computer science would predict female but not male students' sense of belonging uncertainty in the academic domain of computer science.

MATERIALS AND METHODS

Participants

A total of 453 computer science students participated in the study. Four students did not provide any information on our variables of interest, including their gender, and were excluded from the analyses. Our final sample consisted of 449 students (345 male, 104 female) with an average age of 22.03 years (SD = 4.36). Participation rate across both time points of assessment was high with an overall percentage of 94.47%.

Procedure

The study was conducted in winter term 2016/2017 and 1 year later at the Institute of Computer Science at a large German university (more than 30.000 enrolled students), from which permission was obtained beforehand. Questionnaires were distributed in the mandatory functional programming classes of the first-year students, which have to be attended at least 80% of the time in order to pass. Moreover, students have to actively take part in the tutorial classes (max. 30 students), e.g., by handing in weekly exercises that they worked on in groups of two. Data was collected in 26 tutorial groups with an average size of 18.00 students (SD = 4.89). Each of the tutorial groups was assessed at two time points: at the beginning of the lecture period in week two or three (T1), and four weeks later (T2).

All students in the functional programming classes were asked to participate in a study on their "first impressions and experiences at university" and were told that the aim of the study was to "improve the conditions of studying." Students were also informed about their voluntariness of participation, assured of their anonymity, and instructed that all of their data would be kept confidential and be used for research purposes only. In addition, students gave their written consent on top of the questionnaire at each time point of assessment and were informed of their right to withdraw their participation at any time of the study without giving any reason. All participants generated a personal code, allowing us to match the questionnaires of each one person while ensuring an anonymous data processing. Once students gave their consent, research assistants emphasized that there were no right or wrong answers and encouraged participants to answer in whatever way seemed right for them.

The completion of the paper-and-pencil questionnaire took between 15 and 20 min. At the first assessment, students participated without receiving any reward or compensation. With the objective of increasing the motivation and incentive to participate in the second assessment, students received sweets and could participate in a raffle in which they could win a book voucher. In order to take part in the raffle and to be informed about the results of the study, students could write their email address on a separate list after the assessment.

Measures

Belonging Uncertainty

An adapted version of a measure by Walton and Cohen (2011) was used to assess students' subjective level of uncertainty about their belonging within computer science. The original scale of Walton and Cohen (2011) consisted of three items, however, a factor analysis found that one of the items loaded weakly on the common factor. The remaining 2-item scale demonstrated a good internal consistency (α = 0.820), wherefore we decided to apply it in our study. Because students' degree of uncertainty about their social belonging within the particular domain of computer science was of interest in our study, rather than their belonging to the entire college, the content of the two-item measure was adapted to the subject of study: (1) "Sometimes I feel that I belong to this study program, and sometimes I feel that I don't belong to this study program[1]" and (2) "When things don't go well, I often think that maybe I don't belong to this study program." Students indicated their agreement on a 5-point Likert scale (1 = *strongly disagree,* 5 = *strongly agree*), with greater values reflecting a higher level of belonging uncertainty. Both items were found to be internally consistent (α = 0.712), summed and averaged into a composite score. The intraclass correlation (ICC) for students within tutorial groups was determined to be 0.001.

Affective and Academic Social Exclusion

Students' perceived exclusion from non-academic social activities with fellow students (affective exclusion) and subject-related exchange with fellow students (academic exclusion) was assessed using a self-developed scale consisting of four items (e.g., "Sometimes I have the feeling that other students meet privately and I am not included," "I have already noticed that other students engage in subject-related exchange and I am not included"; for the full scale see section "**Supplementary Material**"). The reason why we used a self-developed measure was that there was no adequate measure for our targeted population in terms of age and the university context which encompassed both students' perceived social and academic exclusion. All items used a 5-point Likert response scale (1 = *strongly disagree*, 5 = *strongly agree*) and formed a reliable scale (α = 0.890).[2]

Domain-Specific Academic Self-Efficacy

Students' confidence in their ability to succeed in study-related tasks, even when confronted with difficulties or under challenging circumstances, was assessed using an adapted version of a German scale by Jerusalem and Schwarzer (1986). Whereas the original scale consists of seven items altogether, a shortened two-item version was applied here: (1) "I am confident that I have the competencies to perform well in this subject" and (2) "I can cope with difficult situations and challenges in my studies when I try hard." The corresponding items used a 5-point Likert response scale (1 = *strongly disagree,* 5 = *strongly agree*), for which reliability analysis revealed a sufficient internal consistency (α = 0.700).

Performance Potential Compared to Fellow Students

To tap into the social comparison impression that has been shown to influence students' academic self-concept, an adapted measure by Walton and Cohen (2007) was used to assess students' perception of their individual performance potential compared to the potential of their fellow students. In contrast to the original measure, students were prompted to think about their fellow students before being asked to evaluate their individual potential to succeed in their studies compared to their peers on a percentile scale: "I have more potential than . . . of the students in this subject" (10% = *more potential than 10% of the students,* 90% = *more potential than 90% of the students*, in steps of 10%).

Academic Performance

Since our sample primarily consisted of first-year university students (71.9%) who had not received any grades at university yet, we assessed a proxy variable of previous academic performance of the respondents by asking them to report their average grade obtained in the German school-leaving examination.

Socio-Demographic Data

Participants were asked to provide information on their gender and age. To prevent priming effects based on gender, which could impact students' reports of their sense of belonging uncertainty (Mallett et al., 2011), socio-demographic data were assessed at the very end of the questionnaire.

Statistical Analyses

Data analyses, if not stated differently, were run using Mplus version 8.1 (Muthén and Muthén, 1998–2017). At first, descriptive statistics and bivariate correlations for all variables of interest were calculated. In addition, mean differences and standardized mean differences[3] between male and female students were computed. Prior to running our main analyses, we conducted Little's MCAR test (Little, 1988) within SPSS's Missing Value Analysis option (version 25.0; IBM Corp, 2017) to examine missing data patterns in our sample. This test is implemented as a chi-squared test with the null hypothesis that cases of missing data are missing completely at random (Little and Rubin, 1989). We then estimated missing values in Mplus using full information maximum likelihood estimation (FIML), which has been proven to be superior to other missing-data techniques, such as list- or pairwise deletion, mean substitution or last observation carried forward, with respect to model estimation, bias, and efficacy (Schafer and Graham, 2002; Peugh and Enders, 2004). In order to avoid listwise deletion of individuals with missing data on *x*-variables, independent variables were treated as dependent variables within Mplus as a result of specifying the means and variances of the independent variables (Hox et al., 2015). Because students (level-1 unit) were nested in tutorial groups (level-2 unit), which may violate the assumption of independent observations within regression analyses (e.g., Nimon, 2012; Snijders and Bosker, 2012), we used the TYPE = COMPLEX command in Mplus to take into account the hierarchical data structure and to adjust the standard errors. The multiple linear regression was conducted using a Robust Maximum Likelihood (MLR) estimator. In the main model, we regressed belonging uncertainty at T2 on the T1

[1] Although double-barreled items like this one should generally be avoided because respondents could hold different views about different topics within one item (Simms and Watson, 2007), the authors specifically aimed at measuring students' doubts or uncertainty about their belonging within computer science (cf., Walton and Cohen, 2007).

[2] Noteworthy, Cronbach's alpha is an estimate of reliability under rather restrictive assumptions, wherefore the coefficient almost always underestimates true reliability (Bollen, 1989; Sijtsma, 2009; Eisinga et al., 2013).

[3] Standardized mean differences were calculated using Cohen's d statistic, with small, medium, and large effect sizes (ES) being equivalent to d-values of 0.20, 0.50, and 0.80, respectively (Cohen, 1988). ES were calculated using the following formula: $\Delta = \frac{\beta_1}{\sigma_e}$ (Tymms et al., 1997).

variables perceived affective and academic exclusion, domain-specific academic self-efficacy, perceived relational performance potential as well as each predictor's interaction with gender, while controlling for belonging uncertainty at T1 and students' previous academic performance. Significant interactions were visualized using the web-based data visualization tool interActive (McCabe et al., 2018). Because our analyses focused on effects within persons and because we expected the relevant reference group for participants to be students in their tutorial group, rather than all students in their study program[4], all level-1 variables, except the categorical variable of gender, were entered group-mean centered into the model. Accordingly, slopes are interpreted as the increase in the criterion variable associated with one unit increase in the predictor variable – relative to the tutorial group's mean (for the implication of different centering choices in terms of interpretation see Park, 2008).

RESULTS

The descriptive statistics for our dependent and independent variables are shown in **Table 1**. Means and standard deviations are presented for the total sample as well as separately for male and female students. In addition, gender differences on all relevant variables were examined using linear regression analyses with a dummy variable taking a value of zero for male students and one for female students. Significant mean differences were found for belonging uncertainty at both time points of assessment, domain-specific academic self-efficacy, the perceived performance potential in relation to fellow students, and previous academic achievement. Female students, on average, reported higher levels of belonging uncertainty in computer science than male students both at T1 ($B = 0.515$, $p \leq 0.001$, $d = 0.445$) and T2 ($B = 0.443$, $p \leq 0.01$, $d = 0.402$). Regarding the perceived affective and academic exclusion by fellow students, no significant mean difference between male and female students was found ($B = -0.022$, $p = 0.840$, $d = 0.020$). By contrast, male students reported higher levels of academic self-efficacy in computer science ($B = -0.245$, $p \leq 0.01$, $d = 0.341$), and evaluated themselves, unlike female students, as above average regarding their own performance-related potential in comparison to that of other students in the field ($B = -0.949$, $p \leq 0.001$, $d = 0.480$). Interestingly, female students, however, reported higher school-leaving examination grades and thus, a better academic performance in high school than male students ($B = -0.318$, $p \leq 0.05$, $d = 0.189$).

In **Table 2**, bivariate correlations are shown. All predictor variables, except students' previous academic performance, which we controlled for in the subsequent regression analyses, correlated significantly with our criterion. While social exclusion and gender (males = 0, females = 1) positively correlated with belonging uncertainty, negative correlations were obtained for domain-specific academic self-efficacy, and perceived potential in relation to fellow students. Since there were weak to moderate correlations between some of the explanatory variables, multicollinearity was tested by means of variance inflation factors (VIFs), applying a cut-off value of 10 (Bühner and Ziegler, 2009). VIFs were examined within SPSS (version 25.0; IBM Corp, 2017), and based on a multiple regression analysis of belonging uncertainty on all independent variables. With the lowest VIF-score being 1.026 and the highest being 1.714, no significant inflation of standard errors due to non-orthogonality among the predictors was indicated.

According to Little's MCAR test, which showed that missing data in our sample ranged from 2% to 38% with an overall proportion of 14%, these data points were missing completely at random ($\chi^2 = 76.72$, $df = 68$, $p = 0.219$), indicating that the probability of missingness does not depend on any observed or missing values.

To test the core assumption of our research that perceived exclusion by fellow students, domain-specific academic self-efficacy and perceived performance potential compared to others are relevant predictors of belonging uncertainty within computer science, while controlling for students' previous academic achievement and the initial level of belonging uncertainty, a multiple linear regression analysis was conducted (see **Table 3**).

In line with our first hypothesis, the interaction between students' perceived affective and academic exclusion by fellow students and gender was a significant predictor of belonging uncertainty ($\beta = 0.087$, $p = 0.038$). Although there were no significant mean differences in the perceived affective and academic exclusion between male and female computer science students as depicted in **Table 1**, it was found to be a significant predictor of female students' belonging uncertainty. Thus, female students with higher values of perceived social exclusion relative to their group mean experienced greater uncertainty about their belonging than their male peers (see **Figure 1A**). Consistent with our second hypothesis, the interaction between students' domain-specific academic self-efficacy and gender was found to be predictive of the uncertainty about belonging in the domain of computer science ($\beta = -0.133$, $p = 0.019$). As expected, academic self-efficacy was a more relevant predictor of belonging uncertainty for female than for male students, i.e., female students with lower self-efficacy beliefs in relation to the average degree of academic self-efficacy beliefs in their respective tutorial group were more uncertain about their belonging in computer science than male students (see **Figure 1B**). Contrary to our expectations, the link between the perceived performance potential in comparison to fellow students and feelings of belonging uncertainty did not differ as a function of gender ($\beta = 0.017$, $p = 0.742$). Thus, the slopes of the regression lines did not differ significantly between male and female students. Rather, a significant main effect of the relative potential ($\beta = -0.147$, $p = 0.005$) indicated that the assessment of one's capabilities through social comparisons with others in the same academic domain is a relevant predictor of both male and female students'

[4]We expected the relevant reference group for participants to be students in their tutorial group, rather all students in their study program, because we conducted our study in the tutorial classes of the first-year students. Hence, we expected students to be not yet familiar with many of their fellow students. In contrast to the teacher-centered teaching in lectures, students have to actively take part in the smaller and compulsory tutorial classes, e.g., by handing in weekly exercises that they worked on in groups, thereby getting to know each other faster. We therefore expected students to assess e.g., their performance potential and their affective as well as academic exclusion in relation to the students in their tutorial group.

TABLE 1 | Means, standard deviations, and mean comparisons by gender of the dependent and independent variables.

		Belonging uncertainty T2	Social exclusion T1	Academic self-efficacy T1	Relative potential T1	Belonging uncertainty T1	Academic performance T1
	N	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**
Total	449	2.76 (1.11)	2.88 (1.08)	3.95 (0.73)	5.04 (2.01)	2.82 (1.16)	2.37 (1.67)
Males	345	2.66 (1.07)	2.89 (1.09)	4.00 (0.71)	5.25 (1.79)	2.71 (1.16)	2.44 (1.90)
Females	104	3.07 (1.16)	2.86 (1.05)	3.75 (0.74)	4.24 (2.00)	3.20 (1.15)	2.12 (0.63)
B (*SE*)		0.443 (0.17)	−0.022 (0.11)	−0.245 (0.10)	−0.949 (0.21)	0.515 (0.15)	−0.318 (0.14)
Sig.		0.010**	0.840	0.015*	0.000***	0.001***	0.019*

*All values were estimated using Mplus and full information maximum likelihood estimation (FIML). Standard errors were adjusted for the hierarchical data structure. Mean values (M), standard deviations (SD), and standard errors (SE) were rounded to two decimal places. Regression coefficients (B) were rounded to three decimal places. Gender: 0 = male, 1 = female. *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001.*

TABLE 2 | Correlations of the dependent and independent variables.

	1	2	3	4	5	6	7	VIF
1 Belonging uncertainty T2	1	0.176**	−0.475***	−0.428***	0.681***	0.182	0.167*	–
2 Social exclusion		1	−0.124*	−0.052	0.207***	0.037	−0.008	1.051
3 Academic self-efficacy			1	0.545***	−0.467***	−0.031	−0.142*	1.714
4 Relative potential				1	−0.401***	−0.083*	−0.199***	1.521
5 Belonging uncertainty T1					1	0.062*	0.185***	1.537
6 Academic performance						1	−0.079**	1.026
7 Gender							1	1.086

*N = 449. Values were estimated using Mplus and full information maximum likelihood estimation (FIML). Standard errors were adjusted for the hierarchical data structure. VIF = variance inflation factor of the independent variables (variables 2–7; results were estimated using SPSS). Values were rounded to three decimal places. Gender: 0 = male, 1 = female. *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001.*

uncertainty about belonging. The overall model explained a total of 51.1% of the variance in the outcome variable, with Cohen's f^2 statistic yielding an effect size estimate of 1.04, which corresponds to a large effect (Cohen, 1988).

DISCUSSION

Female students in computer science, like in many other STEM domains, still constitute a numerical minority. Thus, an important goal for research in this field is to study the factors that prevent students from meeting their academic potential. A well-established predictor of minority students' academic underachievement is the worry to not "fit into" the respective academic environment: belonging uncertainty. While much research has focused on the outcomes of belonging uncertainty, we sought to deepen our understanding of the sources of belonging uncertainty. We examined the role of (a) the perceived affective and academic exclusion by fellow students, (b) domain-specific academic self-efficacy beliefs, and (c) the perceived individual performance potential in comparison to that of fellow students as possible predictors of female students' belonging uncertainty. In doing so, this study adds to the literature by extending our conceptual understanding of the uncertainty about belonging.

TABLE 3 | Multiple linear regression for variables at T1 predicting belonging uncertainty at T2.

Explanatory variables	*B*	*SE*	β	*p*
Social exclusion	−0.015	0.045	−0.014	0.744
Academic self-efficacy	−0.099	0.082	−0.063	0.233
Relative potential	−0.084	0.030	−0.147	0.005**
Belonging uncertainty	0.540	0.036	0.561	0.000***
Academic performance	−0.050	0.123	−0.074	0.724
Gender	0.029	0.129	0.011	0.823
Social exclusion X gender	0.204	0.096	0.087	0.038*
Academic self-efficacy X gender	−0.424	0.175	−0.133	0.019*
Relative potential X gender	0.019	0.058	0.017	0.742

*N = 449. All values were estimated using Mplus and full information maximum likelihood estimation (FIML). Standard errors were adjusted for the hierarchical data structure. Values were rounded to three decimal places. SE = standard error of B. Gender: 0 = male, 1 = female. *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001.*

Consistent with our expectations, we found that perceived affective and academic exclusion by fellow students increased female but not male students' belonging uncertainty in computer science. This finding conforms to the assumption that members of underrepresented and negatively stereotyped groups, as in the case of female students in STEM, are particularly sensitive to the quality of their social relationships in competitive academic environments (Walton and Cohen, 2007). Although male and female students did not differ in the extent to which they felt excluded from non-academic social activities and subject-related exchange, the subjective experience of being socially excluded was a relevant explanation for female students' doubts whether they would belong. Thus, our results indirectly support

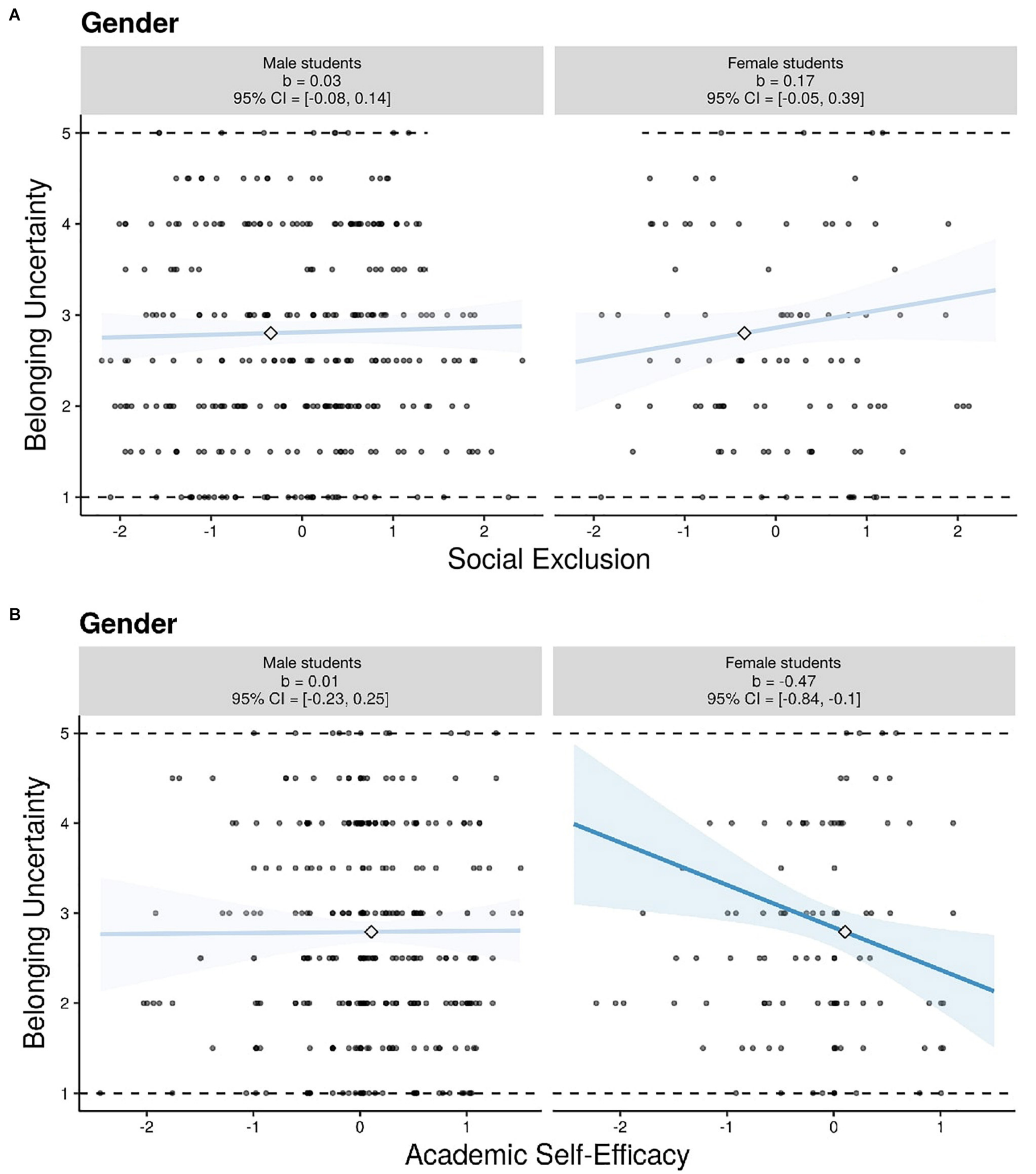


FIGURE 1 | (A) Simple slopes graphs depicting the relationship between social exclusion and belonging uncertainty. Academic self-efficacy, relative potential, the initial level of belonging uncertainty, academic performance, gender, academic self-efficacy X gender, and relative potential X gender were entered as covariates. **(B)** Simple slopes graphs depicting the relationship between academic self-efficacy and belonging uncertainty. Social exclusion, relative potential, the initial level of belonging uncertainty, academic performance, gender, social exclusion X gender, and relative potential X gender were entered as covariates. In each panel, simple slopes are displayed for both levels of the moderator. Each graphic shows the computed 95% confidence region (shaded area), the observed data (gray circles), the minimum and maximum values of the outcome (dashed horizontal lines), and the crossover point (diamond). Regression coefficients (*b*) differ from the ones depicted in **Table 3** due to full information maximum likelihood estimation (FIML) within Mplus. CI = confidence interval. Figures were produced using the interActive data visualization tool (McCabe et al., 2018).

previous research on the cues hypothesis, which holds that subtle situational cues, such as the numerical representation of a social group, can signal a lack of social connectedness and trigger experiences of social identity threat among stereotyped groups (Murphy et al., 2007).

Further and in line with our predictions, we found domain-specific academic self-efficacy beliefs to be a significant predictor, again, of female but not male students' uncertainty about belonging. Previous experimental studies have indicated the negative influence of entity beliefs and negative ability-related stereotypes on self-efficacy beliefs (Hoyt and Blascovich, 2007; Lerdpornkulrat et al., 2012; Wang et al., 2017) and our results seem to confirm such a link. Female students with a threatened social identity in the context of a male-dominated STEM subject appear to be more susceptible to low academic self-efficacy beliefs to the effect that they constitute a relevant source of their sense of belonging uncertainty.

Unexpected, however, was that the extent to which students felt they had lower academic potential than their fellow students was not more important for female students but rather a relevant predictor of both male and female students' uncertainty about belonging in computer science. Both, male and female students' upward social comparisons regarding their academic potential in computer science appear to affect their belonging uncertainty in a similar vein, regardless of the relative number of in- or out-group members. The presence of negative stereotypes about females' abilities in the field may explain differences in male and female assessment of their relative performance potential (cf., Ertl et al., 2017). Yet, our pattern of results suggests that being a member of a potentially stereotyped minority does not imply that females weight their perceived potential more strongly than males when judging their belonging to the specific academic domain.

Belonging uncertainty has frequently been defined as concern about the quality of one's social ties or about whether one would be fully included in positive social relationships (e.g., Walton and Cohen, 2007, 2011). Our findings that domain-specific academic self-efficacy beliefs as well as the perceived performance potential in comparison to peers are relevant sources of students' uncertainty about belonging expand this conceptual view. This is in line with research by Walton and Cohen (2007) who found that female students' sense of belonging was negatively affected by a manipulation that made them believe they had only a few of the skills required in computer science. It appears that belonging uncertainty is rooted in both students' doubts about their social connectedness within an academic domain and concerns about whether they have the abilities to succeed in that domain.

Experimental research has previously suggested a difference between academic and social belonging (Skourletos et al., 2013). But rather than sources of one concept, academic and social belonging were considered two different types of belonging each with a discriminative power in the prediction of achievement-related variables. Similarly, Lewis and Hodges (2015) developed a domain-specific measure of ability uncertainty and found negative correlations with social belonging and academic self-efficacy. In addition, the authors found ability uncertainty to predict academic outcomes, such as students' intent to persist in their psychology or linguistics major. Again, ability fit and social fit within a particular academic domain were conceptualized as two separate types of belonging individuals might question. Interestingly, Lewis and Hodges (2015) examined a sample of psychology and linguistics students who were predominantly female and Caucasian and thus, did not have to contend with negative stereotypes about their intellectual ability within their domains. With the particular group of computer science students, in which females are a stereotyped minority, our results suggest that belonging uncertainty is in fact a result of their lower confidence in their abilities to overcome academic challenges, their feeling of being less competent than their – primarily male – peers, and their perception of being excluded from social and academic exchange with fellow students.

In summary, the results of the present research suggest that conceptualizing belonging uncertainty as the concerns about one's social connectedness in an academic domain, without incorporating stereotyped students' concerns about their academic abilities, and vice versa, might be an incomplete understanding of the uncertainty about belonging in an educational environment.

Although further research is needed to substantiate our findings, it is possible to consider practical implications for educational institutions in terms of how to organize the integration of minority students into their study programs and foster belonging. First, our results show that belonging uncertainty, a strong predictor of students' achievement, persistence, and career aspirations (Cundiff et al., 2013; Woodcock et al., 2013; Walton et al., 2015; Höhne and Zander, in press), is catalyzed by perceived affective and academic social exclusion, particularly among female students. Therefore, study programs may wish to create more opportunities for male and female students to exchange with their peers – both formally, e.g., in specific group learning arrangements in seminars and tutorials, as well as informally, after classes. These activities may be especially relevant at the beginning of students' university careers in order to prevent increasing disparities. Second, our findings suggest that academic self-efficacy beliefs are particularly important for female students' sense of belonging. Given that successful models, credible social persuasion, experiences of mastery, and positive affective states are important sources of self-efficacy (Bandura, 1997; Bong and Skaalvik, 2003), pedagogical support and mentoring from lecturers and graduate students providing these sources may be a promising route to improve female students' feelings of belonging in male-dominated STEM domains.

Limitations and Future Directions

Despite the contribution of the present research to the understanding of female students' uncertainty about belonging in the male-dominated STEM subject of computer science, our study has some limitations that need to be addressed in future research.

To begin with, our study is not experimental in nature, and thus, conclusions about the causal relationship between students' perceived affective and academic exclusion by fellow students, their reported academic self-efficacy in computer science, their perceived performance potential in comparison to peers, and their uncertainty about belonging in computer science should be made with caution. This being said, the fact that our predictor variables were measured prior to belonging uncertainty, that students' perceptions about their social exclusion and academic competencies at the beginning of the semester predicted belonging uncertainty in the midst of the semester, and that we controlled for the initial sense of belonging uncertainty suggest the directionality of the effect. The results of an experimental study by Walton and Cohen (2007) point in the same direction. Here, students were made to believe that they had only limited computer science skills, which lowered female students' sense of belonging. This reasoning is further supported by a study in which Walton et al. (2015) applied an intervention to mitigate doubts about social belonging in engineering. A core element of the intervention was the implication of normality when experiencing doubts. Through written experiences of former first-year students, participants were told that almost all students had worries about fitting in and being accepted during their first year in college but that these concerns would dissipate with time. The authors found that the intervention helped female students to better integrate in their engineering study program and establish friendships with male students (Walton et al., 2015). This suggests that social doubts precede belonging uncertainty rather than resulting from it. Future experimental research or longitudinal designs with cross-lagged analyses could help to further clarify the interrelation between these constructs.

Second, the present study is limited in that it only examined a sample of undergraduate computer science students. It would be an important next step to examine whether this pattern of findings can be applied to female students in other STEM domains and to other social groups that have to contend with negative ability-related stereotypes and who constitute a numerical minority in their respective academic domain. Moreover, longitudinal studies over a longer period of time are needed to investigate the temporal stability of our findings. With regard to the subject of computer science, we decided to assess students at the beginning and in the midst of their first semester because of the high dropout rates in this subject at German universities (Heublein and Schmelzer, 2018) and because previous research could show that doubting one's belongingness in an academic context itself is a predictor of students' persistence and dropout intentions, respectively (Cundiff et al., 2013; Woodcock et al., 2013; Höhne and Zander, in press). Therefore, and because we wanted to investigate the sources of belonging uncertainty, we expected a certain proportion of the students with a high uncertainty about belonging to already have dropped out by the end of the semester, preventing us from obtaining insight into the psychological experiences of this student group. However, studies over a longer period of time would not only advance our theoretical understanding of the concept of belonging uncertainty, but also inform interventions about when they can exert maximal effects on members of affected social groups in different academic domains. Additionally, if we assume that the sources of belonging uncertainty hold for other academic settings and stereotyped social groups, then a positive sense of belonging could stem from the same sources, but with different signs. Considering empirical support for the phenomenon of stereotype lift, i.e., a performance boost caused by downward comparisons with members of a negatively stereotyped outgroup (for a meta-analysis see Walton and Cohen, 2003), it is plausible that similar mechanisms underlie the development of positive belongingness.

Third, although our finding that belongingness to an academic domain is depending on the perception of one's social and ability fit is in line with the results of a study by Lewis and Hodges (2015), who found a significant correlation between ability uncertainty and social belonging, there may be other components that add to the full picture of belonging uncertainty. For example, research in the field of computer science could demonstrate that girls show an increased interest and sense of belonging when introducing them to physical environments that were not considered stereotypical of computer science (Master et al., 2016). Classrooms and other physical environments that signal a stereotypical image of computer science and the people that represent that domain might therefore result in an adverse balance of self-to-prototype matching in female students and, in turn, serve as source of feelings of belonging uncertainty in that domain. Future research is needed to systematically study how adaptive processes of group formation can be initiated and stimulated through learning arrangements and how institutional norms can further contribute to create a learning environment supportive for all students including those constituting a minority.

A fourth limitation involves methodological issues. It should be noted that the present research is restricted to self-reports and thus reflects students' perceptions of their degree of social inclusion and academic competencies, rather than information from external sources, such as students' actual test grades as a measure of academic performance in college and the reports of other peers, teachers, or even observational data regarding their social inclusion. Another methodological limitation concerns our measures. Given that we conducted research in a real-life educational setting with considerable time constraints, some of the scales applied only consisted of a limited number of items. However, it would be desirable to use multi-item measures in future research in order to provide stronger validity to the present results and to develop new measures of students' sense of belonging that take into account both the social and the ability-related component of the construct. Further, to tap into the social component of the ability measure, we used a one-item measure applied by Walton and Cohen (2007), asking students to evaluate their individual performance potential in comparison to that of fellow students on a percentile scale. Given the minority situation of female students, it would have been interesting to add a measure tapping into their perceived performance compared to other female or male students, respectively, and to assess which student group the relevant group of reference is. Another limitation possibly related to the application of few-item scales in our study concerns the weak to moderate correlations between some of our independent variables, especially with regard to the correlation between our measures of academic self-efficacy, and the perceived individual performance potential compared to that of fellow students. Although both concepts are theoretically clearly distinct from each other, they overlap in that both describe perceptions of the self in academic contexts. Here, again, it would be desirable to apply multi-item scales to enhance the predictive validity of the measures and further, to conduct confirmatory factor analyses to quantify the extent of conceptual overlap in future studies. Lastly, disparities in gender representation also became apparent in our study: of our total sample, only 23% were female computer science students, resulting in a relative small share of our focal group of participants in the sample. Although we have a representative sample in terms of gender in our study (Federal Statistical Office of Germany, 2018), the naturally larger standard error in the smaller subgroup makes our testing more conservative, with the result that we might even underestimate gender mean differences (see Gelman and Hill, 2007). Moreover, in moderated multiple regression analyses with unequal subgroup sample sizes, the statistical power of the inferential test cannot exceed the power of a test involving two subgroups, each of the size of $2(n_1)$ with n_1 being the size of the smaller subgroup, regardless of the size of the second subgroup (Arguinis, 2004). The large – and typical – dropout numbers in computer science further contribute to this problem. Future research replicating our findings with larger sample sizes would therefore be desirable and important.

CONCLUSION

As female students face negative stereotypes about their ability in quantitative fields and continue to remain underrepresented in computer science, understanding the factors that explain these phenomena is key for creating stimulating and encouraging educational environments for females studying a male-dominated subject such as computer science. When considering the causes and cures of this existing gender gap, students' uncertainty about belonging is a promising variable to study.

The present study identified male and female sources of belonging uncertainty in the computer sciences and thereby extends our understanding of this theoretical concept. Our results suggest that belonging uncertainty is comprised of both students' concerns about their social connectedness in an academic domain and concerns about their academic abilities. Therefore, conceptualizing belonging uncertainty as regarding only concerns about the quality of one's social relationship in an academic domain leads to an incomplete picture of this phenomenon.

By identifying the sources of the uncertainty about belonging in computer science, our results may serve to inform the institutional organization of minority students' integration into their studies as well as interventions aimed at fostering students' sense of belonging and increasing the share of female students in computer science and other STEM domains.

ETHICS STATEMENT

According to our university guidelines and to national regulations in Germany, no ethics review was required, because the current research can be classified as research using anonymous or no-risk tests, surveys, interviews, or observations. Written informed consent was obtained from all participants prior to every data assessment.

AUTHOR CONTRIBUTIONS

LZ provided the initial idea, conceived, and designed the study. EH conducted the data collection, organized the data base, carried out the statistical analyses, and wrote the first draft of the manuscript. Both authors contributed to the manuscript revision, read, and approved the submitted version of the manuscript.

FUNDING

The publication of this article was funded by the Open Access Fund of the Leibniz University Hannover.

ACKNOWLEDGMENTS

We thank Prof. Dr. Margarita Esponda-Argüero and Prof. Dr. Bettina Hannover for their valuable support and cooperation. We also thank Julia Böttger, Michael Friedrich, Lisa Grützmacher, John Gubernath, Jannika Haase, Martina Oldani, Madeleine Richter, Tobias Schmidt, Dario Sonnenburg, Deborah Staiger, and Jasmin Quambusch for their help with the data collection and entry.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01740/full#supplementary-material

REFERENCES

Appel, M., Weber, S., and Kronberger, N. (2015). The influence of stereotype threat on immigrants: review and meta-analysis. *Front. Psychol.* 6:900. doi: 10.3389/fpsyg.2015.00900

Arguinis, H. (2004). *Regression Analysis for Categorical Moderators*. New York, NY: Guilford Press.

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: W. H. Freeman.

Baumeister, R. F., and Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117, 497–529. doi: 10.1037/0033-2909.117.3.497

Baumeister, R. F., Twenge, J. M., and Nuss, C. K. (2002). Effects of social exclusion on cognitive processes: anticipated aloneness reduces intelligent thought. *J. Pers. Soc. Psychol.* 83, 817–827. doi: 10.1037/0022-3514.83.4.817

Berkman, L. F., and Syme, L. (1979). Social networks, host resistance, and mortality: a nine-year follow-up study of alameda county residents. *Am. J. Epidemiol.* 109, 186–204. doi: 10.1017/CBO9780511759048

Blanton, H., Buunk, B. P., Gibbons, F. X., and Kuyper, H. (1999). When better-than-others compare upward: choice of comparison and comparative evaluation as independent predictors of academic performance. *J. Pers. Soc. Psychol.* 76, 420–430. doi: 10.1037/0022-3514.76.3.420

Blanton, H., Crocker, J., and Miller, D. T. (2000). The effects of in-group versus out-group social comparison on self-esteem in the context of a negative stereotype. *J. Exp. Soc. Psychol.* 36, 519–530. doi: 10.1006/jesp.2000.1425

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York, NY: Wiley.

Bong, M., and Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: how different are they really? *Educ. Psychol. Rev.* 15, 1–40. doi: 10.1023/A:1021302408382

Bühner, M., and Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. Munich: Pearson Education.

Cohen, J. E. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associated, Inc.

Cohen, S., and Janicki-Deverts, D. (2009). Can we improve our physical health by altering our social networks? *Perspect. Psychol. Sci.* 4, 375–378. doi: 10.1111/j.1745-6924.2009.01141.x

Cundiff, J. L., Vescio, T. K., Loken, E., and Lo, L. (2013). Do gender-science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Soc. Psychol. Educ.* 16, 541–554. doi: 10.1007/s11218-013-9232-8

Davies, P. G., Spencer, S. J., Quinn, D. M., and Gerhardstein, R. (2002). Consuming images: how television commercials that elicit stereotype threat can restrain women academically and professionally. *Pers. Soc. Psychol. Bull.* 28, 1615–1628. doi: 10.1177/014616702237644

Deemer, E. D., Lin, C., and Soto, C. (2016). Stereotype threat and women's science motivation: examining the disidentification effect. *J. Career Assess.* 24, 637–650. doi: 10.1177/1069072715616064

Eisenberger, N. I., Lieberman, M. D., and Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science* 302, 290–292. doi: 10.1126/science.1089134

Eisinga, R., Grotenhuis, Mt, and Pelzer, B. (2013). The reliability of a two-item scale: pearson, cronbach, or spearman-brown? *Int. J. Public Health* 58, 637–643. doi: 10.1007/s00038-012-0416-3

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in stem subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

European Institute for Gender Equality (2017). *Gender Segregation in Education, Training and the Labour Market*. Brussels: Council of the European Union.

Federal Statistical Office of Germany (2017). *Bildung und Kultur. Schnellmeldungsergebnisse der Hochschulstatistik zu Studierenden und Studienanfänger/-innen – Vorläufige Ergebnisse – Wintersemester 2017/2018*. Wiesbaden: Statistisches Bundesamt.

Federal Statistical Office of Germany (2018). *Bildung und Kultur. Studierende an Hochschulen –Wintersemester 2017/2018*. Wiesbaden: Statistisches Bundesamt.

Gelman, A., and Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

Good, C., Rattan, A., and Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *J. Pers. Soc. Psychol.* 102, 700–717. doi: 10.1037/a0026659

Hermann, A. D., Leonardelli, G. J., and Arkin, R. M. (2002). Self-doubt and self-esteem: a threat from within. *Pers. Soc. Psychol. Bull.* 28, 395–408. doi: 10.1177/0146167202286010

Heublein, U., and Schmelzer, R. (2018). *Die Entwicklung der Studienabbruchquoten an den Deutschen Hochschulen*. Hannover: Deutsches Zentrum für Hochschul- und Wissenschaftsforschung.

Höhne, E., and Zander, L. (in press). Belonging uncertainty as predictor of dropout intentions among first-year students of the computer sciences. *Z. Erziehwiss.*

Hox, J., van Buuren, S., and Jolani, S. (2015). "Incomplete multilevel data: problems and solutions," in *Advances in Multilevel Modeling for Educational Research: Addressing Practical Issues Found in Real-World Applications*, eds J. R. Harring, L. M. Stapleton, and S. N. Beretvas (Charlotte, NC: Information Age Publishing Inc), 39–62.

Hoyt, C. L., and Blascovich, J. (2007). Leadership efficacy and women leaders' responses to stereotype activation. *Group Process. Intergroup Relat.* 10, 595–616. doi: 10.1177/1368430207084718

Huang, C. (2013). Gender differences in academic self-efficacy: a meta-analysis. *Eur. J. Psychol. Educ.* 28, 1–35. doi: 10.1007/s10212-011-0097-y

IBM Corp (2017). *IBM SPSS Statistics for Windows, Version 25.0*. Armonk, NY: IBM Corp.

Jerusalem, M., and Schwarzer, R. (1986). "Selbstwirksamkeit," in *Skalen zur Befindlichkeit und zur Persönlichkeit*, ed. R. Schwarzer (Berlin: Zentrale Universitäts-Druckerei), 15–28.

Lent, R. W., Brown, S. D., and Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* 45, 79–122. doi: 10.1006/jvbe.1994.1027

Lerdpornkulrat, T., Koul, R., and Sujivorakul, C. (2012). The influence of ability beliefs and motivational orientation on the self-efficacy of high school science students in Thailand. *Aust. J. Educ.* 56, 163–181. doi: 10.1177/000494411205600205

Lewis, K. L., and Hodges, S. D. (2015). Expanding the concept of belonging in academic domains: development and validation of the ability uncertainty scale. *Learn. Individ. Differ.* 37, 197–202. doi: 10.1016/j.lindif.2014.12.002

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *J. Am. Stat. Assoc.* 83, 1198–1202. doi: 10.1080/01621459.1988.10478722

Little, R. J. A., and Rubin, D. B. (1989). The analysis of social science data with missing values. *Sociol. Methods Res.* 18, 292–326. doi: 10.1177/0049124189018002004

Lockwood, P., and Kunda, Z. (1997). Superstars and me: predicting the impact of role models on the self. *J. Pers. Soc. Psychol.* 73, 91–103. doi: 10.1037/0022-3514.73.1.91

London, B., Downey, G., Romero-Canyas, R., Rattan, A., and Tyson, D. (2012). Gender based rejection sensitivity and academic self-silencing in women. *J. Pers. Soc. Psychol.* 102, 961–979. doi: 10.1037/a0026615

Ludwig, P. H. (2010). "Schulische erfolgserwartungen und begabungsselbstbilder bei mädchen – strategien ihrer veränderung," in *Handbuch Mädchenpädagogik*, eds M. Matzner and I. Wyrobnik (Weinheim: Beltz), 145–158.

MacDonald, G., and Leary, M. R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychol. Bull.* 131, 202–223. doi: 10.1037/0033-2909.131.2.202

Mallett, R. K., Mello, Z. R., Wagner, D. E., Worrell, F., Burrow, R. N., and Andretta, J. R. (2011). Do i belong? It depends on when you ask. *Cultur. Divers. Ethnic Minor. Psychol.* 17, 432–436. doi: 10.1037/a0025455

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *J. Educ. Psychol.* 79, 280–295. doi: 10.1037/0022-0663.79.3.280

Marsh, H. W., and O'Mara, A. (2006). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: unidimensional and multidimensional perceptions of self-concept. *Pers. Soc. Psychol. Bull.* 34, 542–552. doi: 10.1177/0146167207312313

Marsh, H. W., and Parker, J. W. (1984). Determinants of student self-concept: is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *J. Pers. Soc. Psychol.* 47, 213–231. doi: 10.1037/0022-3514.47.1.213

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., et al. (2018). The murky distinction between self-concept and self-efficacy: beware of lurking jingle-jangle fallacies. *J. Educ. Psychol.* 111, 331–353. doi: 10.1037/edu0000281

Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., and Köller, O. (2007). The big-fish-little-pond effect: persistent negative effects of selective high schools on self-concept after graduation. *Am. Educ. Res. J.* 44, 631–669. doi: 10.3102/0002831207306728

Master, A., Cheryan, A., and Meltzoff, A. N. (2016). Computing whether she belongs: stereotypes undermine girls' interest and sense of belonging in computer science. *J. Educ. Psychol.* 108, 424–437. doi: 10.1037/edu0000061

McCabe, C. J., Kim, D. S., and King, K. M. (2018). Improving present practices in the visual display of interactions. *Adv. Methods Pract. Psychol. Sci.* 1, 147–165. doi: 10.1177/2515245917746792

Mendoza-Denton, R., Downey, G., Purdie, V. J., Davis, A., and Pietrzak, J. (2002). Sensitivity to status-based rejection: implications for african american students' college experience. *J. Pers. Soc. Psychol.* 83, 896–918. doi: 10.1037/0022-3514.83.4.896

Murphy, M. C., Steele, C. M., and Gross, J. J. (2007). Signaling threat: how situational cues affect women in math, science, and engineering settings. *Psychol. Sci.* 18, 879–885. doi: 10.1111/j.1467-9280.2007.01995.x

Murphy, M. C., and Taylor, V. J. (2012). "The role of situational cues in signaling and maintaining stereotype threat," in *Stereotype Threat: Theory, Process, Application*, eds M. Inzlicht and T. Schmader (New York, NY: Oxford University Press), 17–33.

Muthén, L. K., and Muthén, B. O. (1998–2017). *Mplus User's Guide. Eighth Edition*. Los Angeles, CA: Muthén & Muthén.

Nguyen, H.-H. D., and Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *J Appl Psychol.* 93, 1314–1334. doi: 10.1037/a0012702

Nimon, K. F. (2012). Statistical assumptions of substantive analyses across the general linear model: a mini-review. *Front. Psychol.* 3:322. doi: 10.3389/fpsyg.2012.00322

OECD (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. Paris: OECD Publishing.

OECD (2017). *Bildung auf einen Blick 2017: OECD-Indikatoren*. Bielefeld: W. Bertelsmann Verlag.

Park, H. S. (2008). Centering in hierarchical linear modeling. *Commun. Methods Meas.* 2, 227–259. doi: 10.1080/19312450802310466

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., and Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educ. Psychol.* 34, 29–48. doi: 10.1080/01443410.2013.797339

Peugh, J. L., and Enders, C. K. (2004). Missing data in educational research: a review of reporting practices and suggestions for improvement. *Rev. Educ. Res.* 74, 525–556. doi: 10.3102/00346543074004525

Robbins, S. B., Lauver, K., Le, H., Davis, D., and Langley, R. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychol. Bull.* 130, 261–288. doi: 10.1037/0033-2909.130.2.261

Schafer, J. L., and Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychol. Methods* 7, 147–177. doi: 10.1037/1082-989X.7.2.147

Schuster, C., and Martiny, S. E. (2017). Not feeling good in stem: effects of stereotype activation and anticipated affect on women's career aspirations. *Sex Roles* 76, 40–55. doi: 10.1007/s11199-016-0665-3

Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., and Simons, A. (1991). Ease of retrieval as information: another look at the availability heuristic. *J. Pers. Soc. Psychol.* 61, 195–202. doi: 10.1037/0022-3514.61.2.195

Seaton, M., Marsh, H. W., and Craven, R. G. (2009). Earning its place as a pan-human theory: universality of the big-fish-little-pond effect across 41 culturally and economically diverse countries. *J. Educ. Psychol.* 101, 403–419. doi: 10.1037/a0013838

Shavelson, R. J., Hubner, J. J., and Stanton, G. C. (1976). Self-concept: validation of construct interpretations. *Rev. Educ. Res.* 46, 407–441. doi: 10.3102/00346543046003407

Sijtsma, K. (2009). On the use, misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika* 74, 107–120. doi: 10.1007/s11336-008-9101-0

Simms, L. J., and Watson, D. (2007). "The construct validation approach to personality scale construction," in *Handbook of research methods in personality psychology*, eds R. W. Robins, R. C. Fraley, and R. F. Kruger (New York, NY: Guilford Press), 240–258.

Skourletos, J. C., Murphy, M. C., Emerson, K. T. U., and Carter, E. A. (2013). Social identity and academic belonging: creating environments to minimize the achievement gap among African American and Latino students. *Mod. Psychol. Stud.* 18, 23–29.

Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): an investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles* 67, 617–629. doi: 10.1007/s11199-012-0209-4

Smith, J. L., Lewis, K. L., Hawthorne, L., and Hodges, S. D. (2013). When trying hard isn't natural: women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns. *Pers. Soc. Psychol. Bull.* 39, 131–143. doi: 10.1177/0146167212468332

Snijders, T. A. B., and Bosker, R. J. (2012). *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*, 2nd Edn. Los Angeles, CA: SAGE Publications.

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35, 4–28. doi: 10.1006/jesp.1998.1373

Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *Am. Psychol.* 52, 613–629. doi: 10.1037/0003-066X.52.6.613

Tajfel, H., and Turner, J. C. (1986). "The Social Identity Theory of Intergroup Behaviour," in *Psychology of Intergroup Relations*, eds S. Worchel and L. W. Austin (Chicago: Nelson Hall), 7–24.

Trautwein, U., Lüdtke, O., Marsh, H. W., and Nagy, G. (2009). Within school social comparison: how students perceive the standing of their class predicts academic self-concept. *J. Educ. Psychol.* 101, 853–866. doi: 10.1037/a0016306

Tymms, P., Merrell, C., and Henderson, B. (1997). The first at school: a quantitative investigation of the attainment and progress of pupils. *Educ. Res. Eval.* 3, 101–118. doi: 10.1080/1380361970030201

Uchino, B. N. (2006). Social support and health: a review of physiological processes potentially underlying links to disease outcomes. *J. Behav. Med.* 29, 377–387. doi: 10.1007/s10865-006-9056-5

Valentine, J. C., DuBois, D. L., and Cooper, H. (2004). The relation between self-beliefs and academic achievement: a meta-analytic review. *Educ. Psychol.* 39, 111–133. doi: 10.1207/s15326985ep3902_3

Walton, G. M., and Carr, P. B. (2012). "Social belonging and the motivation and intellectual achievement of negatively stereotyped students," in *Stereotype Threat: Theory, Process, Application*, eds M. Inzlicht and T. Schmader (New York, NY: Oxford University Press), 89–106.

Walton, G. M., and Cohen, G. L. (2003). Stereotype lift. *J. Exp. Soc. Psychol.* 39, 456–467. doi: 10.1016/S0022-1031(03)00019-2

Walton, G. M., and Cohen, G. L. (2007). A question of belonging: race, social fit, and achievement. *J. Pers. Soc. Psychol.* 92, 82–96. doi: 10.1037/0022-3514.92.1.82

Walton, G. M., and Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* 331, 1447–1451. doi: 10.1126/science.1198364

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., and Zanna, M. P. (2015). Two brief interventions to mitigate a "Chilly Climate" transform women's experience, relationships, and achievement in engineering. *J. Educ. Psychol.* 107, 468–485. doi: 10.1037/a0037461

Wang, P., Zhou, P., Tan, C.-H., and Zhang, P.-C. (2017). Effect of self-efficacy in stereotype activation. *Soc. Behav. Pers.* 45, 469–476. doi: 10.2224/sbp.5201

Woodcock, A., Hernandez, P. R., Estrada, M., and Schultz, P. W. (2013). The consequences of chronic stereotype threat: domain disidentification and abandonment. *J. Pers. Soc. Psychol.* 103, 635–646. doi: 10.1037/a0029120

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 September 2019
doi: 10.3389/fpsyg.2019.01941

Differential Contributions of Empathy to Math Achievement in Women and Men

Nermine Ghazy, Eleanor Ratner and Miriam Rosenberg-Lee†*

Department of Psychology, Rutgers University, Newark, NJ, United States

Mathematics forms a foundation for the science, technology, engineering and mathematics (STEM) fields. While considerable work has identified the individual cognitive and external systemic factors that influence math achievement, less is known about personality-like traits that might contribute to success in mathematics, especially among women. This study examines two such traits: systemizing – the tendency to analyze systems and extract underlying rules that govern their behavior – and empathizing – the ability to identify with another's emotions and respond appropriately. Recently Escovar et al. (2016) found that empathizing was a negative predictor of math skills in children, especially among girls, suggesting that women with higher empathy might be particularly disposed to lower math performance. In the first study, 142 participants (71 female) completed two standardized measures of math achievement and questionnaires to gauge the tendency to empathize and systemize. Surprisingly, higher empathy was associated with better math performance in women, while men displayed the expected pattern of lower empathy being related to higher math scores. In a second study, we extend this finding in women (n = 121) to show that individuals who report higher mathematics achievement in university level course work also have higher empathy scores. Further, while positive attitudes toward mathematics tended to decline from elementary school to college, women whose attitudes increased had higher empathy scores than those who declined. Together, these results suggest that while the tendency to empathize is associated with worse math performance in childhood, it may become a protective factor as women progress through their mathematics education.

Keywords: math achievement, STEM – science technology engineering mathematics, gender differences, empathy quotient, systemizing quotient

Edited by:
Silke Luttenberger,
University College of Teacher Education Styria, Austria

Reviewed by:
Annika M. Svedholm-Häkkinen,
University of Helsinki, Finland
Matthew K. Belmonte,
The Com DEALL Trust, India

Correspondence:
Miriam Rosenberg-Lee
miriam.rosenberglee@rutgers.edu

†ORCID:
Miriam Rosenberg-Lee
orcid.org/0000-0003-0768-0347

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *15 January 2019*
Accepted: *07 August 2019*
Published: *06 September 2019*

Citation:
Ghazy N, Ratner E and Rosenberg-Lee M (2019) Differential Contributions of Empathy to Math Achievement in Women and Men. Front. Psychol. 10:1941. doi: 10.3389/fpsyg.2019.01941

INTRODUCTION

At the undergraduate level, fewer women choose science, technology, engineering, and mathematics (STEM) majors and among those who do pursue these fields, many are diverted toward teaching over higher status STEM careers, such as computer science and engineering (Beede et al., 2011). Moreover, women who do work in STEM are more likely to have lower wages than their male counterparts (Beede et al., 2011). At the elementary school level, boys and girls perform equivalently, but the genders diverge over the course of schooling (Hyde et al., 1990), and ultimately, mathematics has among the lowest female participation rate at the graduate level (Leslie et al., 2015).

Considerable research to date has focused on factors that contribute to low STEM participation rates in the United States, such as a lack of female role models (Cheryan et al., 2011) and how gender stereotypes may subconsciously discourage women from pursuing STEM fields (Shapiro and Williams, 2012). These important external, systemic factors interact with internal, individual factors to impact academic achievement and eventually predispose some individuals to avoid STEM careers. Mathematics achievement is a strong predictor of STEM career attainment (Crisp et al., 2009), making it vital to understand how individual differences in personality-like traits influence math achievement and hence pathways to STEM careers.

This study examines two such traits: *systemizing* – the tendency to analyze systems and extract underlying rules that govern their behavior – and *empathizing* – the ability to identify with another's emotions and respond appropriately. While systemizing is hypothesized to relate to math skills (Baron-Cohen, 2009), Escovar et al. (2016) recently found no evidence for a direct relationship in a large sample of children. Instead, they reported that higher levels of empathizing were related to lower math skills, especially in girls. The authors hypothesized that girls who were more attuned to the social environment may be particularly susceptible to stereotyped messages about gender and mathematics, potentially impacting performance over the course of schooling and ultimately lowering participation in STEM fields. Thus, the primary goal of the current study was to examine whether a negative relationship between empathizing and math achievement held in female college students.

The Role of Empathizing and Systemizing in Math Cognition

Theorizing on the role of systemizing and empathizing in math cognition grows out of research on autism spectrum disorder (ASD), a neurodevelopmental disorder marked by social communication deficits and repetitive behaviors or unusual interests (American Psychiatric Association, 2013). Notably, individuals with autism often have remarkable cognitive strengths in domains such as memory, spatial reasoning and mathematics (Treffert, 2009; Iuculano et al., 2014). Baron-Cohen (2009) has suggested that this constellation of behaviors can be related to difficulties empathizing with others, but a relative strength in systematic thinking. Mathematics, with its emphasis on rules and structures is a paradigmatic example of a system, suggesting that individual differences in systemizing should be positively related to variation in mathematics skills. While several studies find increases in systemizing among individuals with autism (Wakabayashi et al., 2007; Auyeung et al., 2009; Greenberg et al., 2018) and higher rates of autistic-like behaviors in mathematicians (Baron-Cohen et al., 2001), there is surprisingly little empirical support relating systemizing directly to math skills among either typically developing individuals or those on the autism spectrum. In fact, the one study to examine this question directly in adult females found no correlation between the tendency to systemize and math achievement (Morsanyi et al., 2012).

In children, Escovar et al. (2016) did find a marginally significant positive relationship between parents' reports on the Systemizing Quotient (SQ) and a measure of mathematical reasoning. However, this weak relation was entirely accounted for by shared variance with measures of intelligence and reading ability. The child version of the SQ is combined with the Empathy Quotient (EQ) into a single instrument, so Escovar et al. (2016) also assessed the relationship between participants' tendencies to empathize and math achievement. Contrary to the prediction of no relationship, they reported that greater empathizing was related to lower performance on a composite of math calculation skills (comprising both timed and untimed measures), and this relation remained even after accounting for intelligence and reading ability (Escovar et al., 2016). Consistent with prior research (Lawrence et al., 2004; Auyeung et al., 2009), girls had significantly higher scores on the EQ, moreover, the relationship between EQ and math scores was stronger in girls than boys, although this interaction effect did not reach significance.

Explaining the Influence of Empathizing on Math Achievement

To further explore the relationships between gender, EQ, and math scores, Escovar et al. (2016) considered the role of math anxiety, a robust predictor of math achievement (Ashcraft, 2002; Wu et al., 2012). Potentially, individuals who are more empathetically tuned may be more prone to experience math anxiety, thus explaining the relationship between EQ and math achievement. Consistent with prior research (Ma, 1999; Wu et al., 2012), math anxiety predicted math performance, but this relationship was independent of the influence of empathizing (Escovar et al., 2016). Instead, they found that parents' report of children's social skills were strongly related to empathizing and mediated the relationship between empathizing and math skills, an effect that was significantly stronger in girls than boys.

Empathy is not a unitary concept and modern approaches define both cognitive empathy, the ability adopt another's perspective and recognize and label their mental state, and emotional empathy (also termed affective empathy or emotional reactivity), the tendency produce the appropriate emotional response to another's emotional state (Rueda et al., 2015). A third category, social skills, describes using the appropriate behaviors in response to the emotions of others (Lawrence et al., 2004). Using factor analysis, Lawrence et al. (2004) established that these three components can be assessed from subsets of the EQ items. In the current study, we sought to further explore the relationship between empathy and math achievement by considering its relationship to these three components.[1] More broadly, a second goal of this study was to understand the factors contributing to the relationship between math achievement and empathy, with respect to math anxiety, social skills and distinct components of empathizing.

Gender Gap in Math Achievement

The gender gap in mathematics achievement is not fixed over the course of schooling. In elementary school there are no consistent

[1] We thank one of the reviewers for suggesting this empathy component analysis.

differences between the genders (Bakker et al., 2019), but the gap grows in middle school, high school and university (Hyde et al., 1990), with men outperforming women. One potential explanation for this pattern is that social, rather than cognitive factors, grow in their influence on women's math achievement.

The phenomenon of stereotype threat – in which awareness of gender stereotypes can reduce performance in stereotyped groups – suggests that individuals with stronger tendencies to be aware of the feelings of others (as in the EQ) and cognizant of social expectations (a component of social skills) may be particularly prone to internalizing these stereotypes by the time they reach adulthood. For example, Beilock et al. (2010) found that students of teachers with greater math anxiety had smaller gains in math achievement over the school year, but only among students who endorsed gender stereotypes. Attitudes toward mathematics are an individual difference trait that also correlates with current math achievement and future math performance (Anttonen, 1969; Ma and Kishor, 1997). We reasoned that among women, changing attitudes toward mathematics may be a rough index of internalization of societal messages about gender and mathematics. Thus, a third goal of this study was to assess whether changes in women's attitudes toward mathematics over the course of schooling impact the relationship between empathizing and math achievement in adulthood.

The Current Studies

In this paper, we sought to investigate the relationship between math achievement and empathizing, systemizing and social skills in adults. In Study 1, we tested whether the association of higher empathy with lower math skills found in children also holds in young adults, especially women. In Study 2, we sought to replicate the results of Study 1 and examine the role of changing attitudes toward mathematics on the relationship between empathizing and math achievement in women.

STUDY 1

Methods and Procedures

Participants

One Hundred and forty two college-aged students (71 women) attending Rutgers University in Newark, NJ, United States, ranging in age from 18 to 25, participated in this study. Participants received credit toward their psychology coursework.

Standardized Mathematics Assessments

Math skills were measured using two subtests of the Woodcock Johnson – 3rd Edition (WJ-III): Calculation and Math Fluency. While both measures were collected by Escovar et al. (2016), that study combined them into a single calculation composite score. Here, we analyzed the measures separately, in order to assess differential effects on timed and untimed math skills. Calculation, a non-timed test, measures a person's ability to complete mathematical problems that increase in difficulty starting from basic arithmetic up to calculus. Math Fluency, a 3-min timed test of single-digit arithmetic problems, requires answering intermixed addition, subtraction, and multiplication questions. Raw scores were converted to age based standardized scores using Woodcock Johnson Compuscore software.

Empathizing Measure

The EQ (Baron-Cohen and Wheelwright, 2004) is a 40-item questionnaire for adults designed to measure a person's empathy. Empathy is defined as being able to identify with another person's emotions and perspective. Responses are given using a 4-point Likert scale (*definitely agree, slightly agree, slightly disagree,* and *definitely disagree*), and a subset of items are reverse coded to reduce acquiescence bias. Participants receive a "0" for a response that does not endorse the trait, and a "1" or "2" for a response that endorses the trait, depending on the strength of the reply. Missing items were replaced by the participants' average score on all other items in that measure. For the EQ, there were ten instances of items missing (0.18% of items) and no participants missed more than one item. Internal reliability was good (Cronbach's $\alpha = 0.831$).

To investigate distinct components of empathy, namely cognitive empathy (EQ-CE), emotional empathy (EQ-EE) and social skills (EQ-SS), we summed scores on the items identified by the maximum factor loadings in Lawrence et al. (2004). This division of items corresponds to the approach taken in a large study of the psychometric properties of the EQ in a Dutch sample (Groen et al., 2015). Further, assigning items uniquely to components reduces the correlation between the components, allowing us to better assess their independent contributions to math skills. Items in the EQ-CE component probe understanding of the mental states of others (*I can pick up quickly if someone says one thing but means another*), while EQ-EE items assess appropriate emotional reactions (*I get upset if I see people suffering on news programs*). Finally, items on the EQ-SS index ease in social situations (*I don't tend to find social situations confusing*).

Systemizing Measure

The SQ (Baron-Cohen et al., 2003) is a 75-item adult self-report questionnaire designed to measure a person's ability to systemize. Systemizing is defined as the tendency to analyze or construct organized schemes. Responses are given using the same 4-point Likert scale as the EQ and a subset of items are reverse coded to reduce acquiescence bias. For the SQ, there were 26 instances of items missing (0.24%) and one participant was missing four items, six were missing two items and the remaining participants missed one item. Scoring follows the same procedures as for the EQ and missing items were replaced by the participants' average score on all other items in the measure. Internal reliability for the SQ was good (Cronbach's $\alpha = 0.892$).

Math Anxiety Measure

The Abbreviated Math Anxiety Research Scale (A-MARS, Alexander and Martray, 1989) is a 25-item questionnaire for adults measuring math anxiety. Participants are asked to rate their level of anxiety in a variety of situations, such as studying for and taking math tests, using a 5-point Likert scale (*not at all, a little, a fair amount, much,* and *very much*).

Total score is the sum of responses and missing values were replaced by that participant's average on other items. There were six instances of items missing (0.17%) and no participants were missing more than one item. Cronbach's alpha for the A-MARS was 0.957.

Social Abilities Measure

The adult self-report version of the social responsiveness scale (SRS, Constantino and Gruber, 2012) was used to broadly assess social skills. Each question on this 65-item measure describes a social behavior or characteristic and uses a 0- to 3-point Likert scale (*not true, sometimes true, often true,* and *almost always true*). Higher scores on the SRS correspond to worse social skills, thus items describing prosocial behaviors were reverse scored. We employed the total raw scores for subsequent analyses, and missing values were replaced by that participant's average on all other items. There were 10 instances of items missing (0.22%), one participant was missing three items, two were missing two items, and the remaining participants missed one item. Cronbach's alpha for the SRS was 0.908.

Demographics Questionnaire

Participants were asked to report their assigned sex status at birth (Male, Female, Intersex) and their gender identification (Male, Female, Gender Variant). No participants reported being assigned intersex at birth, nor identified as gender variant, and all participants reported assigned sex and gender identification matched and were used in subsequent analyses. Participants were also asked about race and social-economic status.

Regression Analyses

Multiple regression analyses were performed to assess the interaction of gender with EQ and SQ in predicting math skills. We first modeled the main effects of gender, EQ and SQ on the Calculation and Math Fluency math measures. Next, we added terms for the interaction of Gender with EQ and Gender with SQ to the model, using dummy coding to indicate female participants. To explore the effects of math anxiety and SRS on Calculation, we also computed main effect models including those terms. And then asked whether adding the interaction terms explained additional variance. To examine the independent contributions of the EQ components on Calculation and Math Fluency, we first modeled the main effects of all the components, along with gender and then added all the interaction terms to the model. Covariates were mean centered before being entered into the model, reducing multicollinearity between predictors, as indicated by variance inflation factors less than 4 in all models (Wood, 1984).

Procedure

Participants completed the Math Fluency and Calculation subtests of the WJ-III, using pencil and paper. Participants then used a laptop computer to complete a series of questionnaires implemented in Qualtrics (Provo, UT, United States). Participants completed the EQ, SQ, A-MARS, SRS, and the basic demographics questionnaire. The full session took 50–60 min.

Results

Gender Differences in Math Achievement, Systemizing and Empathizing

Participant characteristics are listed **Table 1**. Women scored significantly lower than men in the untimed Calculation and in timed Math Fluency subtests of the WJ-III). Women scored significantly lower on the SQ than men. While women had numerically higher scores on the EQ, this difference did not reach significance ($p = 0.108$).

Relations Between Math Achievement, Systemizing and Empathizing

Table 2 lists the correlations between SQ, EQ and math performances in the full sample. Scores on the EQ and SQ were positively correlated, but there was no relationship between EQ or SQ and either of the math measures (all ps > 0.13).

Gender Differences in Relations Between Empathizing, Systemizing and Math Achievement

Next, we considered the relations between EQ, SQ and math achievement, separately in women and men. Here, we found divergent results between the genders, with positive relations of EQ and SQ with math achievement in women, and negative relationships in men (**Table 2** and **Figure 1**). Most notably, in women there was a strong

TABLE 1 | Participant characteristics (Study 1).

	Women (n = 71)		Men (n = 71)		t	p
	M	SD	M	SD		
Age[1]	20.370	1.819	20.280	1.870	−0.287	0.774
Calculation	104.324	12.334	108.464	11.846	2.040	0.043
Math Fluency	94.930	13.364	101.690	13.593	2.988	0.003
EQ	45.004	10.583	42.195	10.134	−1.616	0.108
SQ	59.932	18.904	67.443	18.968	2.363	0.019
Math Anxiety	64.952	19.580	60.672	21.911	−1.227	0.222
SRS	111.325	21.259	116.463	23.510	1.366	0.174

EQ, Empathy Quotient; SQ, Systemizing Quotient; SRS, Social Responsiveness Scale; [1]Two female participants did not report their date of birth; therefore, the sample size for women is n = 69.

TABLE 2 | Pearson correlations between math achievement, EQ, SQ, Math Anxiety, and SRS (Study 1).

	Math Fluency	EQ	SQ	Math Anxiety	SRS
Full Sample					
Calculation	0.464**	−0.027	−0.028	−0.214*	0.078
Math Fluency		0.118	0.126	−0.284**	−0.111
EQ			0.332**	−0.193*	−0.580***
SQ				−0.203*	−0.147
Math Anxiety					0.366**
Women					
Calculation	0.534**	0.233	0.058	−0.282*	−0.189
Math Fluency		0.313**	0.273*	−0.226	−0.336**
EQ			0.300*	−0.119	−0.600***
SQ				−0.362**	−0.264*
Math Anxiety					0.312**
Men					
Calculation	0.349**	−0.262*	−0.190	−0.126	0.294*
Math Fluency		−0.002	−0.104	−0.306**	0.026
EQ			0.442***	−0.297*	−0.551***
SQ				−0.033	−0.093
Math Anxiety					0.439***

*Top panel: Full sample. Middle panel: Women. Bottom panel: Men. Abbreviations as in **Table 1**. *Correlation is significant at the 0.05 level (two-tailed). **Correlation is significant at the 0.01 level (two-tailed). ***Correlation is significant at the 0.001 level (two-tailed).*

positive relationship between EQ and Math Fluency and a marginal relationship with Calculation. While in men, the opposite pattern held, namely lower empathizing was related to higher math achievement for Calculation, although there was no relationship for Math Fluency. For systemizing, we found the same pattern of positive correlations in women and negative correlations in men, but only the relationship between SQ and Math Fluency in women was statistically significant (**Table 2**).

To confirm this pattern of results, we employed multiple regression analyses to determine whether EQ, SQ, or both, interacted with gender in predicting participants' math scores. For Calculation, in Model 1, there was a main effect of gender and as expected no significant effects for EQ or SQ. Adding the interaction terms produced a significant F-change [$F(2,136) = 4.36$, $p = 0.015$] and there was a significant interaction between gender and EQ, but no interaction between gender and SQ (**Table 3**, Model 2). A follow-up analysis, using only EQ, gender and their interaction in predicting Calculation, revealed a main significant negative effect of EQ in men ($t = -2.20$, $p = 0.030$) and a significant positive effect in women ($t = 3.00$, $p = 0.003$), as well as a significant

TABLE 3 | Hierarchical regression analysis of Calculation.

	B	*SE*	Beta	*t*	*p*
Model 1 R^2 = 0.033					
Intercept	108.68	1.48		73.69	<0.001
Gender	−4.56	2.13	−0.18	−2.14	0.034
EQ	0.03	0.107	0.02	0.49	0.805
SQ	−0.05	0.06	−0.07	−0.79	0.432
Model 2 R^2 = 0.091, Sig (F-change$_{model1}$) = 0.015					
Intercept	108.32	1.48		73.19	<0.001
Gender	−4.42	2.08	−0.18	−2.12	0.036
EQ	−0.26	0.16	−0.22	−1.65	0.101
SQ	−0.06	−0.08	−0.09	−0.69	0.489
Gender * EQ	0.53	0.210	0.33	2.54	0.012
Gender * SQ	0.05	0.12	0.05	0.43	0.668

Model 1: Effects of gender, empathy quotient (EQ) and systemizing quotient (SQ). Model 2: additional variance accounted for by interaction of Gender with SQ and EQ.

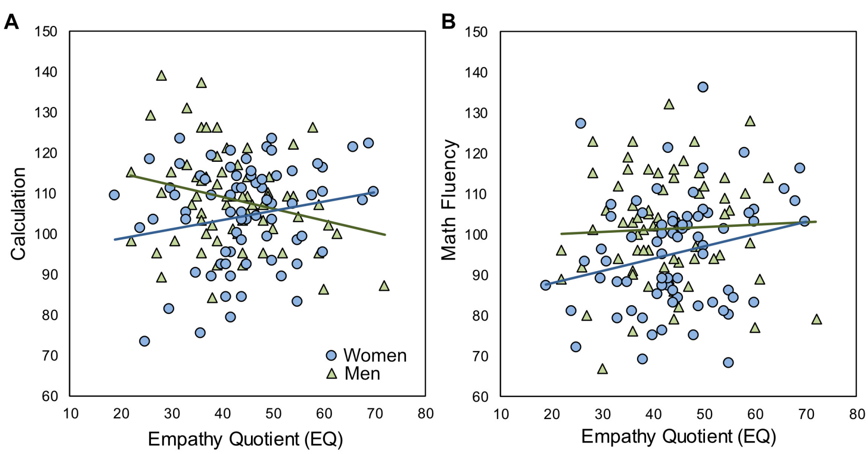


FIGURE 1 | Relations between EQ and Math Achievement (Study 1). **(A)** Women had a positive relationship between performance on the Calculation subtest of the WJ-III and empathy quotient [$r(71) = 0.233$, $p = 0.051$], while men displayed the opposite pattern [$r(71) = -0.262$, $p = 0.028$]. **(B)** Women had a positive relationship between performance on the Math Fluency subtest of the WJ-III and empathy quotient [$r(71) = 0.313$, $p = 0.008$], while there was no relationship in men [$r(71) = -0.0020$, p 0.984].

effect of gender (t = −2.05, p = 0.043). For Math Fluency, there was again a main effect of gender, but no effects of EQ or SQ (**Supplementary Table S1**). Adding the interaction terms produced a significant F-change [$F(2,136)$ = 3.43, p = 0.035], but only the interaction of gender and SQ was marginally significant.

Gender Differences in Relations Between Empathizing and Math Achievement Remain After Controlling for Math Anxiety

Next, we examined the role of math anxiety on the relationship between EQ and Calculation. Consistent with prior research, we found a negative correlation between scores on the A-MARS and performance on Calculation and Math Fluency, in the full sample (**Table 2**). We also found that math anxiety was negatively related to EQ and SQ (**Table 2**). Thus, we sought to determine if math anxiety might explain the relationship between gender and EQ in predicted Calculation scores. Adding math anxiety to the multiple regression model predicting Calculation (Model 3) revealed a main effect of math anxiety and a significant F-change relative to Model 1 [$F(1,137)$ = 6.70, p = 0.011]. However, adding the interaction terms to this model still resulted in a significant F-change [$F(2,135)$ = 5.41, p =0.005] and the interaction between EQ and gender remained significant (**Table 4**, Model 4). Notably, there was also a main effect of EQ in this model, indicating a significant negative relationship between EQ and Calculation in men, as well as the positive relationship in women denoted by the interaction term. These results suggest the role of empathizing on math achievement is not explained by math anxiety.

Gender Differences in Relations Between Empathizing and Math Achievement Are Explained by Social Skills

Next, we examined the role of social skills, as measured by the SRS, on the relationship between EQ and Calculation. Higher scores on the SRS indicate worse social skills and EQ was negatively correlated with SRS in the full group (**Table 2**) and in each gender (**Table 2**). A model (**Table 5**, Model 5) which included a main effect of SRS scores did not significantly improve the fit relative to Model 1 [$F(1,137)$ = 0.75, p =0.390], however adding the interaction of SRS and gender to that model (**Table 5**, Model 6) resulted in significant F-change [$F(1,136)$ = 8.47, p = 0.004] in the predictions of Calculation and a significant interaction between gender and SRS. A follow-up analysis, using only SRS, gender and their interaction in predicting Calculation, revealed a significant positive effect of SRS in men (t = 2.46, p = 0.015) and a significant negative effect in women (t = −2.88, p = 0.005), as well as a significant effect of gender (t = −2.03, p = 0.045). As higher SRS scores indicate worse social skills, these results show, consistent with the EQ, that worse social skills are related to better math performance in men, but the opposite pattern in women.

Finally, we added terms for the interaction of gender with EQ and SQ to Model 6, which did not produce a significant F-change [$F(2,134)$ = 1.30, p = 0.279], and none of the gender interaction terms were significant (**Table 5**, Model 7). Together, these results suggest that both SRS and EQ interact with gender in predicting math skills; however, because of the multicollinearity between the measures, both the main effects and interactions are not significant when included in the same model.

TABLE 4 | Hierarchical regression analysis of Calculation accounting for Math Anxiety.

	B	*SE*	Beta	*t*	*p*
Model 3 R^2 = 0.078 Sig (F-change$_{model1}$) = 0.011					
Intercept	108.41	1.45		74.818	<0.001
Gender	−4.03	2.10	−0.17	−1.92	0.057
EQ	−0.02	0.11	−0.01	−0.15	0.881
SQ	−0.06	0.06	−0.10	−1.11	0.268
Math Anxiety	−0.13	0.05	−0.22	−2.59	0.011
Model 4 R^2 = 0.147, Sig (F-change$_{model3}$) = 0.005					
Intercept	107.77	1.45		74.24	<0.001
Gender	−3.75	2.04	−0.15	−1.84	0.068
EQ	−0.37	0.16	−0.12	−2.37	0.019
SQ	−0.04	0.08	−0.06	−0.45	0.651
Math Anxiety	−0.15	0.05	−0.25	−2.96	0.004
Gender * EQ	0.64	0.21	0.36	3.10	0.002
Gender * SQ	−0.03	0.11	−0.03	−0.23	0.818

Calculation Model 3: Effects of gender, empathy quotient (EQ), systemizing quotient (SQ) and math anxiety. Model 4: additional variance accounted for by interaction of Gender with SQ and EQ.

TABLE 5 | Hierarchical regression analysis of Calculation accounting for SRS.

	B	*SE*	Beta	*t*	*p*
Model 5 R^2 = 0.038 Sig (F-change$_{model1}$) = 0.390					
Intercept	108.65	1.48		76.59	< 0.001
Gender	−4.50	2.13	−0.19	−2.11	0.037
EQ	0.09	0.13	0.08	0.69	0.494
SQ	−0.05	0.06	−0.08	−0.83	0.408
SRS	0.05	0.06	0.09	0.86	0.390
Model 6 R^2 = 0.095, Sig (F-change$_{model5}$) = 0.004					
Intercept	108.38	1.44		75.22	<0.001
Gender	−4.65	2.08	−0.19	−2.24	0.027
EQ	0.07	0.13	0.06	0.55	0.580
SQ	−0.06	0.06	−0.10	−1.06	0.290
SRS	0.16	0.07	0.29	2.40	0.018
Gender * SRS	−0.26	0.09	−0.32	−2.91	0.004
Model 7 R^2 = 0.112, Sig (F-change$_{model6}$) = 0.276					
Intercept	108.37	1.47		73.51	<0.001
Gender	−4.52	2.08	−0.19	−2.18	0.031
EQ	−0.08	0.19	−0.07	−0.41	0.685
SQ	−0.09	0.09	−0.14	−1.02	0.310
SRS	0.12	0.07	0.23	1.68	0.096
Gender * SRS	−0.17	0.11	−0.21	−1.52	0.131
Gender * EQ	0.30	0.25	0.18	1.18	0.240
Gender * SQ	0.07	0.12	0.08	0.63	0.529

Model 5: Effects of gender, empathy quotient (EQ), systemizing quotient (SQ) and social responsiveness scale (SRS) scores. Model 6: additional variance accounted for by interaction of Gender with SRS. Model 7: additional variance accounted for by interaction of Gender with SQ and EQ.

Gender Differences in the Relations Between Emotional Empathy and Math Achievement

To assess the construct validity of the EQ components of Cognitive Empathy (EQ-CE), Emotional Empathy (EQ-EE) and Social Skills (EQ-SS), we first examined their correlation with SRS scores. As SRS is an index of social skills, we expected the strongest relationship to be between SRS and the EQ-SS component. Indeed, we found that SRS correlated most strongly with EQ-SS [$r(142) = -0.532$, $p < 0.001$], then with EQ-EE [$r(142) = -0.423$, $p < 0.001$] and finally with EQ-CE [$r(142) = -0.313$, $p < 0.001$]. Next we examined gender differences between the EQ components. While there was no significant difference between men and women on the full EQ (**Table 1**), women had significantly higher scores on the EQ-EE component [$t(140) = -3.557$, $p = 0.001$], but not EQ-CE or EQ-SS (all ps > 0.53, see **Figure 2**).

In unplanned follow-up analyses, we sought to assess which aspect of the EQ drove the positive relationship between empathy and mathematical achievement, in women. In the full sample, there was a marginal positive relationship between EE-SS and Math Fluency [$r(142) = 0.152$, $p = 0.071$]. None of the other components were related to math achievement in the full sample (all ps > 0.28). However, in women, the EQ-EE correlated significantly with math achievement for both Calculation and Math Fluency and there was a marginal relationship between EQ-SS and Math Fluency (**Supplementary Table S2**). In men there was a marginal negative relationship between EQ-SS and Calculation (**Supplementary Table S2**). No other components correlated with math achievement in men.

To further quantify these gender differences, we performed multiple regression analyses using the three EQ components with the two math measures as dependent variables. In a first model we examined the main effects of gender and each of the components; we then added the interaction terms of each of the components to the model. For Calculation, adding the interaction terms resulted in a significant F-change relative to Model S1 [$F(3,134) = 4.26$, $p = 0.007$, **Supplementary Table S3**, Model S2] and there was a significant interaction of gender and the EQ-EE component, and marginal interactions of gender with EQ-CE and EQ-SS. For Math Fluency, adding the interaction terms resulted in a marginal F-change relative to Model S3 [$F(3,134) = 2.58$, $p = 0.056$] and only EQ-EE interacted significantly with gender (**Supplementary Table S4**, Model S4).

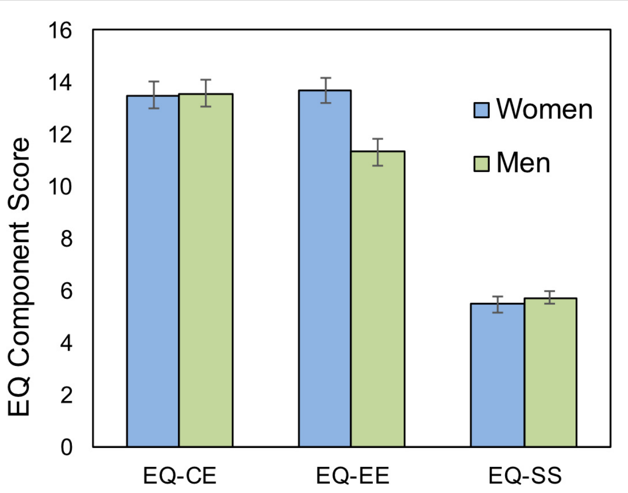


FIGURE 2 | Scores on the Empathy Quotient (EQ) component scores (Study 1). Women scored higher than men on the Emotional Empathy (EE) component of the EQ, but there was no differences on the Cognitive Empathy (CE) and Social Skills (SS) components.

Discussion

In Study 1, we examined the relationship between empathizing, systemizing and math achievement in women and men. We found that among women, both EQ and SQ were positively related to math achievement while the opposite pattern held in men. There was a significant interaction between gender and empathizing in the prediction of Calculation scores, but not Math Fluency, and a marginal interaction between gender and SQ in predicting Math Fluency, but no effect in Calculation. Moreover, the results for Calculation were independent of the strong negative relationship between math anxiety and math achievement found in both men and women.

These results in women are surprising given the negative relationship between math achievement and empathizing found in children by Escovar et al. (2016). Although there was no significant difference in the strength of these effects between boys and girls, follow up analyses found a stronger role for social skills among girls than boys on math performance. Specifically, social skills mediated the relationship between EQ and math achievement in girls, but not boys. Those results are in line with a negative relationship of math achievement with empathizing and social skills more broadly in females. Yet, in the current study we found a positive relationship in women. Moreover, SRS also tracked with EQ in this sample, that is, better social skills were related to better math performance in women, and worse performance in men. Further, including SRS and EQ in the same model (Model 7) removed the interaction between EQ and gender because of the multicollinearity between EQ and SRS.

Additional support for the role of social skills in math performance came from the examination of EQ components scores. There was a marginal relationship between EQ-SS scores and Math Fluency in the full sample and a marginal interaction between gender and EQ-SS in Calculation scores. Less expected were the gender differences in EQ-EE, both in total scores and in the interaction between gender and EQ-EE in predicting Math Fluency and Calculation scores. Cognitive empathy has been associated with theory of mind (Lawrence et al., 2004), which has been linked to executive functions (Perner and Lang, 1999), and we found a marginal interaction between gender and this component in predicting Calculation scores.

Given the unexpected positive relationship between empathizing and math skills found among women in Study 1, we conducted a second study aimed at replicating the direction of this effect. We also sought to assess the role of another individual difference trait known to be related to math skills: attitudes toward mathematics. Surveys of attitudes toward math assess the thoughts and feelings that students bring to mathematics beyond whether it induces anxiety. Notably, attitudes toward math predict math achievement both concurrently (Ma and Kishor, 1997) and longitudinally (Anttonen, 1969) and tend to decline over the course of schooling (Watt, 2004). Further, females tend to have lower levels of positive attitudes toward math (Watt, 2004). Coupled with the growing gap in achievement between the genders starting in middle school (Hyde et al., 1990), these findings suggest that changes in attitudes toward mathematics over the course of schooling may correspond to the internalization of gender stereotypes regarding math achievement. Thus, we asked participants to retrospect over their attitudes from elementary school to the college level.

Our primary goal in Study 2 to was to replicate the unexpected finding of a positive relationship between EQ and math achievement in female college students. Therefore, we restricted our sample to women only. Further, our online method of data collection precluded obtaining standardized measures of math achievement. Instead, we asked students to report their grade point averages (GPA) on college-level mathematics coursework, as well as complete questionnaire measures. Our second goal, was to replicate the association between the emotional empathy components of the EQ and math achievement. Our final goal was to examine the relationship between changes in attitudes toward mathematics and math skills and see if that could explain the relationship between EQ and math achievement in women. We predicted that positive attitudes would decline from elementary school into adulthood, as this is a crucial period when women may be discouraged from pursuing STEM careers. Further, we explored whether changing attitudes toward mathematics could explain the positive relationship between empathizing and math achievement found in Study 1.

STUDY 2

Methods and Procedures

Participants

Hundred and twenty one college-aged female students attending Rutgers-University in Newark, NJ, United States, participated in the study, ranging in age from 18 and 27. As in Study 1, participants received credit toward their psychology coursework.

Empathizing, Systemizing and Math Anxiety Measures

Scoring and missing data procedures for the EQ, SQ, and A-MARS followed the same procedures as in Study 1, including computing the component scores from the EQ. Six participants were missing one item (0.12%) and Cronbach's alpha for EQ was 0.845. For the SQ there were 21 instances of missing items, two participants were missing two items and the remainder were missing one for a total of 0.23% missing data. Cronbach's alpha for the SQ was 0.832. For the A-MARS, 7 participants were missing one item for a total of 0.24% missing data and Cronbach's alpha was 0.964.

Retrospective Attitude Toward Math Measure

The Attitudes Toward Mathematics questionnaire is a 40-item instrument that measures a person's attitudes toward mathematics beyond math anxiety (Tapia and Marsh, 2004). We adapted this instrument to produce a brief measure of participants' attitudes toward math over the course of schooling, which we termed Retrospective Attitudes Toward Mathematics (R-ATM). Participants were asked to rate statements on 5-point Likert scale (*strongly disagree, disagree, neutral, agree,* and *strongly agree*) regarding three core attitudes toward math: ability (*I was good at mathematics*), enjoyment (*I liked mathematics*) and importance (*I thought it was important to do well in math class*). We asked participants to retrospect over four stages of schooling: elementary school, middle school, high school and college. Higher values reflect more positive attitudes toward math. We computed a measure of participants change in attitudes by fitting a line to the four time points, using the formula for slope in a linear regression. There were five instances of missing items (0.41%), one participant was missing all three items from elementary school and two other participants were missing one item each. Missing items were replaced by the average of the other items for that time bin. The participant missing all elementary school items was excluded from the repeated measures ANOVA, but included in the slope analyses. The Cronbach's alpha for the R-ATM was 0.882.

Math Measure

Participants were asked to report their GPA on college level math courses. Previous research has found high correlations between students self-report academic grades and their actual grades, with slightly higher reliability for mathematics relative to language studies (Sticca et al., 2017). We only considered responses from participants who reported results on the 4.0 scale used at Rutgers University – Newark (n = 80). Math GPA deviated from normality (*skewness* = −0.774, *SE* = 0.269; *kurtosis* = 0.500, *SE* = 0.532), so following Coyle et al. (2011) we computed the square of Math GPA (*skewness* = −0.162, *SE* = 0.269; *kurtosis* = −0.755, *SE* = 0.532). As deviations from normality can distort error estimates we used the transformed (squared) Math GPAs for the analyses, but report the original values in **Table 6**.

TABLE 6 | Participant characteristics (Study 2).

	M	*SD*
Age[1]	20.864	1.927
Math GPA[2]	3.103	0.712
EQ	44.975	10.401
SQ	57.924	15.104
Math Anxiety	71.625	24.062

GPA, Grade Point Average. Abbreviations as in ***Table 1****. [1] Three participants did not report their date of birth; therefore, the sample size is n = 118. [2] Only 80 participants reported Math GPA on 4.0 scale.*

Procedure

Data were collected online and consisted entirely of questionnaires and demographic information. Participants completed the EQ, SQ, and R-ATM. Finally, participants completed the demographics questionnaire, including their GPA in college level math. The procedure took 30–45 min.

Results

Relations Between Math Achievement, Systemizing and Empathizing

Participant characteristics are listed **Table 6**. Math achievement, as measured by self-report of Math GPA, was positively related to empathizing in this female-only sample, although the relationship was only marginally significant (**Table 7** and **Figure 3**). This result extends the findings of Study 1 by replicating the positive relationship between EQ and math achievement in women, now in the domain of self-reported math performance at the college-level. Despite the strong positive correlation between EQ and SQ, there was no relationship between SQ and math in this sample. Next, we examined EQ component scores and their relations with Math GPA. Consistent with Study 1, EQ-EE had the strongest relationship of the EQ components with math skills, although it was only marginally significant [$r(80) = 0.218$, $p = 0.052$]. There was no relationship between Math GPA and EQ-CE [$r(80) = 0.163$, $p = 0.148$], or EQ-SS [$r(80) = 0.086$, $p = 0.451$].

Relations Between Empathizing and Math Achievement After Controlling for Math Anxiety

Math anxiety negatively correlated with math achievement, as measured by Math GPA (**Table 7**), while there was no relationship between math anxiety and EQ, either in the full sample (**Table 7**), or among the subset of individuals who reported their Math GPA [$r(80) = -0.114$, $p = 0.315$]. Using partial correlations, after controlling for math anxiety, there was no longer even a marginal relationship between EQ and math achievement [$r(80) = 0.180$, $p = 0.112$].

Relations Between Attitudes Toward Math, Math Achievement and Empathizing

Next, we considered the progression of participants' attitudes toward mathematics over the course of formal schooling. Participants' scores on the R-ATM measure declined from elementary, to middle, and into secondary school, and university (**Figure 4**). A repeated measures ANOVA revealed a significant main effect of time [$F(1,119) = 17.538$, $p < 0.001$]. Within-participant contrasts revealed a marginal decline from elementary to middle school [$F(1,119) = 3.498$, $p = 0.064$], a significant decline from middle to high school [$F(1,119) = 18.832$, $p < 0.001$], but no significant decrease from high school to university [$F(1,119) = 1.196$, $p = 0.276$] (**Figure 3**).

Next, we examined whether change in participants attitudes toward math was related to math achievement and the relationship between math achievement and EQ. Slope R-ATM was highly correlated with math skills [$r(80) = 0.425$, $p < 0.001$], such that greater declines in attitudes were associated with worse Math GPA. There was also a marginal correlation between

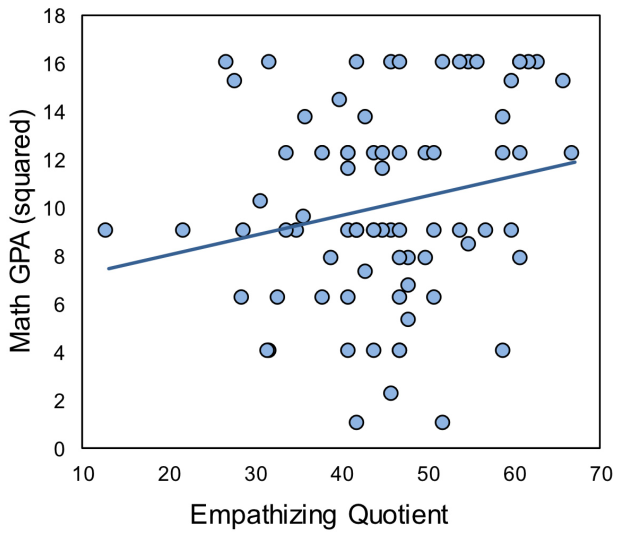


FIGURE 3 | Relations between EQ and Math GPA (Study 2). Participants' self-report Math GPA (Grade Point Average) squared was marginally correlated with the scores on the Empathy Quotient [$r(80) = 0.210$, $p = 0.061$].

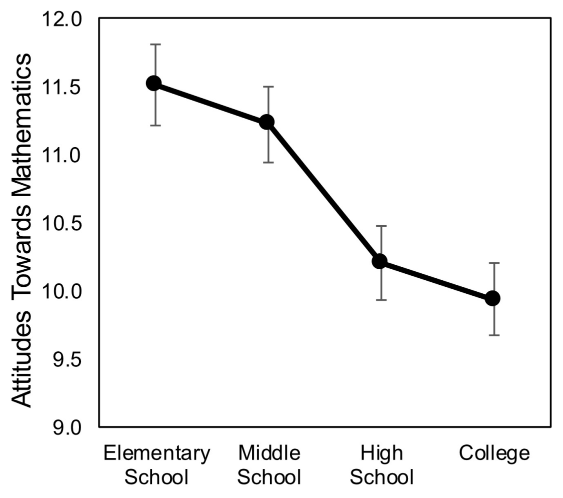


FIGURE 4 | Retrospective attitudes toward math (Study 2). Participants' attitudes toward mathematics declined over the course of their schooling.

TABLE 7 | Pearson correlations between Math Grade Point Average (GPA), EQ, SQ, and Math Anxiety in women (Study 2).

Measures	EQ	SQ	Math Anxiety
Math GPA[1]	0.210	0.136	−0.423***
EQ		0.475***	−0.019
SQ			0.130

Abbreviations and significance as in ***Table 2***. *[1]Only 80 participants reported Math GPA, so correlations with this measure are for n = 80. The square of Math GPA was used for correlational analyses. All other measures have n = 121. ***Correlation is significant at the 0.001 level (two-tailed).*

EQ and Slope R-ATM [$r(121) = 0.167$, $p = 0.067$], such that participants with higher EQ scores declined less than those with lower EQ scores. Further, partial correlations revealed that controlling for Slope R-ATM, EQ was no longer even marginally related to math skills [$r(80) = 0.157$, $p = 0.168$].

Finally, to quantify the relationship between changing attitudes toward math and EQ, we noted that 27 out of 121 participants had positive Slope R-ATM values, indicating increasing attitudes toward math (and 15 were flat). As would be expected by the correlational analysis, these participants had significantly higher EQ scores than the 79 students who declined [$M = 49.05$, $SD = 10.89$ vs. $M = 43.64$, $SD = 9.44$, $t(104) = 2.468$, $p = 0.015$]. We report this result to illustrate that this difference of 6 points represents more than half of a standard deviation in the distribution of EQ scores ($M = 44.98$, $SD = 10.40$, **Table 6**). Together these results suggest that individuals with higher EQ scores tend to decline less in their attitudes toward mathematics over the course of schooling.

Discussion

Consistent with Study 1, we found that math achievement, as measured by self-report of GPA, was positively related to empathizing in a large sample of female college students, although the effect was only marginally significant. Moreover, attitudes toward mathematics declined over the course of schooling, but empathizing acted as a protective factor, with higher EQ levels amongst those whose attitudes improved relative to those who declined. Together these results confirm the unexpected finding of Study 1 and further suggest a mechanism for reconciling these results with those of Escovar et al. (2016), as discussed in the Section "General Discussion."

GENERAL DISCUSSION

Consistent with prior literature, we found that men had a greater tendency to systemize than women (Groen et al., 2015), however, we did not find a significant difference between the genders in the tendency to empathize. In a recent meta-analysis, Groen et al. (2015) report that effect sizes for the SQ tend to be larger than for EQ, consistent with our results. Interestingly, when we examined the EQ components, women scored significantly higher than men on the emotional empathy component, but not on the cognitive empathy or social skills components (**Figure 2**), in line with prior work showing that EQ-EE scores display the strongest gender differences (Groen et al., 2015, 2018) among neurotypical individuals.

In two samples, we found that for women, scores on the EQ were positively related to math achievement. In men, correlations were either negative or flat between EQ and math achievement. This result in women is surprising given the finding that higher EQ is related to lower math performance in children, especially girls (Escovar et al., 2016). To better understand the individual differences driving these effects we examined several correlates and components of empathy, including cognitive empathy, emotional empathy, and social skills. We also considered other known predictors of math skills, namely math anxiety and attitudes toward mathematics.

Social Skills and the Relationship Between Empathizing and Math Achievement, in Women and Men

Empathy and social skills are intimately linked, as the capacities to identify the emotions of others and respond appropriately are crucial precursors to fluent social interactions. While strong social skills could be an asset in a STEM career, prior research suggests that math intensive majors (mathematics, computer science, and physics) have some of the poorest social skill levels among science majors (Baron-Cohen et al., 2001) as do math-intensive majors relative to humanities and social sciences (Valla et al., 2010). Specifically, social skills, as measured by the social interaction factor of the autism quotient (AQ) and the reading the mind's eye task predicted men's choice of field, while there was no relationship in women (Valla et al., 2010).

In the current study, we found some support for the negative relationship of empathizing and social skills with math achievement in men (**Table 2**). However, rather than no relationship, among women we found that better social skills were related to better math performance. In Study 1, we found a significant main effect of SRS scores and significant interaction of gender with SRS in predicting Calculation scores, indicating both that men with worse social skills had better math performance and that women with better social skills had better math performance (**Table 5**, Model 6). Moreover, SRS explained the interaction of gender and EQ in predicting math skills (**Table 5**, Model 7). While the EQ component results were less definitive, we did find a marginal positive relationship with Math Fluency in the full group and in women, and a marginal negative relationship with Calculation in men, resulting in a marginal interaction between EQ-SS and gender in predicting Calculation.

Individuals with strong math and verbal skills tend to choose STEM careers at lower rates than individuals with strong math, but moderate verbal skills. This first group is overwhelmingly female and values interacting with people as part of their career (Wang et al., 2013) In contrast, men are over represented among those with strong math but moderate verbal skills, and may pursue STEM majors for lack of other options. Potentially, women with high verbal and math skills may also have strong social skills, further alienating them from less welcoming STEM fields (Ceci et al., 2009). In the current study, we only examined math performance. Further work is needed to assess how social skills might impact not just math achievement but the choice of STEM career.

Cognitive Empathy and Math Achievement

Cognitive empathy has been used synonymously with theory of mind (Lawrence et al., 2004), which itself has been linked to executive functions (Perner and Lang, 1999), which in turn, strongly predicts math achievement (Clark et al., 2010; Cragg and Gilmore, 2014). Thus, we might expect that the cognitive empathy aspects of empathy would be most closely aligned with math skills. Yet, there were no differences between the genders in scores on the EQ-CE, nor did it correlate with math performance in either sample of women (although there was a marginal interaction between EQ-CE and gender in predicting Calculation scores). Another way to assess the impact of executive functions on the relationship between EQ and math achievement comes from differences between the two standardized math measures. Calculation likely draws more on working memory, while Math Fluency engages task switching (between operations) and inhibitory control (inhibiting answers from different operations). Yet, we did not find consistently stronger relationships for one measure or the other. Future work, targeted at comprehensive cognitive assessments of executive function (and intelligence and other academic skills) is needed to fully characterize the extent, and specificity, of the relationship between EQ and math achievement in women.

Changing Attitudes Toward Math and Emotional Empathy

The current studies revealed a consistent finding that EQ is positively related to math achievement in college-attending women, yet Escovar et al. (2016) reported a negative relation in children. Clearly, longitudinal studies, which track students' math achievement and tendency to empathize from elementary school to college are needed to confirm whether, within the same individuals, empathy can shift from a negative to a positive predictor of math skills.

If these results are borne out in such studies, one possible explanation is that empathy itself changes over development, accounting for its changing relationship to math achievement. Decety and Lamm (2006) have characterized empathy as involving both a bottom–up component of sharing the emotions of others, and a top–down component of regulating those emotions. Moreover, while the bottom–up component is developed by the age of 3 years, the top–down component continues to develop from childhood into adolescence and adulthood (Decety, 2010). From this perspective, we would expect that girls with higher empathy might more readily absorb messages about gender stereotypes, impacting math performance in childhood (Beilock et al., 2010). But in adulthood, individuals with high empathy may also have a stronger capacity to regulate emotions and hence distance themselves from these stereotypes. That is, women with higher empathy may be better able to recognize that cultural stereotypes are just rules of society rather than laws of nature, while those with lower empathy may be more likely to internalize these stereotypes.

Two sets of results in the current studies support this conclusion. First, contrary to the general pattern of declining attitudes toward math (Watt, 2004), a subset of women in Study 2 actually increased their positive attitudes toward math, and these had higher EQ scores relative to those who declined. While we cannot assess the direction of this effect, it is possible that women with a greater tendency to empathize may be more able to maintain a positive attitude toward mathematics in the face of negative societal messages. Second, emotional empathy had the highest correlation with Math GPA in Study 2 and the interaction of EQ-EE and gender significantly predicted both math measures in Study 1. These results suggest that the ability to react appropriately to the emotions of others may be a particularly useful skill in navigating these waters.

Math Anxiety and Math Achievement

Math anxiety is one of the most robust, non-cognitive predictors of math achievement (Ma, 1999). Consistent with this body of research, we found that a self-report measure of math anxiety correlated with math performance, both with standardized measures of math achievement (Study 1) and performance in college-level course work (Study 2). There were no differences between men and women in the rates of math anxiety, nor was the relationship between math anxiety and math achievement modulated by gender, as both genders displayed a negative relationship between math anxiety and math skills.

This pattern of consistent relations between the genders contrasts the finding of no main effect of EQ on math achievement, but instead an interaction between gender and EQ. Further, statistically accounting for anxiety did not change the interaction between EQ and gender in the prediction of Calculation performance in Study 1. However, the inclusion of math anxiety did affect the relationship between math GPA and EQ in Study 2. The negative relationship between math anxiety and math achievement in women was much stronger in Study 2 than in Study 1 (**Tables 2**, **7**), potentially reflecting a greater influence of anxiety on math performance during actual coursework relative to the lower stakes of volunteer laboratory experiment. Larger samples are needed to further asses the unique contributions of EQ on math achievement, independent of the effects of math anxiety. That said, these results further highlight the gendered nature of the relationship between empathizing and math skills, suggesting that empathy may be a particularly important factor for understanding women's achievement in mathematics.

Systemizing and Math Achievement

Although the primary focus of these studies was empathizing, we also examined the role of systemizing. In fact, there is considerable theoretical support for the role of systemizing in math achievement, as mathematics is a paradigmatic example of a system. Yet, there is surprisingly little empirical support for the role of systemizing and math achievement. For example, Escovar et al. (2016) found a marginally significant relationship between SQ and performance on the Applied Problems subtest of the WJ-III, which was not significant in either gender. In Study 1, we found some hints of a positive relationship between SQ and math in women (e.g., Math Fluency), but the opposite pattern in men. Moreover, directly comparing these effects found no significant main effects or interactions between gender and SQ in predicting math achievement. In Study 2 there was no relationship between self-report of math achievement and SQ. Together these results suggest that there may be a small effect of SQ on math achievement among women, and studies with larger samples will be needed to detect these relationships.

These findings, coupled with other null results between SQ and math achievement (Morsanyi et al., 2012), raise the possibility that the Systemizing Quotient may be a poor measure of the construct of systemizing. Notably, our findings of strong positive correlations between EQ and SQ, in both men and women, conflicts with both Empathizing – Systemizing Theory (Baron-Cohen, 2009), which suggests they should be negatively correlated, and the Extreme Male Brain Theory (Baron-Cohen, 2002), which suggests they should be independent. In the current study, we used multiple regression to account for any collinearity between the measures. In the broader context, other measures of systemizing might be more informative. For example, Valla and colleagues found that among women, higher scores on the details/pattern factor of the AQ was related to pursuing degrees in fields which involve more systemizing. However, other putative measures of systemizing, such as the embedded figures test and a Go/No Go task were unrelated to field of study in either gender (Valla et al., 2010). How best to conceptualize and measure systemizing remains an open question, especially in applications beyond the original purpose of characterizing autism phenotypes.

CONCLUSION

In this study, we examined the relationships between math achievement, and two individual traits, empathizing and systemizing. Contrary to our expectations, we found a positive relationship between math achievement and empathizing in women, that is, higher empathy was related to better math performance, both on standardized measures of math achievement and college-level math attainment. In contrast, among men, lower empathy was related to better math performance. These differential patterns illustrate the complex role gender plays in the path to a STEM career. More broadly, these results contradict the view that traditionally female traits are incompatible with success in mathematics. In fact, rather than empathy acting as an obstacle, it seems to support women's mathematics achievement, potentially facilitating their pursuit of STEM careers.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations for written (pen/paper or electronic) informed consent of the Arts and Sciences Institutional Review Board, at Rutgers University. The protocol was approved by the Arts and Sciences Institutional Review Board at Rutgers University. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

MR-L conceived and designed the studies. NG and ER collected the data for Study 1. NG collected the data for Study 2. MR-L, NG, and ER analyzed the data and wrote the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01941/full#supplementary-material

REFERENCES

Alexander, L., and Martray, C. (1989). The development of an abbreviated version of the mathematics anxiety rating scale. *Meas. Eval. Couns. Dev.* 22, 143–150. doi: 10.1080/07481756.1989.12022923

American Psychiatric Association. (2013). *Neurodevelopmental Disorders Diagnostic and Statistical Manual of Mental Disorders*, 5Edn Edn. Washington, DC: American Psychiatric Association.

Anttonen, R. G. (1969). A longitudinal study in mathematics attitude. *J. Educ. Res.* 62, 467–471. doi: 10.1080/00220671.1969.10883904

Ashcraft, M. H. (2002). Math anxiety: personal, educational, and cognitive consequences. *Curr. Dir. Psychol. Sci.* 11, 181–185. doi: 10.1111/1467-8721.00196

Auyeung, B., Wheelwright, S., Allison, C., Atkinson, M., Samarawickrema, N., and Baron-Cohen, S. (2009). The children's empathy quotient and systemizing quotient: sex differences in typical development and in autism spectrum conditions. *J. Autism. Dev. Disord.* 39, 1509–1521. doi: 10.1007/s10803-009-0772-x

Bakker, M., Torbeyns, J., Wijns, N., Verschaffel, L., and De Smedt, B. (2019). Gender equality in 4- to 5-year-old preschoolers' early numerical competencies. *Dev. Sci.* 22:e12718. doi: 10.1111/desc.12718

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends Cogn. Sci.* 6, 248–254. doi: 10.1016/s1364-6613(02)01904-6

Baron-Cohen, S. (2009). Autism: the empathizing-systemizing (E-S) theory. *Ann. N.Y.Acad. Sci.* 1156, 68–80. doi: 10.1111/j.1749-6632.2009.04467.x

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., and Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high–functioning autism, and normal sex differences. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 361–374. doi: 10.1098/rstb.2002.1206

Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism. Dev. Disord.* 34, 163–175. doi: 10.1023/b:jadd.0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism. Dev. Disord.* 31, 5–17.

Beede, D., Julian, T., Langdon, D., McKittrick, G., Khan, B., and Doms, M. (2011). *Women in STEM: A Gender Gap to Innovation*. Washington, DC: U.S. Department of Commerce.

Beilock, S. L., Gunderson, E. A., Ramirez, G., and Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1860–1863. doi: 10.1073/pnas.0910967107

Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., and Kim, S. (2011). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Soc. Psychol. Personal. Sci.* 2, 656–664. doi: 10.1177/1948550611405218

Clark, C. A. C., Pritchard, V. E., and Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Dev. Psychol.* 46, 1176–1191. doi: 10.1037/a0019672

Constantino, J. N., and Gruber, C. P. (2012). *Social Responsiveness Scale (SRS)*. Torrance, CA: Western Psychological Services.

Coyle, T., Snyder, A., Pillow, D., and Kochunov, P. (2011). SAT predicts GPA better for high ability subjects: implications for Spearman's Law of Diminishing Returns. *Pers. Individ. Dif.* 50, 470–474. doi: 10.1016/j.paid.2010.11.009

Cragg, L., and Gilmore, C. (2014). Skills underlying mathematics: the role of executive function in the development of mathematics proficiency. *Trends Neurosci. Educ.* 3, 63–68. doi: 10.1016/j.tine.2013.12.001

Crisp, G., Nora, A., and Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: an analysis of students attending a hispanic serving institution. *Am. Educ. Res. J.* 46, 924–942. doi: 10.3102/0002831209349460

Decety, J. (2010). The neurodevelopment of empathy in humans. *Dev. Neurosci.* 32, 257–267. doi: 10.1159/000317771

Decety, J., and Lamm, C. (2006). Human empathy through the lens of social neuroscience. *Sci. World J.* 6, 1146–1163. doi: 10.1100/tsw.2006.221

Escovar, E., Rosenberg-Lee, M., Uddin, L. Q., and Menon, V. (2016). The empathizing-systemizing theory, social abilities, and mathematical achievement in children. *Sci. Rep.* 6:23011. doi: 10.1038/srep23011

Greenberg, D. M., Warrier, V., Allison, C., and Baron-Cohen, S. (2018). Testing the empathizing-systemizing theory of sex differences and the extreme male brain theory of autism in half a million people. *Proc. Natl. Acad. Sci. U.S.A.* 115, 12152–12157. doi: 10.1073/pnas.1811032115

Groen, Y., den Heijer, A. E., Fuermaier, A. B. M., Althaus, M., and Tucha, O. (2018). Reduced emotional empathy in adults with subclinical ADHD: evidence from the empathy and systemizing quotient. *Atten. Defic. Hyperact. Disord.* 10, 141–150. doi: 10.1007/s12402-017-0236-7

Groen, Y., Fuermaier, A. B., Den Heijer, A. E., Tucha, O., and Althaus, M. (2015). The empathy and systemizing quotient: the psychometric properties of the dutch version and a review of the cross-cultural stability. *J. Autism. Dev. Disord.* 45, 2848–2864. doi: 10.1007/s10803-015-2448-z

Hyde, J. S., Fennema, E., and Lamon, S. J. (1990). Gender differences in mathematics performance: a meta-analysis. *Psychol. Bull.* 107, 139–155. doi: 10.1037/0033-2909.107.2.139

Iuculano, T., Rosenberg-Lee, M., Supekar, K., Lynch, C. J., Khouzam, A., Phillips, J., et al. (2014). Brain organization underlying superior mathematical abilities in children with autism. *Biol. Psychiatry* 75, 223–230. doi: 10.1016/j.biopsych.2013.06.018

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., and David, A. S. (2004). Measuring empathy: reliability and validity of the empathy quotient. *Psychol. Med.* 34, 911–920. doi: 10.1017/S0033291703001624

Leslie, S.-J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 262–265. doi: 10.1126/science.1261375

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *J. Res. Math. Educ.* 30, 520–540. doi: 10.2307/749772

Ma, X., and Kishor, N. (1997). Assessing the relationship between attitude toward mathematics and achievement in mathematics: a meta-analysis. *J. Res. Math. Educ.* 28, 26–47. doi: 10.2307/749662

Morsanyi, K., Primi, C., Handley, S. J., Chiesi, F., and Galli, S. (2012). Are systemizing and autistic traits related to talent and interest in mathematics and engineering? Testing some of the central claims of the empathizing–systemizing theory. *Br. J. Psychol.* 103, 472–496. doi: 10.1111/j.2044-8295.2011.02089.x

Perner, J., and Lang, B. (1999). Development of theory of mind and executive control. *Trends Cogn. Sci.* 3, 337–344. doi: 10.1016/s1364-6613(99)01362-5

Rueda, P., Fernández-Berrocal, P., and Baron-Cohen, S. (2015). Dissociation between cognitive and affective empathy in youth with Asperger Syndrome. *Eur. J. Dev. Psychol.* 12, 85–98. doi: 10.1080/17405629.2014.950221

Shapiro, J. R., and Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles* 66, 175–183. doi: 10.1007/s11199-011-0051-0

Sticca, F., Goetz, T., Bieg, M., Hall, N. C., Eberle, F., and Haag, L. (2017). Examining the accuracy of students' self-reported academic grades from a correlational and a discrepancy perspective: evidence from a longitudinal study. *PLoS One* 12:e0187367. doi: 10.1371/journal.pone.0187367

Tapia, M., and Marsh, G. E. (2004). An instrument to measure mathematics attitudes. *Acad. Exch. Q.* 8, 12–21.

Treffert, D. A. (2009). The savant syndrome: an extraordinary condition. A synopsis: past, present, future. *Philos. Trans. R. Soc. B.Biol. Sci.* 364, 1351–1357. doi: 10.1098/rstb.2008.0326

Valla, J. M., Ganzel, B. L., Yoder, K. J., Chen, G. M., Lyman, L. T., Sidari, A. P., et al. (2010). More than maths and mindreading: sex differences in empathizing/systemizing covariance. *Autism Res.* 3, 174–184. doi: 10.1002/aur.143

Wakabayashi, A., Baron-Cohen, S., Uchiyama, T., Yoshida, Y., Kuroda, M., and Wheelwright, S. (2007). Empathizing and systemizing in adults with and without autism spectrum conditions: cross-cultural stability. *J. Autism Dev. Disord.* 37, 1823–1832. doi: 10.1007/s10803-006-0316-6

Wang, M.-T., Eccles, J. S., and Kenny, S. (2013). Not lack of ability but more choice: individual and gender differences in choice of careers in science. *Technol. Eng. Math. Psychol. Sci.* 24, 770–775. doi: 10.1177/0956797612458937

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th- through 11th-grade australian students. *Child Dev.* 75, 1556–1574. doi: 10.1111/j.1467-8624.2004.00757.x

Wood, F. S. (1984). Comment. *Am. Stat.* 38, 88–90. doi: 10.1080/00031305.1984.10483173

Wu, S. S., Barth, M., Amin, H., Malcarne, V., and Menon, V. (2012). Math anxiety in second and third graders and its relation to mathematics achievement. *Front. Psychol.* 3:162. doi: 10.3389/Fpsyg.2012.00162

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 13 September 2019
doi: 10.3389/fpsyg.2019.02065

Social or Economic Goals? The Professional Goal Orientation of Students Enrolled in STEM and Non-STEM Majors in University

Ilka Wolter[1], Lisa Ehrtmann[1], Tina Seidel[2] and Barbara Drechsel[3]*

[1]Competencies, Personality, Learning Environments, Leibniz Institute for Educational Trajectories, Bamberg, Germany, [2]Friedl Schöller Endowed Chair for Educational Psychology, School of Education, Technical University of Munich (TUM), Munich, Germany, [3]Institute of Psychology, University of Bamberg, Bamberg, Germany

OPEN ACCESS

Edited by:
Silke Luttenberger, University of Teacher Education Styria, Austria

Reviewed by:
Elizabeth Stearns, University of North Carolina at Charlotte, United States
Denise Wilson, University of Washington, United States

***Correspondence:**
Ilka Wolter
ilka.wolter@lifbi.de

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *15 January 2019*
Accepted: *26 August 2019*
Published: *13 September 2019*

Citation:
Wolter I, Ehrtmann L, Seidel T and Drechsel B (2019) Social or Economic Goals? The Professional Goal Orientation of Students Enrolled in STEM and Non-STEM Majors in University. Front. Psychol. 10:2065. doi: 10.3389/fpsyg.2019.02065

Various studies try to disentangle the gender-specific competencies or decisions that lead to a career in a STEM field and try to find a way to encourage more women to pursue this kind of career. The present study examines differences in the meaning of work (i.e., their professional goal orientation) of students who are enrolled in STEM or non-STEM programs in tertiary education. Based on the background that gender stereotypes associate women and men with communal or agentic roles respectively, we expected that women in STEM subjects differ in their professional goal orientation from women in non-STEM programs. More precisely, women who are enrolled in a STEM major are expected to be less oriented to social and communal goal orientations than women in non-STEM university programs. In a sample of 5,857 second-year university students of the German National Educational Panel Study, three profiles of professional goal orientation were confirmed in a latent profile analysis. As expected, women were more oriented toward social aspects of occupations, whereas men more likely belonged to a profile with high importance for economic aspects of occupations. Moreover, students enrolled in STEM programs more likely belonged to the profile of economic goal orientation. There was, however, no interaction of gender and STEM program: Women in STEM fields did not differ in their occupational goal orientation from women enrolled in non-STEM programs. Based on these findings and on a goal congruity perspective, future interventions aiming at overcoming the underrepresentation of women in STEM fields should consider the individual meaning of work and the goals that are associated with STEM occupations.

Keywords: STEM, goal orientation, latent profile analysis, university students, gender stereotypes, meaning of work

THEORETICAL BACKGROUND

Gender segregation in the labor market and in university majors is a widely known and consistent pattern of previous empirical research (e.g., Leuze and Strauß, 2009; Bechmann et al., 2012; Ochsenfeld, 2012; Hausmann and Kleinert, 2014), showing in detail that women are especially underrepresented in STEM fields (i.e., science, technology, engineering, and mathematics). This segregation is particularly concerning because STEM fields are mostly characterized by higher prestige and income (e.g., Oh and Lewis, 2011). However, research also showed that the mechanisms of this gap need further investigation and differentiation, in particular with respect to STEM related subdisciplines. The gender ratio is by far not as consistent as it seems at a first glance (c.f. Gisler et al., 2018). An analysis of the official statistics of first-year university students in the winter term of 2017/2018 for Germany, as a highly STEM-industrialized western country, showed the unequal proportion of women and men who are enrolled in different subdisciplines within STEM majors: While women are in fact overrepresented in subjects such as biology (62.0%), pharmacy (68.6%), or architecture (57.9%), they are decidedly underrepresented in subjects like physics (28.7%), engineering (22.3%), or computer sciences (21.1%) (German Federal Bureau of Statistics, 2018).

Not only recently, there is an international call towards researchers and policy-makers to focus on developing interventions to encourage women to engage in STEM majors in universities and to follow careers in these domains (c.f. Liben and Coyle, 2014). It is important to mention that the debate also emphasizes gender differences in abilities as an explanation for gender gaps in STEM fields. Gender differences in domain-specific competencies such as mathematics or reading are consistently shown in empirical research (see PISA 2015; OECD, 2016). Furthermore, boys are on average better at spatial tasks than girls, who are in turn better at verbal tasks (for a review, see Spelke, 2005). However, most studies only find small differences and the current discussion focuses more on the proposition of the gender similarities hypothesis (Hyde, 2005, 2014).

The most recent PISA study in 2015 placed a particular focus on the science competence of 15-year-olds in Germany and investigated whether students in adolescence already show aspirations towards STEM occupations (for Germany: Schiepe-Tiska et al., 2016b). The findings revealed for the example of Germany that 27% of boys, yet, only 18% of girls indicated that they would consider pursuing a career in science at the age of 30. The study further confirmed that there are differences with respect to certain subdisciplines. Boys are more interested in mathematics and information technology, whereas girls are more interested in health related occupations within STEM fields (Schiepe-Tiska et al., 2016a). The results also confirmed that science competence is rather unimportant in predicting adolescents' aspiration to pursue a STEM career, but instead their instrumental motivation and their enjoyment of science tasks were positively related to their aspiration toward STEM (Schiepe-Tiska et al., 2016a,b).

Thus, in the recent scientific discussion, gender differences in domain-specific competencies are not considered sufficient in fully explaining differences in choices for STEM careers, and researchers focus more often on affective-motivational and other non-cognitive explaining factors.

Psychological Mechanisms for Educational and Occupational Choices Into Science Technology Engineering Mathematics Fields

In looking more closely into factors that might be able to explain differences in choosing a STEM vs. non-STEM occupation or study course, psychological research provides a variety of explanations with respect to educational decisions and motivation as well as stereotypes and images of STEM subjects.

Gender-specific educational decisions and motivation are often displayed in the wide known metaphor of the "pipeline issue" (Clark Blickenstaff, 2005; Maltese and Tai, 2011; Cannady et al., 2014). This line of discussion describes the phenomenon that fewer girls than boys choose to study STEM subjects already in secondary school which then again leads to less women who decide to study STEM subjects in university or to work in STEM-related occupations (c.f. for Germany: German Federal Bureau of Statistics, 2018). Moreover, there is also the issue of the "leaky pipeline" (Alper, 1993) describing that, in addition to the lower proportion of women who start a STEM career, they were also more likely to drop-out during the course of an education within a STEM field. The pipeline-metaphor is criticized not only for suggesting a linear path within a STEM career and for neglecting the role of gatekeepers in this process but also for providing a seemingly easy fix for policy-makers (Cannady et al., 2014). Also, a report by the Committee on Barriers and Opportunities in Completing 2-year and 4-year STEM degrees appointed by the National Academies of Sciences, Engineering, and Medicine (2016) points out that there is a variety of paths within STEM careers and that it is not advisable to propose only a linear route. The report also states that this inaccurate image of STEM careers is a reason why most efforts to intervene fail because it does not acknowledge the more complex ways and challenges that students face within their STEM education. In that same respect, previous research also argues that explanations for the high attrition rate in STEM majors in general are a lack or loss of interest in STEM subjects, poor teaching of STEM faculty, or inadequate advising and help with academic difficulties (e.g., Seymour, 1992). Therefore, without the intention to simplify the obstacles for women to choose a STEM career, from our point of view, the metaphor still quite well symbolizes the smaller probability to enter a STEM field as well as the larger drop-out for women in STEM careers; however, it does not provide an extensive image, especially with the aim to derive intervention programs.

In fact, the reasons for these drop-outs are manifold: Studies, for example, confirmed external factors such as a discrimination against women in hiring processes within STEM fields (e.g., Moss-Racusin et al., 2012). It is, however, also often discussed that there are internal factors which lead women to opt out of a STEM career, for example, the difficult compatibility of different roles with respect to a reasonable work-family-balance within those fields (e.g., Blair-Loy, 2003; Diekman et al., 2010). Seymour (1992) reported that more than half of the students who switched from STEM majors to non-STEM majors indicated the rejection of STEM careers or the associated lifestyle, respectively, as a concern that contributed to their decision. As a consequence, there is a small number of women who work in or study STEM fields in tertiary education in Germany (German Federal Bureau of Statistics, 2018). One prominent theory that is often consulted is the expectancy-value theory (c.f. Wigfield and Eccles, 1992) which describes motivational and self-evaluative aspects of career decisions that are influenced by individual characteristics, but also formed by socialization, and expectations from teachers (e.g., Beilock et al., 2010; Upadyaya and Eccles, 2014) or parents (e.g., Räty et al., 2002; Tenenbaum and Leaper, 2003; Lindberg et al., 2008). Additionally, according to this theory, the value that is assigned to certain subjects or fields of occupation plays an important role for educational decisions. For example, Lent et al. (1994) argued in the social-cognitive career theory that individuals' self-efficacy or expected outcomes are relevant for the development of occupational interests which in turn are related to occupational success, aspirations, and decisions (Lent et al., 1994, 2010). Lent et al. (2018) conducted a meta-analysis combining the findings of 143 studies and showed that perceived support and perceived barriers are in general relevant for individuals' choice goals, but also that in particular for women perceived barriers are relevant for their outcome expectations (Lent et al., 2018).

Another theoretical angle to describe the phenomenon of underrepresentation of women in STEM fields is related to the stereotypes and images that are associated with these domains. A lot of empirical studies have confirmed the gender-science and gender-math stereotypes in implicit association tests according to which mathematics and sciences are perceived as male-stereotyped domains (e.g., Nosek et al., 2009; Plante et al., 2009; Steffens et al., 2010; Cvencek et al., 2011; Steffens and Jelenec, 2011; Passolunghi et al., 2014). Moreover, empirical evidence shows that those stereotypes also predict gender differences in science and mathematical achievements (e.g., Nosek et al., 2009). Additionally, alternative studies focused on the perceived images of STEM subjects and the self-to-prototype matching strategy (c.f. Niedenthal et al., 1985). Previous research shows, for example, that students who indicated physics as their favorite subject are perceived as intelligent, but at the same time as unpopular and unattractive (Hannover and Kessels, 2004).

Goal Congruity Approach

According to the goal congruity theory, individuals strive to live in congruence to their goals and to the perceived expectations of their environment; therefore, an individual's communal goal orientation might discourage them from pursuing a STEM career (c.f. Diekman and Steinberg, 2013; Diekman et al., 2015). Early on, Bakan (1966) proposed that there are agentic and communal motivations, and these motivations were stated as relevant to social judgment and self-concepts (c.f. warmth-competence distinction by Fiske et al., 2007). Even though communal traits are valued in men and women, gender norms particularly associate them to women (e.g., Diekman and Goodfriend, 2006). Against the background of previous research on gender stereotypes (i.e., the association of attributes, traits, tasks to either gender group, c.f. Eagly, 1987; Eagly and Wood, 2016), women's and men's meaning of work or their professional goal orientation should differ according to these stereotypes. Gender stereotypes are defined as how women and men are perceived and what is expected of them (i.e., descriptive and prescriptive component of stereotypes, Eagly, 1987). Research showed that gender stereotypes are categorized along the dimensions of "people-things" (Su et al., 2009) or "communal-agentic traits" (e.g., Abele, 2003; Abele and Wojciszke, 2007; Fiske et al., 2007). These stereotypes describe that women are more likely associated to activities (including occupations) related to interactions with people, caring, and taking social responsibilities (e.g., as a communal role). Men, however, are stereotypically associated to handling "things," securing the financial situation of the family (e.g., breadwinner), and being in charge (e.g., as an agentic role) (e.g., Eagly and Wood, 2016). Therefore, individuals' occupational goals and their meaning of work is expected to differ for women and men: Women are expected to be oriented toward social aspects and a satisfactory work-family balance, yet, the professional goal orientations of men should reflect more competitive, individualistic, and economic goals.

The current study focuses on a particular aspect of occupational goal orientation, namely individuals' meaning of work. Against the empirical and theoretical background, students' meaning of work (MOW) could be of particular interest in explaining the gender-gap in STEM fields. Individuals' meaning of work are described by the significance and value that is associated to work and occupations (c.f. Ruiz-Quintanilla, 1991; Claes and Ruiz-Quintanilla, 1993). We expect that students in STEM and non-STEM programs in university might be differentiated with respect to their goal orientation according to gender stereotypes. Students in STEM programs are aspiring to higher income and secure and prestigious jobs (Oh and Lewis, 2011), whereas the more heterogeneous group of students in non-STEM majors might be characterized by placing more importance to social goals or expressive aspects of occupations. Furthermore, women should in general strive for a comfortable work situation and flexible hours to balance their work with their family life. Men, however, should be aspiring more economic security according to the breadwinner model (c.f. social role theory, Eagly and Wood, 2016). We further expect that there is an interaction with students' gender: Women who decide to study a STEM subject should be less in line with female stereotypical goals such as care-taking and societal or social responsibilities than women who decide to choose a non-STEM major. Men, however, should probably show different goal orientations when they pursue a communal career (c.f. Croft et al., 2015) than men who pursue STEM subjects. Yet, since non-STEM occupations are very heterogeneous and not only consist of communal occupations, we do not expect large differences for men. Since STEM fields are also related to higher income and prestige (Oh and Lewis, 2011) students' socio-economic background should explain interindividual differences in students' professional goal orientations. To sum up, the current study aims at a comparison of students in STEM and non-STEM programs by examining differences in their professional goal orientation.

RESEARCH AIMS AND HYPOTHESES

Previous studies focused on the impact of competencies, motivation, or expectations on individuals' decisions to pursue a career in a STEM field. To add to the various aspects of explaining the underrepresentation of women in STEM, we argue that individuals' professional goal orientation (i.e., meaning of work) might also be related to educational and occupational decisions and tenacity. This should be true especially for women whose major subjects are within STEM fields and therefore more often associated to men. Therefore, our hypotheses were as follows:

(1a) Specific profiles of students' professional goal orientation (i.e., meaning of work) can be detected. Next to a profile with overall high and a profile with overall low goal orientations, we hypothesize two specific profiles that are described as either focused on social, and well-being aspects of working (c.f. communal role) or on economic and autonomy aspects of working (c.f. agentic role). We further hypothesize that (1b) women more likely belong to the profile of high social goal orientations, whereas men more likely belong to the profile of high economic goal orientations.

(2) Students who pursue a career in a STEM field as compared to students in non-STEM majors differ in their professional goal orientations. (2a) Students who are enrolled in non-STEM majors are more likely members of a profile with social and well-being aspects of working, whereas (2b) students who are enrolled in STEM fields more likely belong to the profile of economic or autonomy aspects of working. Furthermore, we expect (2c) an interaction of students' gender and their study major. Women in STEM majors should less likely belong to a profile of social goal orientations than women in non-STEM majors.

MATERIALS AND METHODS

Sample

In the current study, we did a secondary data analysis with data from a sample of $N = 13{,}113$ university students in their first to second academic year (i.e., wave 1; winter term 2010/2011; starting cohort (5) of the German National Educational Panel Study (NEPS; Blossfeld et al., 2011). The German National Educational Panel study provides longitudinal data from six representative starting cohorts within a multi-cohort sequence design (starting with a birth cohort and up to adulthood). In this study, we used data from the NEPS starting cohort five of university students in their first academic year. For registered researchers, the data are available as Scientific Use File and more information on the design, cohort information, and the measurements are documented on www.neps-data.de. We aimed at identifying first-year university students who were enrolled in their majors in either STEM or non-STEM fields. In order to distinctly identify the STEM-related subjects of the students, we excluded 4,252 students (32.4%) of the sample who studied toward a teaching degree (either bachelor degree or state examination for teaching degree) because most of them have a combination of two to three subjects which are not necessarily within the same categories with respect to STEM and non-STEM. However, this exclusion criterion was not based on any reasons implying less relevance of this particular group of students. We surely acknowledge the major role (prospective) teachers play in modeling (gender-typical) behavior, especially in science (c.f. Stout et al., 2011). We further excluded students who are enrolled in a university of applied sciences ($n = 2{,}967$; $n = 5$ students had missing values on the type of university, $n = 3$ students studied abroad or indicated to have no university). We argue that there is no coherent theoretical outline to include both types of institutions (i.e., university or university of applied sciences) for this research question since universities of applied sciences are more directly oriented toward the labor market and also show differences with respect to the provided study majors; for example, universities of applied sciences provide more courses in engineering than universities.

The sample consists of $n = 5{,}883$ students who are enrolled in a university in Germany ($n = 3$ students had missing values on their major program at university). We used the information about students' first major subject in university, yet, $n = 183$ students indicated to study in a bachelor program with two major subjects. From this subsample, we excluded $n = 29$ students because their first and second major subjects were not coherently both in either STEM or non-STEM fields. Students' minor subjects were not taken into account. The final sample for our analyses was $n = 5{,}857$. The students were asked about their goal orientation (i.e., meaning of work) in their second year in university (i.e., wave 3, summer term 2012).

Research Instruments

STEM and Non-STEM Program

First-year students' main study majors in university were categorized as STEM or non-STEM majors according to the categorization of major subjects in the winter term of 2010 of the German Federal Bureau of Statistics (2018) ($n = 3$ missing values on first major subject). This categorization subsumes STEM fields for all subjects in the area of mathematics and science (mathematics and science in general, mathematics, physics, astronomy, chemistry, pharmacy, biology, geological science, and geography) as well as engineering (engineering in general, mining industry/metallurgy, mechanical engineering/process engineering, electrical engineering/ information engineering, traffic engineering/nautical science, architecture/interior design, city and regional planning, construction engineering, surveying and mapping, industrial engineering, computer sciences, and materials sciences/materials engineering). We categorized the students' subjects accordingly. Students were more often enrolled in non-STEM majors ($n = 3{,}597$; 61.4%) than in STEM majors ($n = 2{,}257$; 38.6%) which is comparable to the proportion of students in STEM majors (38.0%) and non-STEM majors in Germany in the winter term 2010/2011 (reference: students in first semester at university; German Federal Bureau of Statistics, 2011, p. 34). Of the students enrolled in STEM majors 42.3% ($n = 954$) were enrolled in engineering, whereas 57.7% ($n = 1{,}303$) were enrolled in mathematics and science.

In reference to the group of students in non-STEM majors, the largest groups were students enrolled in law, economic, and social sciences (46.9%, n = 1,688) followed by students in language and cultural studies (29.5%, n = 1,061). The smaller groups were students enrolled in medicine and health-related sciences (15.3%, n = 552), agricultural, forestry, and nutritional sciences (2.5%, n = 90), arts (3.4%, n = 123), sports (1.3%, n = 47), and veterinary medicine (1.0%, n = 36). This distribution of first-year students is comparable to the expected proportion of students in the subgroup of non-STEM majors who were enrolled in winter term 2010/2011 at German universities (c.f. German Federal Bureau of Statistics, 2011; students only in non-STEM majors were enrolled as follows: 54.0% law, economic and social sciences; 28.0% language and cultural studies; 7.2% medicine and health related sciences; 3.3% agricultural, forestry and nutritional sciences; 5.6% arts; 1.5% sports; and 0.4% veterinary medicine).

Students' Professional Goal Orientations

Students' professional goal orientations (i.e., meaning of work) (c.f. Ruiz-Quintanilla, 1991) describe the importance of goals and activities associated to occupations independent of the individual's current situation. These occupational goal orientations were measured on a six-point scale from "1 not important at all" to "6 very important". The theoretically expected six subscales of this questionnaire were (1) learning (e.g., "Opportunity to learn new things", two items), (2) social orientation (e.g., "A work that is useful for the society", three items), (3) autonomy (e.g., "Own decision making competence", two items), (4) economic aspects (e.g., "Good chances to move up the career ladder", three items), (5) comfort aspects (e.g., "Pleasant working hours", two items), (6) expressive aspects (e.g., "Diverse tasks", four items). Before including these dimensions into the latent profile analysis, however, we checked the factor structure in a confirmatory factor analysis. The model fit for the original factor model was not satisfactory with χ^2 = 4252.40, df = 89, p < 0.001, CFI = 0.80; TLI = 0.73, RMSEA = 0.089 and hinted to a problem in the dimension of expressive aspects of goal orientations. Therefore, we conducted an exploratory factor analysis to check the empirical validity of these dimensions.

Student's Economic Situation

Students' economic situation was measured *via* a question regarding the perceived difficulties to provide things or to pay fees for the study course ("How hard is it for you and your family to pay for the things that you need for your academic studies, for example, travel expenses, books, or tuition fees?"). This question was measured on a five-point scale from "1, very difficult" to "5, very easy" and afterwards recoded so that higher values indicated a higher financial hardship of the according student. Students indicated on average a medium burden due to financial issues, M = 2.51, SD = 0.97, range 1–5, missing values n = 14.

Analysis Plan

In order to test the hypotheses, an exploratory factor analysis as well as a latent profile analysis were conducted using Mplus Version 8 (Muthén and Muthén, 2017). The preparation of the data set and some of the preliminary descriptive analyses were conducted in SPSS. First, the exploratory factor analysis with the scale goal orientation was conducted to empirically test the dimensions of this construct. Second, the latent profile analysis was conducted with a comparison of latent profile solutions from two to five profiles. For the latent profile analysis, the latent profile indicators were the factors of students' professional goal orientations. To test the profile specific hypotheses, the automatic three-step method implemented by Mplus through the R3STEP command was used (c.f. Asparouhov and Muthén, 2014). Afterwards, the thereby established profile memberships were fixed and used in a multinomial logistic regression as dependent variables with auxiliary variables (i.e., the predictor variables). Independent variables in this analysis were students' gender (male −0.5; female 0.5) and whether students were enrolled in a STEM major or a non-STEM major (non-STEM −0.5; STEM 0.5) as well as the interaction of students' gender and the STEM vs. non-STEM programs. Furthermore, the control variable students' economic situation (grand-mean centered) as well as the interaction of economic situation and STEM major was included in the analysis. The Mplus syntax and model outputs for the exploratory factor analysis and the latent profile analyses are available under: https://osf.io/k86ny/.

Cases with missing values on either dependent or independent variables were excluded from the multinomial logistic regression analysis but included in the latent profile analysis. However, there were only n = 17 missing cases (<1%) on the predictor variables; therefore, listwise deletion should not lead to biased results in this analysis.

RESULTS

Preliminary Analyses

Sample Descriptives

In the first step, we compared the two groups of students with respect to the gender distribution and their economic situation. As expected, 74.3% of women were enrolled in a non-STEM major as compared to STEM major subjects, whereas 54.6% of men were enrolled in a STEM major as compared to non-STEM programs. The unequal proportion of women and men in STEM majors was statistically significant, χ^2 = 512.71, df = 2, p < 0.001.

Furthermore, student's economic situation differed between the two groups, however, with only a small effect size: Students in STEM majors reported slightly less burden (M = 2.46, SD = 0.95) than students who were enrolled in non-STEM majors (M = 2.55, SD = 0.98), $t_{(5838)}$ = 3.42, p = 0.001, d = 0.09.

Exploratory Factor Analyses

In the second step, comparative exploratory factor analyses were conducted to test the subdimensions of students' occupational goal orientation. The model fits for the factor solutions of a unidimensional up to a six-factor model were compared by the conventions described for example in Hu and Bentler (1999). The comparison of model fit information (as displayed in **Table 1**) suggested a five- or six-factor solution for this scale. Yet, the subdimensions between the five- and six-factor models only differed slightly and did not add meaningful information to the goal orientations because none of the items loaded substantially better on the sixth factor. Therefore, the five-factor model with RMSEA = 0.044; CFI = 0.973, and TLI = 0.936 was chosen for the following analysis.

The five dimensions are labeled as oriented toward (1) social, (2) psychosocial health, (3) economic, (4) autonomy, and (5) motivational goal orientations. The *social factor* (measured by three items; Cronbach's alpha = 0.71) is described by the importance of useful work, of helping others, and doing something meaningful in work. The *psychosocial health factor* (measured by five items; Cronbach's alpha = 0.59) comprises a high significance of a good workplace climate, good working hours, and physical working conditions as well as secure employment, and a high match of skills and demands in the workplace. The *economic factor* (measured by two items; Cronbach's alpha = 0.69) combines the relevance of a good payment and opportunities for advancements. The *autonomy factor* (measured by three items; Cronbach's alpha = 0.72) is described by the importance of being independent, of having the authority to decide, and of being in charge. The *motivational factor* (measured by three items; Cronbach's alpha = 0.65) includes the importance of learning new things, facing manifold tasks, and having interesting work.

Descriptive Analyses of Students' Goal Orientations

The means and standard deviations of the factors of the construct goal orientation (i.e., meaning of work) with the five-factor solution are displayed in **Table 2** for the overall sample and separate for students in STEM and in non-STEM majors.

In general, students in non-STEM majors show higher goal orientations across almost all dimensions, except economic goals which were higher for students enrolled in STEM majors.

TABLE 1 | Model comparison of one- to six-factor solution (exploratory factor analyses) for students' goal orientations.

	AIC	BIC	RMSEA	CFI	TLI
1-factor	244490.23	244810.65	0.123	0.559	0.491
2-factors	240593.58	241014.13	0.101	0.746	0.658
3-factors	237731.97	238245.98	0.074	0.884	0.814
4-factors	236556.24	237157.02	0.058	0.941	0.886
5-factors	235891.88	236572.77	0.044	0.973	0.936
6-factors	235671.32	236425.64	0.038	0.984	0.952

The largest difference between students in non-STEM and STEM majors existed in the dimension of social goal orientation [$t_{(5852)} = 11.80$, $p < 0.001$, $d = -0.32$], followed by autonomy goal orientation [$t_{(5852)} = 7.67$, $p < 0.001$, $d = -0.21$], motivational goal orientation [$t_{(5852)} = 4.97$, $p < 0.001$, $d = -0.13$], and rather small differences in students' economic goal orientation [$t_{(5852)} = 3.53$, $p < 0.001$, $d = 0.09$], and the psychosocial health factor of the goal orientations [$t_{(5852)} = 2.59$, $p = 0.010$, $d = -0.07$]. The intercorrelations of the subdimensions of the construct are displayed in **Table 3**.

All subdimensions of the five-factor model of students' goal orientation are moderately correlated with each other (between $r = 0.13$ for economic and motivational factor to $r = 0.40$ for social and motivational factor); additionally, there is a zero correlation between the social and the economic factor of students' goal orientation ($r = 0.01$).

Results of Latent Profile Analysis

First, simple latent profile analyses with profiles differing from three to six were conducted to determine if the presumed four-profile solution was acceptable. According to the comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) (e.g., Nylund et al., 2007) and considering the Vuong-Lo-Mendell-Rubin likelihood ratio test (Vuong, 1989; Lo et al., 2001), the three-profile model had the overall best acceptable fit (AIC = 61308.17, BIC = 61455.03, VLMR = 770.32, $p = 0.015$) compared to a two-profile model (AIC = 62066.49, BIC = 62173.30, VLMR = 2763.74, $p < 0.001$), to a four-profile model (AIC = 60753.40, BIC = 60940.31, VLMR = 566.78, $p = 0.209$), and to a five-profile model (AIC = 60412.26, BIC = 60639.22, VLMR = 353.14, $p = 0.350$). The Bayesian information criterion (BIC), the Akaike information criterion (AIC), the Vuong-Lo-Mendell-Rubin likelihood ratio test value, and the entropy of each latent profile is displayed in **Table 4**.

The substantial decrease in AIC and BIC as well as a slightly better entropy (albeit it still points to less distinguishable profiles) results in accepting the three-profile solution. Even though the four-profile solution shows slightly better fit indices with respect to AIC and BIC, the sizes of the profiles become very small (one profile consists of $n = 184$ individuals; 4.8% of the sample). The results of the three-profile latent model are displayed in **Figure 1**.

The findings of the latent profile analysis resulted in one general profile of students' goal orientations and two more or less specific goal orientation profiles, partially confirming our hypothesis 1a. The general profile of goal orientation is characterized

TABLE 2 | Descriptive analyses of subdimensions of students' goal orientations.

Occupational goal orientation	sample (*N* = 5,857)		Non-STEM (*n* = 3,597)		STEM (*n* = 2,257)		Mean difference (df = 5,852)		
	M	SD	*M*	SD	*M*	SD	*t*	*p*	*d*
Social (three items)	4.78	0.80	4.88	0.80	4.63	0.79	11.80	<0.001	–0.32
Psychosocial health (five items)	4.69	0.63	4.70	0.63	4.66	0.64	2.59	0.010	–0.07
Economic (two items)	4.68	0.87	4.65	0.88	4.72	0.85	3.53	<0.001	0.09
Autonomy (three items)	4.29	0.80	4.36	0.80	4.19	0.80	7.67	<0.001	–0.21
Motivational (three items)	5.20	0.59	5.24	0.59	5.16	0.59	4.97	<0.001	–0.13

by a profile with high overall goal orientation ("high goal orientation profile"; n = 3,144 students, 53.7% of the sample). The findings did not confirm a profile with relatively low goal orientation since students showed in general rather high goal orientations in all aspects of occupational goals. Furthermore, students in all profiles barely differed in their motivational goal orientation. The specific goal orientation profiles included one "social goal orientation profile" (n = 1,611 students, 27.5% of the sample) with relatively high importance toward a meaningful job for society and helping others as well as by trend toward the security of employment and pleasant working conditions in the workplace. The second specific goal orientation profile was characterized by high importance to economic goals as compared to the other goals and was therefore labeled as the "economic goal orientation profile" (n = 1,102 students, 18.8% of the sample).

The percentages of female and male students as well as students in non-STEM and STEM program within the three profiles of goal orientation are displayed in **Table 5**.

TABLE 3 | Intercorrelation of subdimensions of students' goal orientations.

	1	2	3	4	5
Social (1)	1.00	0.32	0.01	0.25	0.40
Psychosocial health (2)		1.00	0.38	0.28	0.28
Economic (3)			1.00	0.36	0.13
Autonomy (4)				1.00	0.37
Motivational (5)					1.00

TABLE 4 | Model fits of two- to five-profiles solutions from latent profile analyses of students' goal orientations.

	AIC	BIC	VLMR	Entropy
2-profiles	62066.49	62173.30	$p < 0.001$	0.588
3-profiles	61308.17	61455.03	$p = 0.015$	0.613
4-profiles	60753.40	60940.31	$p = 0.209$	0.654
5-profiles	60412.26	60639.22	$p = 0.350$	0.657

VLMR, Vuong-Lo-Mendell-Rubin likelihood ratio test.

Results of Multinomial Logistic Regression Analysis

The results of the multinomial logistic regression model are displayed in **Table 6**. As expected in hypothesis 1b, women were more likely than men in the social goal orientation profile (b = 0.85, SE = 0.12, $p < 0.001$, OR = 2.33) and in the high goal orientation profile (b = 0.91, SE = 0.08, $p < 0.001$, OR = 2.48) in reference to the economic goal orientation profile. As a consequence, men were more likely in the economic goal orientation profile relative to the social goal orientation profile and relative to the high goal orientation profile. Students who are enrolled in a STEM program were more likely in the economic goal orientation profile than in the social goal orientation profile (b = 0.54, SE = 0.19, p = 0.004, OR = 1.72) and less likely in the high goal orientations (b = −0.42, SE = 0.13, p = 0.001; OR = 0.66), confirming our hypothesis 2b. That means that students who are enrolled in non-STEM programs were more likely in the social goal orientation profile and in the high goal orientation profile, confirming our hypothesis 2a. Students in STEM or non-STEM programs had equal probabilities to belong to the social goal orientation profile in reference to the high goal orientation profile (b = 0.12, SE = 0.14, p = 0.396, OR = 1.13). Overall, contrary to our hypothesis, there were no significant interactions of students' gender and them being enrolled in a STEM field, hypothesis 2c was therefore not confirmed. Concerning the control variable, there were no significant effects of student's economic situation with respect

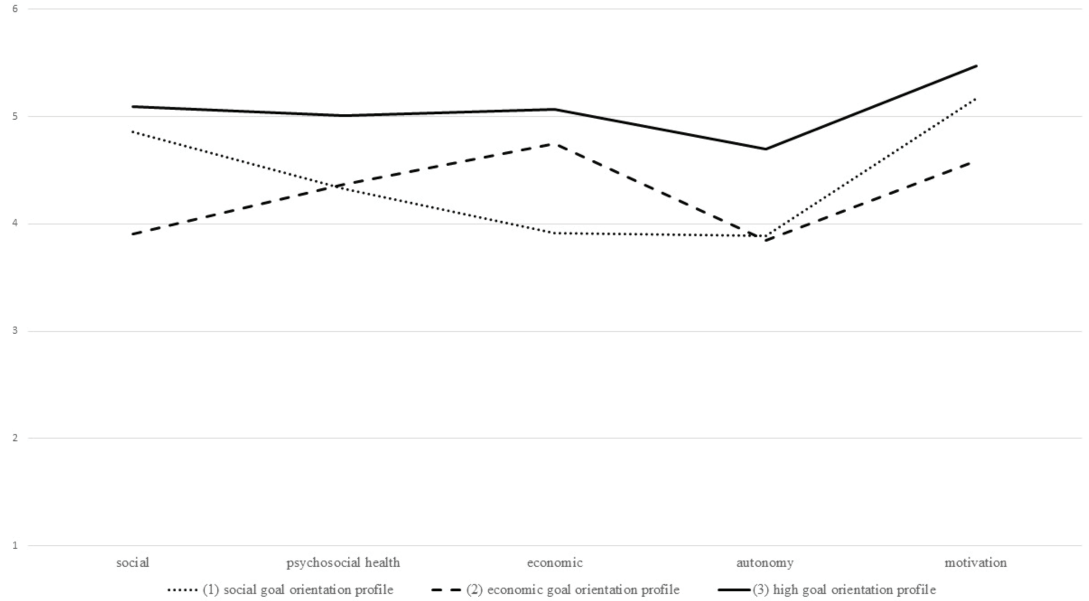


FIGURE 1 | Latent profile analysis of students' goal orientation: three-profile solution.

TABLE 5 | Frequencies (in percent) for the allocated profiles of students' goal orientation separate for female and male students as well as for students in non-STEM and STEM programs.

	Female students	Male students	Non-STEM program	STEM program
Social goal orientation profile	57.1	42.9	63.7	36.3
Economic goal orientation profile	40.7	59.3	51.6	48.4
High goal orientation profile	60.0	40.0	63.7	36.3

TABLE 6 | Results of multinomial logistic regression analysis.

Ref. Profile			Estimate	SE	p	OR
Social						
	Economic	Intercept	−0.29			
		Economic situation	0.03	0.07	0.636	1.03
		STEM	0.54	0.19	0.004	1.72
		Gender	−0.85	0.12	<0.001	0.43
		Economic × STEM	0.10	0.14	0.481	1.11
		Gender × STEM	−0.02	0.25	0.922	0.98
	High	Intercept	0.59			
		Economic situation	0.07	0.05	0.160	1.07
		STEM	0.12	0.14	0.396	1.13
		Gender	0.06	0.09	0.501	1.06
		Economic × STEM	0.02	0.11	0.840	1.02
		Gender × STEM	−0.21	0.17	0.223	0.81
Economic						
	High	Intercept	0.88			
		Economic situation	0.04	0.04	0.377	1.04
		STEM	−0.42	0.13	0.001	0.66
		Gender	0.91	0.08	<0.001	2.48
		Economic × STEM	−0.08	0.11	0.461	0.92
		Gender × STEM	−0.23	0.18	0.208	0.79

Gender × STEM, interaction term between STEM and gender; Economic × STEM, interaction term between STEM and economic situation; STEM, lower value non-STEM; Gender, lower value male; OR, odds ratio.

to the probability for either profile of goal orientation (e.g., social vs. economic profile of goal orientation: $b = 0.03$, SE = 0.07, $p = 0.636$, OR = 1.03). Furthermore, the interaction of students' financial situation and STEM major was also not relevant for their profile of goal orientation (e.g., social vs. economic profile of goal orientation: $b = 0.10$, SE = 0.14, $p = 0.481$, OR = 1.11).

DISCUSSION

Against the background of women's underrepresentation in STEM fields (e.g., German federal employment office, 2018), the aim of this study was to investigate latent profiles of students' occupational goal orientations (i.e., their meaning of work) in their second year in university. Furthermore, we aimed at comparing the goal orientations of two groups of students who were enrolled either in a STEM or in a non-STEM study program in higher education. The results first confirmed that there were differentiated profiles of students' professional goal orientations. In more detail, there were three profiles of students' goal orientations along the five dimensions of goal orientations labeled (1) social, (2) psychosocial health, (3) economic, (4) autonomy, and (5) motivational factors. First, there was one profile that was characterized by very high occupational goal orientation in general. Additionally, two profiles that are more specific were confirmed. One profile was characterized by a relatively high orientation toward social goals and well-being aspects combined with low orientation toward economic goals. The other specific profile was characterized by a relatively high orientation toward economic goals combined with rather low orientations toward social goals. In general, students (and therefore profiles) did not differ much in their motivational goal orientations. Consequently, the entropy of our profile solution was not very high, because at least this dimension, but also by trend the dimensions psychosocial health, and autonomy did not sufficiently differentiate between students' profiles. Overall, it is likely that there was less variation in the motivational factor because students are in general highly motivated to learn new things and to challenge themselves with diverse tasks when entering higher (tertiary) education.

As expected, gender does make a difference: Women were on average more than two times as likely in the social or high goal orientation profile than in the economic goal orientation profile, whereas men were more likely members of the economic goal orientation profile. This is in line with previous research showing that women are more associated with communal roles and taking care of others (e.g., Abele, 2003; Eagly and Wood, 2016). These aspects are reflected in the social dimension of the goal orientations, which include doing a meaningful work, and helping others. Men, however, are more associated with agentic roles (Abele and Wojciszke, 2007; Fiske et al., 2007) which are linked to being in charge, being the family breadwinner (c.f. social role theory, Eagly, 1987), and being competitive. This agentic role—but especially the breadwinner role—is much more reflected in the economic goal orientation profile, even though the autonomy dimension only shows a slightly higher magnitude. It is plausible to assume that entering university is in general associated with high independence and autonomy and that this is why we did not find meaningful differences in this dimension.

However, contrary to our hypotheses, we did not find an interaction of student's gender and STEM affiliation. Women in STEM did not differ in their membership to either of the profiles from women in non-STEM fields. The same applies to men: Men in STEM fields showed equal probabilities for the latent profiles as men in non-STEM fields. This is particularly important with respect to recent efforts and campaigns to get women involved in STEM fields (c.f. Liben and Coyle, 2014; Diekman et al., 2015). It seems from our study that women show in general a relatively higher orientation toward communal goals (e.g., Abele, 2003) which is in line with research on gender stereotypes (e.g., Fiske et al., 2007). Women in STEM programs, however, do not show a different pattern with respect to their goal orientations than their peers in non-STEM study programs according to our study. One explanation might be, however, that the results are not definitive and clear to interpret with respect to our hypothesis because many subjects are included in the broader area of STEM fields that are considered female gender-typed, such as biology or health-related subjects (c.f. Schiepe-Tiska et al., 2016a,b). Moreover, since we measured students' goal orientations in second year in university, it might be possible that students already altered or adapted their goals to perceived expectations along with their academic studies. Future research should try to disentangle those potential effects with a longitudinal design and an alternative categorization of major subjects in university.

Yet, as a consequence from our study, intervention plans should perhaps focus on a better fit of social and communal goals with a STEM career to further increase the proportion of women entering and persisting in those fields. The current scientific discussion reflects on many approaches to helping women to persist and to succeed in STEM programs (c.f. van den Hurk et al., 2019; for a review). Recent efforts of universities and other stakeholders are focusing, for example, on mentoring programs for women in STEM fields and further on providing positive role models (e.g., Drury et al., 2011). These approaches might not only increase individual's sense of belonging and feeling welcomed at an institution (e.g., Dasgupta, 2011; Ramsey et al., 2013) but also provide important information networks. Consequently, those approaches would pick up ideas to increase the support for students in STEM majors to provide help with academic difficulties and give constructive advice since these are often reported concerns, especially by students who switched from STEM to non-STEM majors (e.g., Seymour, 1992).

Furthermore, students economic situation (i.e., in this study: difficulties to pay for all expenses) was not relevant for the probability of profile membership, even though it was expected to be an indicator of further reasons to pursue a STEM career or university major, respectively, since the salaries in STEM occupations are on average relatively high (e.g., Oh and Lewis, 2011). Moreover, in this study, results showed that students in STEM majors were almost two times as likely in the profile of economic goals as compared to the profile of social goal orientation. However, our findings do not confirm the relevance of an interaction of STEM and economic situation for students' goal orientations and further research is needed with different indicators of socio-economic background and more information on the financial situation of the students.

Our study is limited to a cross-sectional analysis: Even though the students were examined at two time points (first and second year in university), we do not have data with repeated measures of their goal orientation. It would greatly increase the interpretation of our findings if there were longitudinal measures of students' occupational goal orientations. This would enable researchers to analyze not only the initial goal orientation—maybe even before entering tertiary education—but also to provide the analyses to examine if and how the goal orientations change over the educational years in university. It is plausible to assume that there are not only selection processes but also socialization processes that are relevant for university students' professional goal orientations. Overall, the differences between the groups in our study were not as pronounced as we expected. Gender differences and differences of students in STEM or non-STEM majors were similar and more pronounced in the comparison of economic and social goal orientation profiles. Since previous research showed a more differentiated pattern of gender differences in subdisciplines within the area of STEM subjects (c.f. German Federal Bureau of Statistics, 2018; Gisler et al., 2018), future research might also focus on differentiating STEM fields in more detail.

In conclusion, our findings did not confirm differences in the professional goal orientations between women (and men) in STEM and non-STEM majors in tertiary education. However, women were more oriented toward social aspects of occupations, whereas men were more oriented toward economic aspects of occupations. Furthermore, students enrolled in STEM majors allocated more importance to economic goals than social goals. Intervention plans to increase the proportion of women in STEM fields in tertiary education and the labor market should, according to our findings, probably focus more on the congruity of students' goal orientations with future career prospects of university degrees in STEM fields. Previous research showed that while boys are overall more interested in mathematics and information technology, girls are in fact interested in health-related occupations within STEM fields (Schiepe-Tiska, et al., 2016a,b). One potential approach might perhaps be to highlight the manifold occupations that a major in a STEM fields opens up.

ETHICS STATEMENT

The analyses of this paper are a secondary analyses of data published previously (Blossfeld et al., 2011). Data sources used for the analyses were the cohort of first-year students (doi: 10.5157/NEPS:SC5:11.0.0) of the German National Educational Panel Study (Blossfeld et al., 2011). All students from this cohort gave informed consent to participate in the panel by providing their phone number for being contacted for telephone interviews after being informed about the purposes of the study. Specific information about the recruitment process can be found in the field report of the study (Steinwede and Aust, 2012). All data analyses were performed via a download access at LIfBi in Bamberg, Germany that provided a controlled privacy environment for data access. Furthermore, an ethics approval for the analyses was obtained by the local ethics committee.

AUTHOR'S NOTE

This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort First-Year Students, doi: 10.5157/NEPS:SC5:11.0.0. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS

is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg in cooperation with a nationwide network.

AUTHOR CONTRIBUTIONS

IW performed the statistical analyses and drafted the manuscript. LE, TS and BD contributed to the conception of this study and discussed the results of the analyses and contributed to the manuscript and the revisions. All listed authors read and approved the submitted manuscript.

FUNDING

The publication of this article was funded by the Open Access Fund of the Leibniz Association.

REFERENCES

Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: findings from a prospective study. *J. Pers. Soc. Psychol.* 85, 768–776. doi: 10.1037/0022-3514.85.4.768

Abele, A. E., and Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *J. Pers. Soc. Psychol.* 93, 751–763. doi: 10.1037/0022-3514.93.5.751

Alper, J. (1993). The pipeline is leaking women all the way along. *Science* 260, 409–411. doi: 10.1126/science.260.5106.409

Asparouhov, T., and Muthén, B. (2014). Auxiliary variables in mixture modeling: 3-step approaches using Mplus. *Struct. Equ. Model.* 21, 329–341. doi: 10.1080/10705511.2014.915181

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion*. Chicago: Rand McNally.

Bechmann, S., Dahms, V., Tschersich, N., Frei, M., Leber, U., and Schwengler, B. (2012). Fachkräfte und unbesetzte Stellen in einer alternden Gesellschaft. Problemlagen und betriebliche Reaktionen [Skilled workers and vacant jobs in an aging society. Problems and organizational reactions]. IAB Research Report, 13. Available at: http://doku.iab.de/forschungsbericht/2012/fb1312.pdf (Accessed August 4, 2019).

Beilock, S. L., Gunderson, E. A., Ramirez, G., and Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proc. Natl. Acad. Sci. USA* 107, 1860–1863. doi: 10.1073/pnas.0910967107

Blair-Loy, M. (2003). *Competing devotions: Career and family among women executives*. Cambridge: Harvard University Press.

Blossfeld H.-P., Roßbach H.-G., and von Maurice J. (eds.) (2011). *Education as a lifelong process: The German National Educational Panel Study (NEPS). Zeitschrift für Erziehungswissenschaft/special issue.* Vol. 14, Wiesbaden: VS Verlag für Sozialwissenschaften.

Cannady, M. A., Greenwald, E., and Harris, K. N. (2014). Problematizing the STEM pipeline metaphor: is the STEM pipeline metaphor serving our students and the STEM workforce? *Sci. Educ.* 98, 443–460. doi: 10.1002/sce.21108

Claes, R., and Ruiz-Quintanilla, S. A. (1993).Work meaning patterns in early career. CAHRS Working Paper Series. Paper 277. Available at: http://digitalcommons.ilr.cornell.edu/cahrswp/277 (Accessed August 4, 2019).

Clark Blickenstaff, J. (2005). Women and science careers: leaky pipeline or gender filter? *Gend. Educ.* 17, 369–386. doi: 10.1080/09540250500145072

Croft, A., Schmader, T., and Block, K. (2015). An underexamined inequality: cultural and psychological barriers to men's engagement with communal roles. *Personal. Soc. Psychol. Rev.* 19, 343–370. doi: 10.1177/1088868314564789

Cvencek, D., Meltzoff, A. N., and Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child Dev.* 82, 766–779. doi: 10.1111/j.1467-8624.2010.01529.x

Dasgupta, N. (2011). Ingroup experts and peers as social vaccines who inoculate the self-concept: the stereotype inoculation model. *Psychol. Inq.* 22, 231–246. doi: 10.1080/1047840X.2011.607313

Diekman, A. B., Brown, E. R., Johnston, A. M., and Clark, E. K. (2010). Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol. Sci.* 21, 1051–1057. doi: 10.1177/0956797610377342

Diekman, A. B., and Goodfriend, W. (2006). Rolling with the changes: a role congruity perspective on gender norms. *Psychol. Women Q.* 30, 369–383. doi: 10.1111/j.1471-6402.2006.00312.x

Diekman, A. B., and Steinberg, M. (2013). Navigating social roles in pursuit of important goals: a communal goal congruity account of STEM pursuits. *Soc. Personal. Compass* 7, 487–501. doi: 10.1111/spc3.12042

Diekman, A. B., Weisgram, E. S., and Belanger, A. L. (2015). New routes to recruiting and retaining women in STEM: policy implications of a communal goal congruity perspective. *Soc. Issues Policy Rev.* 9, 52–88. doi: 10.1111/sipr.12010

Drury, B. J., Siy, J. O., and Cheryan, S. (2011). When do female role models benefit women? The importance of differentiating recruitment from retention in STEM. *Psychol. Inq.* 22, 265–269. doi: 10.1080/1047840X.2011.620935

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, New Jersey: Lawrence Erlbaum.

Eagly, A. H., and Wood, W. (2016). "Social role theory of sex differences" in *The Wiley Blackwell encyclopedia of gender and sexuality studies*. eds. N. Naples, R. C. Hoogland, M. Wickramasinghe, and W. C. A. Wong (New Jersey: Wiley)

Fiske, S. T., Cuddy, A. J. C., and Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends Cogn. Sci.* 11, 77–83. doi: 10.1016/j.tics.2006.11.005

German Federal Bureau of Statistics (2011). Bildung und Kultur. Studierende an Hochschulen. [Education and culture. Students at Universities]. Available at: https://www.destatis.de/GPStatistik/servlets/MCRFileNodeServlet/DEHeft_derivate_00006845/2110410117004.pdf (Accessed August 4, 2019).

German Federal Bureau of Statistics (2018). Students enrolled in STEM courses. Official statistics about institutions of higher education. Available at: https://www.destatis.de/EN/Themes/Society-Environment/Education-Research-Culture/Institutions-Higher-Education/Tables/students-in-stem-courses.html (Accessed August 4, 2019).

German Federal Employment Office (2018). *MINT – Berufe [STEM - Occupations]. Statistik der Bundesagentur für Arbeit*. Berichte: Blickpunkt Arbeitsmarkt, Nürnberg.

Gisler, S., Kato, A., Lee, S., and Leung, D. (2018). One size does not fit all: gender inequity in STEM varies between subfields. *Ind. Organ. Psychol.* 11, 314–318. doi: 10.1017/iop.2018.21

Hannover, B., and Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learn. Instr.* 14, 51–67. doi: 10.1016/j.learninstruc.2003.10.002

Hausmann, A.-C., and Kleinert, C. (2014). Berufliche Segregation auf dem Arbeitsmarkt. Männer- und Frauendomänen kaum verändert. IAB Research Short Report, 9. Available at: http://doku.iab.de/kurzber/2014/kb0914.pdf (Accessed August 4, 2019).

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55.

Hyde, J. S. (2005). The gender similarities hypothesis. *Am. Psychol.* 60, 581–592. doi: 10.1037/0003-066X.60.6.581

Hyde, J. S. (2014). Gender similarities and differences. *Annu. Rev. Psychol.* 65, 373–398. doi: 10.1146/annurev-psych-010213-115057

Lent, R. W., Brown, S. D., and Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* 45, 79–122. doi: 10.1006/jvbe.1994.1027

Lent, R. W., Paixão, M. P., Silva, J., and Leitão, L. M. (2010). Predicting occupational interests and choice aspirations in Portuguese high school students: a test of social cognitive career theory. *J. Vocat. Behav.* 76, 244–251. doi: 10.1016/j.jvb.2009.10.001

Lent, R. W., Sheu, H.-B., Miller, M. J., Cusick, M. E., Penn, L. T., and Truong, N. N. (2018). Predictors of science, technology, engineering, and mathematics choice options: a meta-analytic path analysis of the social–cognitive choice model by gender and race/ethnicity. *J. Couns. Psychol.* 65, 17–35. doi: 10.1037/cou0000243

Leuze, K., and Strauß, S. (2009). Lohnungleichheiten zwischen Akademikerinnen und Akademikern: Der Einfluss von fachlicher Spezialisierung, frauendominierten Fächern und beruflicher segregation [wage inequality between male and female university graduates: the influence of occupational specialization, female-dominated subjects and occupational segregation]. *Z. Soziol.* 38, 262–281. doi: 10.1515/zfsoz-2009-0401

Liben, L. S., and Coyle, E. F. (2014). Developmental interventions to address the STEM gender gap: exploring intended and unintended consequences. *Adv. Child Dev. Behav.* 47, 77–115. doi: 10.1016/bs.acdb.2014.06.001

Lindberg, S. M., Hyde, J. S., and Hirsch, L. M. (2008). Gender and mother-child interactions during mathematics homework: the importance of individual differences. *Merrill-Palmer Q.* 54, 232–255. doi: 10.1353/mpq.2008.0017

Lo, Y., Mendell, N. R., and Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika* 88, 767–778. doi: 10.1093/biomet/88.3.767

Maltese, A. V., and Tai, R. H. (2011). Pipeline persistence: examining the association of educational experiences with earned degrees in STEM among US students. *Sci. Educ.* 95, 877–907. doi: 10.1002/sce.20441

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., and Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proc. Natl. Acad. Sci. USA* 109, 16474–16479. doi: 10.1073/pnas.1211286109

Muthén, L. K., and Muthén, B. O. (2017). *Mplus user's guide. 8th Edn*. Los Angeles, CA: Muthén & Muthén.

National Academies of Sciences, Engineering, and Medicine (2016). *Barriers and opportunities for 2-year and 4-year STEM degrees: systemic change to support students' diverse pathways*. Washington, DC: The National Academies Press.

Niedenthal, P. M., Cantor, N., and Kihlstrom, J. F. (1985). Prototype matching: a strategy for social decision making. *J. Pers. Soc. Psychol.* 48, 575–584. doi: 10.1037/0022-3514.48.3.575

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. USA* 106, 10593–10597. doi: 10.1073/pnas.0809921106

Nylund, K. L., Asparouhov, T., and Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct. Equ. Model.* 14, 535–569. doi: 10.1080/10705510701575396

Ochsenfeld, F. (2012). Gläserne Decke oder goldener Käfig: Scheitert der Aufstieg von frauen in erste Managementpositionen an betrieblicher Diskriminierung oder an familiären Pflichten? [glass ceiling or golden cage: is discrimination in the workplace or duties in the family preventing women from promotion to early management positions?]. *Kölner Z. Soziol. Sozialpsychol.* 64, 507–534. doi: 10.1007/s11577-012-0178-1

OECD (2016). *PISA 2015 results (volume I): Excellence and equity in education*. Paris: PISA, OECD Publishing.

Oh, S. S., and Lewis, G. B. (2011). Stemming inequality? Employment and pay of female and minority scientists and engineers. *Soc. Sci. J.* 48, 397–403. doi: 10.1016/j.soscij.2010.11.008

Passolunghi, M. C., Rueda Ferreira, T. I., and Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learn. Individ. Differ.* 34, 70–76. doi: 10.1016/j.lindif.2014.05.005

Plante, I., Theoret, M., and Favreau, O. E. (2009). Student gender stereotypes: contrasting the perceived maleness and femaleness of mathematics and language. *Educ. Psychol.* 29, 385–405. doi: 10.1080/01443410902971500

Ramsey, L. R., Betz, D. E., and Sekaquaptewa, D. (2013). The effects of an academic environment intervention on science identification among women in STEM. *Soc. Psychol. Educ.* 16, 377–397. doi: 10.1007/s11218-013-9218-6

Räty, H., Vänskä, J., Kasanen, K., and Kärkkäinen, R. (2002). Parents' explanations of their child's performance in mathematics and reading: a replication and extension of Yee and Eccles. *Sex Roles* 46, 121–128. doi: 10.1023/A:1016573627828

Ruiz-Quintanilla, S. A. (1991). Introduction: the meaning of work. *Eur. Work Org. Psychol.* 1, 81–89. doi: 10.1080/09602009108408514 (Accessed August 4, 2019).

Schiepe-Tiska, A., Rönnebeck, S., Schöps, K., Neumann, K., Schmittner, S., Parchmann, I., et al. (2016a). "Naturwissenschaftliche Kompetenz bei PISA 2015: Ergebnisse des internationalen Vergleichs mit einem modifizierten Testansatz" in *PISA 2015: Eine Studie zwischen Kontinuität und Innovation*. eds. K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme, and O. Köller (Münster: Waxmann), 45–98.

Schiepe-Tiska, A., Simm, I., and Schmidtner, S. (2016b). "Motivationale Orientierungen, Selbstbilder und Berufserwartungen in den Naturwissenschaften in PISA 2015" in *PISA 2015: Eine Studie zwischen Kontinuität und Innovation*. eds. K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme, and O. Köller (Münster: Waxmann), 99–132.

Seymour, E. (1992). "The problem iceberg" in science, mathematics, and engineering education: student explanations for high attrition rates. *J. Coll. Sci. Teach.* 21, 230–238.

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *Am. Psychol.* 60, 950–958. doi: 10.1037/0003-066X.60.9.950

Steffens, M. C., and Jelenec, P. (2011). Separating implicit gender stereotypes regarding math and language: implicit ability stereotypes are self-serving for boys and men, but not for girls and women. *Sex Roles* 64, 324–335. doi: 10.1007/s11199-010-9924-x

Steffens, M. C., Jelenec, P., and Noack, P. (2010). On the leaky math pipeline: comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *J. Educ. Psychol.* 102, 947–963. doi: 10.1037/a0019920

Steinwede, J., and Aust, F. (2012). *Methodenbericht. NEPS Startkohorte 5 – CATI-Haupterhebung Herbst 2010 B52*. Bonn: infas Institut für angewandte Sozialwissenschaften.

Stout, J. G., Dasgupta, N., Hunsinger, M., and McManus, M. A. (2011). STEMing the tide: using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *J. Pers. Soc. Psychol.* 100, 255–270. doi: 10.1037/a0021385

Su, R., Rounds, J., and Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychol. Bull.* 135, 859–884. doi: 10.1037/a0017364

Tenenbaum, H. R., and Leaper, C. (2003). Parent-child conversations about science: the socialization of gender inequities? *Dev. Psychol.* 39, 34–47. doi: 10.1037/0012-1649.39.1.34

Upadyaya, K., and Eccles, J. S. (2014). How do teachers' beliefs predict children's interest in math from kindergarten to sixth grade? *Merrill-Palmer Q.* 60, 403–430. doi: 10.13110/merrpalmquar1982.60.4.0403

van den Hurk, A., Meelissen, M., and van Langen, A. (2019). Interventions in education to prevent STEM pipeline leakage. *Int. J. Sci. Educ.* 41, 150–164. doi: 10.1080/09500693.2018.1540897

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57, 307–333. doi: 10.2307/1912557

Wigfield, A., and Eccles, J. (1992). The development of achievement task values: a theoretical analysis. *Dev. Rev.* 12, 265–310. doi: 10.1016/0273-2297(92)90011-P

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 24 September 2019
doi: 10.3389/fpsyg.2019.02146

The More Interest, the Less Effort Cost Perception and Effort Avoidance

Juyeon Song[1], *Sung-il Kim*[2]* *and Mimi Bong*[2]

[1] Department of Education, Korea National University of Education, Cheongju, South Korea, [2] Department of Education, Brain and Motivation Research Institute (bMRI), Korea University, Seoul, South Korea

The present study aims to investigate what factors determine students' engagement in mathematics. We examined the predictive relationships between interest, effort cost (i.e., the cost of making the effort), and three forms of academic engagement: persistence, cognitive engagement, and effort avoidance. In addition, we examined gender differences in these relationships. We recruited 546 8th and 9th graders for this study. Consistent with previous research, interest worked as a strong positive predictor of persistence and cognitive engagement, and it predicted effort avoidance negatively. Moreover, interest negatively predicted the perception of effort cost, which in turn positively predicted effort avoidance. Gender differences were found in the mean values of effort avoidance and in the prediction by interest of the perception of effort cost. Male students reported higher effort avoidance than female students, and the prediction by interest of the perception of effort cost was stronger among female students than among male students. These findings provide new insights into students' engagement in mathematics and the role of interest and effort cost in it.

Keywords: interest, effort cost, academic engagement, mathematics, gender difference

OPEN ACCESS

Edited by:
Rebecca Lazarides,
University of Potsdam, Germany

Reviewed by:
Mercedes Inda-Caro,
Universidad de Oviedo, Spain
Andrea Westphal,
University of Potsdam, Germany

****Correspondence:***
Sung-il Kim
sungkim@korea.ac.kr

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *13 January 2019*
Accepted: *05 September 2019*
Published: *24 September 2019*

Citation:
Song J, Kim S and Bong M (2019) The More Interest, the Less Effort Cost Perception and Effort Avoidance. Front. Psychol. 10:2146. doi: 10.3389/fpsyg.2019.02146

INTRODUCTION

Many researchers and educators have been interested in how to make students study mathematics deeply and persistently. Interest is one of the most representative motivators that facilitate engagement (Eccles, 2016). Although interest researchers have slightly different definitions of interest, empirical evidence in a wide variety of theoretical frameworks consistently suggests that interest has a role in promoting academic engagement (Hidi et al., 2004; Sansone and Thoman, 2005; Reeve et al., 2015).

Recent motivation researchers have been trying to identify maladaptive motivators that hinder students' academic engagement. Although previous research has identified trait procrastination and task difficulty as personality or environmental factors that hinder students' engagement, there is still insufficient understanding of motivational factors that can explain negative motivation in academic engagement. One potential factor is effort cost which is defined as the perception of effort required to study mathematics (Battle and Wigfield, 2003; Jiang et al., 2018). Research studying the role of both positive and negative motivators together is expected to deepen our understanding of psychological mechanisms of students' engagement in learning. Thus, the objective of this study was to examine the role of interest and effort cost together in predicting various forms of engagement.

Especially, gender differences in STEM (Science, Technology, Engineering, and Mathematics) areas have received worldwide attention. The previous studies have reported that boys are more likely than girls are to have positive motivation for math (Else-Quest et al., 2010). However, studies on gender differences in mathematics have focused primarily on the mean difference in math achievement and motivational variables, but few studies have examined whether male and female students have different or the same motivational paths in mathematics-related learning and decision-making. In addition, there is a report that the gender gap is decreasing even at the mean levels (Hyde and Linn, 2006; Gaspard et al., 2017). In this study, we thus aimed to examine whether gender differences still exist in the mean levels of math motivation and engagement and whether gender differences exist in the relationship between them.

THEORETICAL BACKGROUND

Relationship Between Interest and Engagement

Academic engagement is defined as an investment into knowledge or skills that can yield meaningful in-depth learning (Newmann et al., 1992; Mayer, 2002). Students need to use cognitive strategies and persistent effort in order to produce meaningful learning. In this sense, persistent effort and cognitive strategy are viewed as the most representative forms of academic engagement (Greene and Miller, 1996; Fredricks et al., 2004). By contrast, effort avoidance is a different form of engagement, in which students participate in learning minimally with little meaningful learning taking place (Song et al., 2017). An example of this is an attitude that avoids trying to understand difficult parts or solve difficult problems.

The most representative role of interest is to increase deep levels of cognitive strategy uses and persistence. Interest is a motivational construct derived from inherent enjoyment that people feel in the process of performing a task (Schiefele, 1991; Sansone and Thoman, 2005). Interest has two distinct features. First, it can lead students to become intrinsically motivated and internally regulate their behavior (Ryan and Deci, 2000; Isen and Reeve, 2005). Second, it has a strong connection to positive emotion generated by engaging in a task (Higgins, 2006). The relationship between interest and engagement is reciprocal. Interest can be situationally triggered by engaging in a specific task. This type of the interest is called *situational interest* (Krapp, 1999; Hidi and Renninger, 2006). In this case, engagement seems to precede interest. The major role of situational interest is to focus attention. Situational interest may not last over time or in other situations (Hidi and Baird, 1986; Hidi, 1995). However, *individual interest,* which is defined as a relatively stable personal interest related to a particular domain, task, or activity, can function as high level of motivation and lead to the use of higher-order cognitive strategies with persistent engagement and learning (Krapp, 1999; Hidi and Renninger, 2006). In the present study, individual interest is postulated as a positive predictor for persistence or cognitive engagement based on Eccles' Expectancy-Value Theory and Hidi and Renninger's Interest Theory (Hidi and Renninger, 2006; Eccles, 2016).

Researches on the relationship between interest and effort avoidance are insufficient compared to studies on the relationship between interest and persistence or cognitive engagement. However, considering the positive role of interest in engagement, it can be expected that interest is negatively related to effort avoidance. In addition, investigating the relationships of effort cost with interest and effort avoidance will enable us to explain the negative link between interest and effort avoidance.

Relationship Between Interest and Effort Cost

Effort consumes mental or physical energy. Thus, people tend to avoid participating in a task when it requires a large amount of efforts (Inzlicht et al., 2018). In this regard, effort is perceived as costly (Eccles et al., 1983). This negative perception of time, energy, or amount of work put into a task is named 'effort cost' (Eccles et al., 1983; Gaspard et al., 2015). In fact, effort and time are typically considered as primary costs when people make a decision (Botvinick et al., 2009; Vassena et al., 2014).

However, effort is not always perceived as costly. When the task is associated with feeling enjoyment, the effort may be no longer considered as costly (Inzlicht et al., 2018). This may due to two notable functions of interest: resource replenishment and effortless attention. First, Thoman et al. (2011) sought to explain why some people can engage in an interesting task even after their resources have been depleted; in a series of three studies, they discovered that interest has a resource replenishment function. In their research, participants were first depleted by a task (e.g., a Stroop task) and then asked to perform one of three emotionally stimulating tasks that evoked either interest, positive emotion, or neutral emotion. Following this, the participants then engaged in a subsequent, unrelated task for as long as they wanted to. Participants who had been given the interesting task persisted longer in the subsequent task than did those who had been asked to complete either the positive- or neutral-emotion task. Interestingly, this result was observed only when the participants' energy had already been depleted before performing the second, emotionally stimulating task. The authors interpreted these differences in persistence as being a consequence of the resource-replenishment function of interest. They also tried to elucidate the underlying mechanism behind the resource-replenishment function and thus tested positive emotion and increased competence as potential mechanisms but were not able to identify the mechanism in question.

Automatic or effortless attention is another important function of interest. For example, a more-interesting text requires less time to read them than a less-interesting one, and people who read more-interesting texts perform better on recall tests than do those who read less-interesting texts (McDaniel et al., 2000). Individuals can quickly and effortlessly focus their attention on a target task when its characteristics, such as novelty and relevance, provoke their interest (Pekrun, 1992; McDaniel et al., 2000; Hidi et al., 2004; Hidi and Ainley, 2008). In this process, the cognitive effort or resources required to concentrate on a task can be preserved by automatic engagement and action. Effortless or automatic attention accordingly enables participants to focus more on deeper cognitive engagement (McDaniel et al., 2000; Linnenbrink and Pintrich, 2004).

Both resource replenishment and effortless attention are strongly connected to the effort or energy that individuals invest in a certain task. According to the research summarized above, interest allows for the recovery of previously drained energy levels through its resource-replenishment function and reduces cognitive effort itself by automatically drawing the attention of the participants. These studies provide an interesting insight into the relationship between interest and cost, especially effort cost. Considering the two functions of interest, it could be expected that interest would be a negative predictor of the perceived effort cost required to complete a task. In other words, even if people do the same amount of work, if they are interested in the task, they may be less aware of the effort required for the task. In addition, efforts combined with interest can be considered even valuable, rather than costly (Inzlicht et al., 2018). Despite this, few studies have examined the relationship between interest and effort cost. Some recent research has reported negative correlations between interest and effort cost, but the relationship between interest and effort cost was not a focus of the studies (Gaspard et al., 2015; Jiang et al., 2018).

Both interest and effort cost are students' subjective perceptions rather than objective ones. Therefore, they can affect each other. For example, how much of a burden the effort required for a task feels like may depend on the degree of interest the individual has in the task, even if the task requires the same amount of effort. The opposite is also possible. One recent study has shown that task values, including interest, and costs can predict each other, although predictions differ depending on school years (Part et al., 2018), meaning that evidence for a causal relationship between interest and effort cost remains questionable. Therefore, in the present study, we first tried to explore whether there were students who had high interest and low effort costs, or if there were students who were both highly perceived by using a person-centered approach. Next, we sought to examine the role of interest in the perception of effort cost theoretically based on two functions of interest mentioned above. Specifically, interest in a task could lower the perception of effort cost because of the two functions of interest: resource replenishment and effortless attention (Hidi et al., 2004; Hidi and Ainley, 2008; Thoman et al., 2011).

Relationship Between Effort Cost and Engagement

The perception of cost is found to be related to the intention to engage in a task and the intention to quit (Eccles and Wigfield, 1995; Battle and Wigfield, 2003; Kurzban et al., 2013; Perez et al., 2014). Avoidance-related intentions (e.g., intent to drop out) and behaviors (e.g., disengagement and procrastination) have been particularly identified as unique consequences of task costs such as effort cost (Perez et al., 2014; Jiang et al., 2018).

Although there were other costs such as opportunity cost and psychological cost, Flake et al. (2015) found that effort cost was the most frequent cost-related response (i.e., 42%) when students were asked to describe the features of the class which motivated them the least. Similarly, Perez et al. (2014) reported that only effort cost significantly and consistently predicted the intent to leave by STEM majors over time, whereas beliefs about competence, task value, opportunity cost, and psychological cost did not. Neuroscience research has also shown that individuals tend to avoid a highly demanding task when they can choose the task (Croxson et al., 2009; Kool et al., 2010; McGuire and Botvinick, 2010). Therefore, among task costs, the present study especially focused on effort cost for two reasons: (1) the unique functions of interest could be the rationale for the link between interest and effort cost, and (2) given the previous findings, effort cost seemed to show better prediction than other costs in the explanation of academic engagement. Linking the relationship between interest, effort cost, and engagement, interest is expected to directly and indirectly predict different forms of engagement by lowering the perception of effort cost.

Gender Differences

According to a meta-study, there is no difference in math achievement between boys and girls, but boys are more likely to have high confidence, intrinsic motivation, and extrinsic motivation compared to girls (Else-Quest et al., 2010). Although there is no gender difference in the perception of values such as importance or usefulness, interest in mathematics has been somewhat consistently higher for male students than for female students (Pajares and Graham, 1999). However, recently, there has been a finding that there is no gender difference in interest (Gaspard et al., 2017). Therefore, there is still a need to accumulate more up-to-date data on gender differences in math motivation. Also, no gender difference in effort cost was found in the previous research.

Regarding gender difference in academic engagement, Hyde and Linn's (2006) recent meta-study provided evidence for gender similarities in mathematics and science. They showed that there is no gender difference in mathematics grades and that there is no gender difference in complex problem solving in elementary and middle school. They say that in 2001, American women received 48% of the bachelor's degrees in mathematics. These findings suggest that the motivation and persistence of men and women in mathematics may have reached similar levels. Furthermore, one study found that girls believed less about their math abilities than boys did, but displayed fewer work-avoidance goals, which aim to get as little involvement in a task as possible (Chouinard and Roy, 2008). This result means that boys' engagement might be more likely to appear as maladaptive, such as effort avoidance, than does that of girls.

Gender differences mentioned above are about differences in mean levels. Like this, gender difference studies in mathematics have focused on differences in mean levels. Thus, there is less understanding of gender differences in the relationship between engagement and motivational beliefs such as interest and effort cost. Only some researchers have reported higher interest-achievement correlations in boys than in girls, meaning that boys are more influenced by their math interests than girls (Reeve and Hakel, 2000; Denissen et al., 2007). Therefore, it is meaningful to investigate whether there would be gender differences in the relationships between math interests, perceived effort cost, and engagement.

Present Research

The purposes of this study were (1) to identify the role of interest and effort cost in persistence, cognitive engagement, and effort avoidance and (2) to examine gender differences in these roles. We preliminarily conducted a latent profile analysis to identify learners who had high interest and low perceived effort cost and who were high on both. We then used a structural equation model to test a hypothesized model. As **Figure 1** shows, we hypothesized interest would positively predict persistence and cognitive engagement and negatively predict effort avoidance. By contrast, a recent research (Jiang et al., 2018) has shown that effort cost plays a more important role in predicting effort avoidance positively than in predicting persistence and cognitive engagement negatively. In addition, we tested a hypothesis that interest could negatively predict effort cost. Thus, effort cost was tested as a mediator especially in the relationship between interest and effort avoidance. Last, we run a multi-group analysis for the examination of gender differences in relationships among interest, effort cost, and engagement variables.

MATERIALS AND METHODS

Participants

Five-hundred and sixty-three 8th and 9th graders from three middle schools in two metropolitan cities in South Korea voluntarily participated in this survey. The age of the students ranged from 13 to 16, and the average age was 14.12. Of the 563 students, six students did not complete the survey, and another eleven responded in a way that suggested their answers were insincere (e.g., choosing the same number for more than half the items in the entire survey). The size of the final sample used in this study was thus 546 (305 boys and 241 girls; 325 eighth graders and 221 ninth graders).

Procedure

All the schools participated in the study were informed of the study and were given an opt-out option. The survey was administered during a regular class period. Parents were informed about the study through school announcement, and none of the parents raised doubts about this study. All participants voluntarily joined study and signed a written informed consent form. This survey study, collecting only students' perceptions of learning and motivation without personal identifying information, did not include any vulnerable participants. The participants were also assured that their responses would not be disclosed to their parents or teachers but used only for research purposes. To reassure the students about the privacy of their responses, they sealed their questionnaire with a sticker after completing the survey. The research protocol was approved by the Institutional Review Board (IRB), Korea University.

Measures

The students responded to all items using a 7-point Likert-type scale, where 1 indicated the strongest disagreement and 7 indicated the strongest agreement for all items. All items referred to mathematics, because the students were expected to have formed clear perceptions of the interest and effort cost associated with this subject (see **Appendix**). In addition, mathematics is a subject in Korea that is used to define academic success.

Interest

We used five interest items included in the Student Motivation in the Learning Environment Scale (SMILES; Bong et al., 2012) based on interest literature (Hidi and Harackiewicz, 2000; Hidi and Renninger, 2006). Reliability and validity of the scale have been systematically examined by using Korean sample in a previous research (Bong et al., 2012). In the validation study, the scale exhibited sufficient reliability ($\alpha = 0.94$). Factor analysis supported its construct validity. High correlation of the SMILES interest scale with the existing interest scale ($r = 0.77$) supported its concurrent validity (Marsh et al., 2005). The reliability of the scale was acceptable in the present study ($\alpha = 0.85$).

Effort Cost

For effort cost items were adopted from Gaspard et al. (2015), in which the scale reliability was reliable ($\rho = 0.90$). The scale

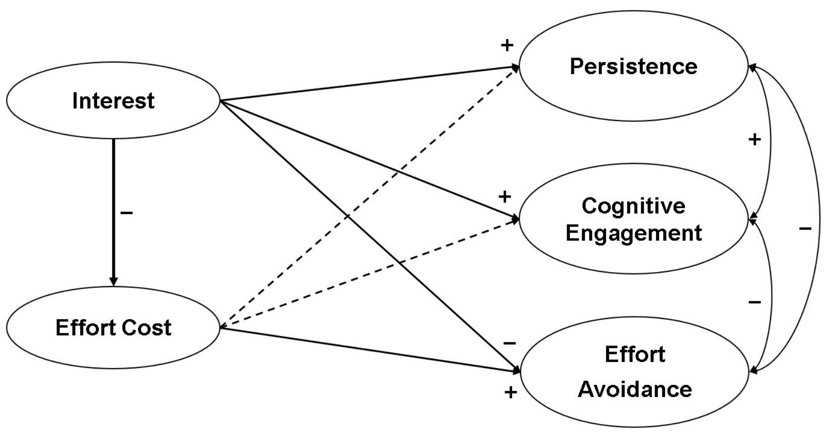


FIGURE 1 | A hypothesized model. Dashed lines indicated a nonsignificant path. For clarity, indicators, measurement errors, and disturbance terms are not presented.

produced a reliable Cronbach's α coefficient (α = 0.87) in the present study.

Persistence

Persistence is measured with seven items (e.g., "If a mathematics problem is really hard, I keep working on it"), which were adopted from Miserandino (1996). The reliability was reported as 0.94 in the previous research using the original English version of the scale and reported as 0.89 in the previous research using the Korean version of the scale (Song et al., 2012). In the present study, the scale showed a reliable Cronbach's α coefficient as 0.89.

Cognitive Engagement

Eight cognitive engagement items (e.g., "I try to understand rather than just memorize how to solve a mathematics problem") were adopted from the cognitive engagement scale included in the SMILES (Bong et al., 2012). Cognitive engagement items were developed based on Bloom's taxonomy and consisted of six levels of cognitive processing: remembering, understanding, applying, analyzing, evaluating, and creating (Krathwohl, 2002). Because the original scale items referred to academic study in general, all items in this study were thus revised to specifically refer to mathematics. The reliability of the scale was acceptable (α = 0.74).

Effort Avoidance

To measure effort avoidance, three items were adopted from Song et al. (2017) research. The validity and reliability of the scale has already been systematically tested (Song et al., 2017). The original items were revised to refer to mathematics. The scale produced a reliable Cronbach's α coefficient (α = 0.80).

Data Analysis

In this study, students were nested in 19 classes. Although multilevel analysis was not an original purpose of the present study, data had a complex structure. Thus, we applied design-based correction of standard errors to avoid underestimation of standard errors considering that the data had complex data structures (McNeish et al., 2017). The full information maximum likelihood (FIML) approach was applied to handle some missing data (0.0 to 0.88% of all items). We first conducted Latent Profile Analysis (LPA). The number of profiles was determined based on AIC (Akike Information Criterion), BIS (Bayesian Information Criterion), SABIC (Sample-size Adjusted BIC), adjusted LMR(Lo-Mendell-Rubin adjusted Likelihod Ratio Test), BLRT (Parametric Bootstrapped Likelihod Ratio Test), and entropy (McLachlan and Peel, 2000; Dias and Vermunt, 2006; Nylund et al., 2007).

Next, we conducted a structural equation modeling (SEM) with a whole sample in order to test the hypothesized model. Persistence and cognitive engagement scales have many items. They were parceled into three indicators after considering the small sample size compared to parameters to be estimated (Kline, 2011). Chi-square statistics, the Tucker-Lewis Index (TLI), comparative fit index (CFI), and the root mean square error of approximation (RMSEA) were used to evaluate the overall fit of the models. For the CFI and TLI, a coefficient above 0.90 indicates a suitable fit (Hu and Bentler, 1999) and for the RMSEA, values under 0.05 represent a close approximate fit, and values between 0.05 and 0.08 suggest an acceptable fit (Browne and Cudeck, 1993).

We then conducted a multi-group analysis to examine gender differences in predictive paths. For this, we tested measurement, covariance, and structural invariance over gender. To test measurement invariance, we constrained the factor loadings to be equal across gender. Then, we additionally constrained covariances between persistence, cognitive engagement, and effort avoidance to test covariance invariance. According to Chen (2007) and Cheung and Rensvold (2002), a decrease of less than 0.01 in the fit of the more parsimonious model on the CFI and TLI, an increase of less than 0.015 in the RMSEA, and an increase in the SRMR of 0.030 were considered as support for the more constrained model. Last, to test gender differences in path coefficients, we added constraints of the path coefficients one by one and tested the chi-square difference for every addition of a single path constraint. All analyses were conducted in Mplus 7.31.

RESULTS

Descriptive Statistics and Correlations

Table 1 presents the descriptive statistics for all the scales. The presence of moderate mean scores, low skewness (less than |0.36|), and low kurtosis (less than | 0.53|) indicates that the scales all produced a range of scores that had an approximately normal distribution (Kline, 2011). The correlations between all latent variables are also presented in **Table 1**. As expected, interest was negatively correlated with effort cost ($r = -0.42$, $p < 0.001$). Interest was positively correlated with persistence ($r = 0.81$, $p < 0.001$) and cognitive engagement ($r = 0.78$, $p < 0.001$), and it was negatively correlated with effort avoidance ($r = -0.47$, $p < 0.001$). Conversely, effort cost was negatively correlated with persistence ($r = -0.35$, $p < 0.001$) and cognitive engagement ($r = -0.27$, $p < 0.001$), and it was positively correlated with effort avoidance ($r = 0.51$, $p < 0.001$). Persistence was highly and positively correlated with cognitive engagement ($r = 0.97$, $p < 0.001$). Effort avoidance was negatively correlated with persistence ($r = -0.66$, $p < 0.001$) and cognitive engagement ($r = -0.53$, $p < 0.001$).

TABLE 1 | Descriptive statistics and latent correlations.

	1	2	3	4	5
1. Interest	–				
2. Effort cost	−0.42	–			
3. Persistence	0.81	−0.35	–		
4. Cognitive engagement	0.78	−0.27	0.97	–	
5. Effort avoidance	−0.47	0.51	−0.66	−0.53	–
M	3.73	3.86	4.41	4.34	3.31
SD	1.29	1.36	1.07	1.16	1.27
Skewness	0.04	0.08	−0.15	−0.30	0.36
Kurtosis	−0.42	−0.53	0.12	0.27	−0.12
α	0.85	0.87	0.89	0.74	0.80

N = 546. All correlation coefficients were significant at p < 0.001.

TABLE 2 | Latent profile solutions.

N of Profile	AIC	BIC	Adjusted BIC	Adjusted LMR	Entropy	Class sizes (%)					
2	7097.035	7165.877	7115.087	<0.001	0.775	0.53	0.47				
3	6773.654	6868.312	6798.475	0.141	0.845	0.16	0.60	0.24			
4	**6591.932**	**6712.406**	**6623.523**	**0.052**	**0.849**	**0.13**	**0.39**	**0.10**	**0.37**		
5	6554.862	6701.151	6593.222	0.485	0.821	0.10	0.16	0.32	0.37	0.05	
6	6512.120	6684.225	6557.249	0.331	0.837	0.03	0.09	0.15	0.35	0.05	0.33

The final selected model is shown in bold.

Latent Profiles

The number of profiles was determined based on AIC (Akike Information Criterion), BIS (Bayesian Information Criterion), SABIC (Sample-size Adjusted BIC), adjusted LMR(Lo-Mendell-Rubin adjusted Likelihod Ratio Test), BLRT (Parmametric Bootstrapped Likelihod Ratio Test), and entropy. As shown in **Table 2**, all indices supported a 4-profile solution. First, the lower the scores of AIC, BIS, and SABIC, the better the fit (Nylund et al., 2007). All three indices for each model are plotted in **Supplementary Figure S1**. AIC, BIC, and SABIC continued to go down as more latent profiles were added. However, slopes appeared to flatten out between 4 and 6 profiles. BLRT for each model was statistically significant. However, adjusted LMR was significant only up to four profiles, supporting the 4-profile solution (McLachlan and Peel, 2000). As seen in **Supplementary Figure S2**, entropy was also the highest at 0.849 in 4-profile solutions. Thus, the 4-profile classification was considered the most accurate (Dias and Vermunt, 2006).

As expected, interest, persistence, and cognitive engagement were either all high or all low. Mean levels of effort cost and effort avoidance appeared to be opposite to them. As shown in **Figure 2**, four profiles were identified. Profile 1 showed high levels of interest, persistence, and cognitive engagement but low levels of effort cost and effort avoidance. Students in this profile 1 seemed to have the most positive perceptions of math. Both Profile 2 and 3 displayed moderate levels in all variables. Profile 2 displayed relatively higher levels of interest, persistence, and cognitive engagement but lower levels of effort cost and effort avoidance than the average, whereas Profile 3 presented relatively lower levels of interest, persistence, and cognitive engagement but higher levels of effort cost and effort avoidance than the average. Lastly, Profile 4 showed low levels of interest, persistence, and cognitive engagement but high levels of effort cost and effort avoidance.

Test of the Hypothesized Model

The model showed acceptable fit to the data, χ^2 (125, N = 546) = 376.576, $p < 0.001$ (CFI = 0.938, TLI = 0.924, RMSEA = 0.061, SRMR = 0.056). As **Table 3** shows, consistent with the hypotheses, interest positively predicted persistence ($\beta = 0.80$, $p < 0.001$) and cognitive engagement ($\beta = 0.82$, $p < 0.001$) and negatively predicted effort avoidance ($\beta = -0.31$, $p < 0.001$). Moreover, interest negatively predicted effort cost ($\beta = -0.42$, $p < 0.001$). Effort cost did not predict either persistence ($\beta = -0.02$, $p = 0.584$) or cognitive engagement ($\beta = 0.08$, $p = 0.153$), but positively predicted effort avoidance ($\beta = 0.38$, $p < 0.001$). By doing so, effort cost significantly mediated the path from interest to effort avoidance ($\beta = -0.16$, $p < 0.001$, see **Table 4**). The standardized path coefficients of the total effect of interest on effort cost was thus −0.47 ($p < 0.001$). Persistence and cognitive engagement were highly correlated with each other ($r = 0.93$, $p < 0.001$). For robustness check, we examined whether similar results were found when excluding either persistence or cognitive engagement from the hypothetical model. As a result, model fits were similar. Significance of path coefficients did not change.

Tests of Gender Differences

Table 5 shows descriptive statistics for each female and male sample. The presence of moderate mean scores, low skewness, and low kurtosis support that the scales all produced a range of scores that had an approximately normal distribution in both the female and the male sample (Kline, 2011). Gender difference was found only in the mean value of effort avoidance ($t = 3.04$, $p = 0.003$).

Multi-group analysis was performed to investigate gender differences in the predictive relationships in the model. First, we checked measurement invariance by constraining all factor loading to be equal. As shown in **Table 6**, the measurement invariance model (Model 1) fitted the current data well, χ^2(281, N = 546) = 759.25, $p < 0.001$ (CFI = 0.920, TLI = 0.913, RMSEA = 0.079, SRMR = 0.068). We next tested a covariance invariance model by constraining covariance invariances between persistence, cognitive engagement, and effort avoidance to be equal (Model 2). Compared to the measurement invariance model, no decreases in the GFI and TLI and the low increases in the RMSEA and SRMR (ΔCFI = 0.000, ΔTLI = 0.001, ΔRMSEA = 0.000, ΔSRMR = 0.001) provide support for the covariance invariance model, indicating there were no gender differences in the covariance coefficients.

To examine gender difference in regression coefficients, we consecutively compared the chi-square of models from Model 3 to Model 9. For example, to test gender difference in the path from interest to effort cost, we additionally constrained the path to be equal across gender (Model 3) and compared the chi-square of Model 3 with that of Model 2, in which all paths were freely estimated but factor loadings and covariance were equally constrained ($\Delta df = 1$). To test gender difference in the path from interest to persistence, we again additionally constrained the path to be equal (Model 4) and compared the chi-squares between Model 3 and Model 4 ($\Delta df = 1$). For all seven path coefficients, gender differences were examined in the same way, and gender differences were found only in the path from interest to effort cost. **Table 3** presents results from the final model. As **Figure 3**

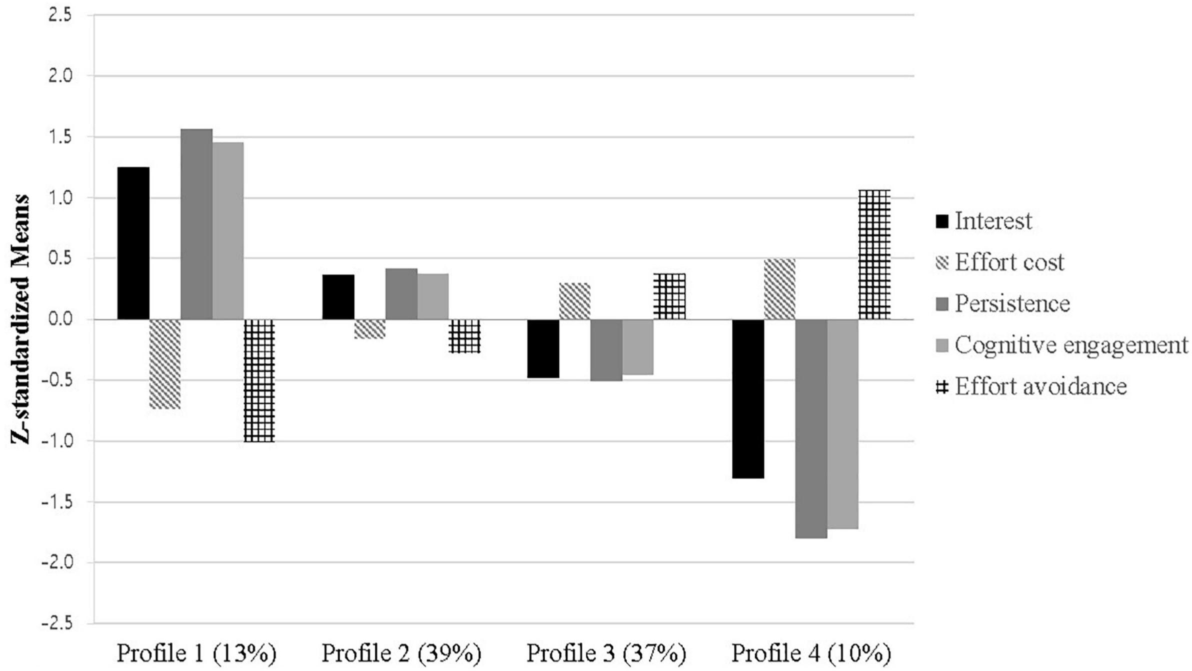


FIGURE 2 | Z-standardized means of profiles.

TABLE 3 | Results from SEM and multi-group analysis.

	Whole			Female students			Male students		
Path	**β**	***SE***	***p***	**β**	***SE***	***p***	**β**	***SE***	***p***
Interest → Effort cost	−0.422	0.06	<0.001	−0.663	0.05	<0.001	−0.278	0.07	<0.001
Interest → Persistence	0.802	0.04	<0.001	0.784	0.05	<0.001	0.805	0.03	<0.001
Effort cost → Persistence	−0.016	0.03	0.584	−0.024	0.03	0.479	−0.023	0.03	0.480
Interest → Cog eng	0.815	0.04	<0.001	0.786	0.06	<0.001	0.843	0.04	<0.001
Effort cost → Cog eng	0.075	0.05	0.153	0.081	0.06	0.160	0.081	0.06	0.160
Interest → Eff avoid	−0.307	0.05	<0.001	−0.297	0.05	<0.001	−0.299	0.05	<0.001
Effort cost → Eff avoid	0.380	0.05	<0.001	0.390	0.05	<0.001	0.367	0.05	<0.001
Persistence ↔ Cog eng	0.928	0.03	<0.001	0.901	0.05	<0.001	0.956	0.05	<0.001
Persistence ↔ Eff avoid	−0.572	0.07	<0.001	−0.706	0.08	<0.001	−0.541	0.05	<0.001
Cog eng ↔ Eff avoid	−0.374	0.06	<0.001	−0.414	0.055	<0.001	−0.387	0.06	<0.001
R^2									
Effort cost	0.178	0.05	<0.001	0.439	0.06	<0.001	0.077	0.04	0.038
Persistence	0.655	0.06	<0.001	0.639	0.07	<0.001	0.658	0.05	<0.001
Cog eng	0.618	0.05	<0.001	0.540	0.06	<0.001	0.679	0.05	<0.001
Eff avoid	0.337	0.04	<0.001	0.394	0.06	<0.001	0.285	0.04	<0.001

Cog eng, cognitive engagement; Eff avoid, effort avoidance.

shows, the prediction of interest for effort cost was much larger for female students ($\beta = -0.66$, $p < 0.001$) than for male students ($\beta = -0.28$, $p < 0.001$).

DISCUSSION

The present study examined the predictive relationship between interest, effort cost, and academic engagement. We also verified if there are gender difference. Frist, consistent with previous interest research (Schiefele, 1992; Mayer, 1998; Renninger et al., 2002), we observed that interest played a significant role in predicting students' persistence and cognitive engagement. We further found that interest negatively predicted effort avoidance as well. Second, interest negatively predicted effort cost, which in turn positively predicted effort avoidance, indicating that effort cost mediated the relationship between interest and effort avoidance. In addition, a latent profile analysis supported our

TABLE 4 | Total, direct, and indirect effects of indirect paths.

Path			Total			Direct			Indirect		
			β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Interest → Effort cost	→	Persistence	0.81	0.03	<0.001	0.80	0.04	<0.001	0.01	0.01	0.577
	→	Cog eng	0.78	0.03	<0.001	0.82	0.04	<0.001	−0.03	0.02	0.169
	→	Eff avoid	−0.47	0.05	<0.001	−0.31	0.05	<0.001	−0.16	0.03	<0.001
Female											
Interest → Effort cost	→	Persistence	0.81	0.03	<0.001	0.81	0.03	<0.001	0.01	0.01	0.439
	→	Cog eng	0.82	0.03	<0.001	0.84	0.04	<0.001	−0.02	0.02	0.240
	→	Eff avoid	−0.40	0.05	<0.001	−0.30	0.05	<0.001	−0.10	0.03	<0.001
Male											
Interest → Effort cost	→	Persistence	0.80	0.05	<0.001	0.78	0.05	<0.001	0.02	0.02	0.479
	→	Cog eng	0.73	0.04	<0.001	0.79	0.06	<0.001	−0.05	0.04	0.164
	→	Eff avoid	−0.56	0.05	<0.001	−0.30	0.05	<0.001	−0.26	0.04	<0.001

Cog eng, cognitive engagement; Eff avoid, effort avoidance.

TABLE 5 | Descriptive statistics and mean difference by gender.

Variable	Female students (n = 241)					Male students (n = 305)						
	M	*SD*	*S*	*K*	α	*M*	*SD*	*S*	*K*	α	*t*	*d*
Interest	3.63	1.19	0.11	−0.16	0.85	3.80	1.35	−0.03	−0.58	0.87	1.54	0.13
Effort cost	3.79	1.28	0.03	−0.30	0.87	3.93	1.37	0.08	−0.60	0.87	1.22	0.11
Persistence	4.37	1.04	−0.02	−0.06	0.90	4.44	1.09	−0.24	0.25	0.89	0.76	0.07
Cognitive engagement	4.38	0.93	0.03	0.17	0.83	4.38	1.00	−0.27	0.39	0.86	0.12	0.00
Effort avoidance	3.12	1.21	0.27	−0.58	0.80	3.45	1.30	0.40	0.06	0.79	3.04*	0.26

*S, skewness; K, kurtosis. *p < 0.01.*

hypothesis. Levels of interest and effort costs varied depending on the type of profiles. If one was high, the other was low. Lastly, we found that boys were more likely to avoid effort than were girls. Apart from this, there was no difference between boys and girls in the mean of interest, effort cost, persistence, and cognitive engagement. Rather, the negative prediction for the perception of effort cost by interest was found to be even greater for girls than for boys.

Importance of Interest in Academic Engagement

Interest has been identified as a potent motivator, containing both emotional and cognitive aspects. Individual interest for mathematics is developed by both enjoyment and internalizing importance of the task, which is positively linked to various forms of student engagement, such as persistence, effort, deep-levels of cognitive engagement, and classroom engagement (Sansone and Thoman, 2005; Hidi and Renninger, 2006; Kim et al., 2015). In particular, cognitive engagement using a wide range of cognitive strategies from memorizing to critical thinking is essential to fully understand the learning material (Krathwohl, 2002; Mayer, 2002). Our study showed that interest was largely related to cognitive engagement, covering a full spectrum of cognitive processing such as outlining knowledge, applying knowledge to real life, and connecting new and old knowledge, consistent with previous studies (Walker et al., 2006; Wang et al., 2017). Our finding that interest positively predicts persistence also demonstrates that math interest of students helps them continue their studies even if they have difficulty in studying mathematics. Furthermore, interest is negatively linked to effort avoidance, which is a maladaptive form of engagement (e.g., minimizing effort investment and skipping difficult parts). That is, if students are less interested in mathematics, they may be reluctant to put more effort into math learning.

In sum, math interest plays a pivotal role in studying mathematics because it plays two important roles together: positively predicting persistence and cognitive engagement and negatively predicting effort avoidance. Given the strong connection between interest and engagement, researchers and educators need to help students to develop their interest in mathematics by making the learning environment more enjoyable.

Effort Cost as an Avoidance Motivation

Among various forms of cost (e.g., effort cost, opportunity cost, and psychological cost) effort cost has been identified as the most salient feature of their least motivating classes. Example statement was "*these courses are too intense, requiring too much time, or being too rigorous*" (Flake et al., 2015). Perez et al. (2014) also discovered that only effort cost was strongly involved in the intent to leave a STEM major, whereas task value did not contribute to this decision.

TABLE 6 | Model fit statistics for a multi-group analysis.

M	Hypothesis	χ^2	*df*	CFI	TLI	RMSEA	SRMR	$\Delta\chi^2$
1	Measurement invariance	637.104	281	0.921	0.914	0.068	0.068	
2	Covariance invariance	638.856	284	0.921	0.915	0.068	0.069	
3	Interest → Effort cost	657.751	285	0.917	0.911	0.069	0.082	18.90**
4	Interest → Persistence	658.403	286	0.917	0.912	0.069	0.082	0.65
5	Effort cost → Persistence	661.924	287	0.917	0.911	0.069	0.088	3.52
6	Interest → Cog eng	663.975	288	0.917	0.911	0.069	0.090	2.05
7	Effort cost → Cog eng	665.721	289	0.916	0.912	0.069	0.091	1.75
8	Interest → Eff avoid	666.789	290	0.916	0.912	0.069	0.093	1.07
9	Effort cost → Eff avoid	668.118	291	0.916	0.912	0.069	0.092	1.33
10	Final model	646.462	290	0.921	0.917	0.067	0.071	

*M, model; Cog eng, cognitive engagement; Eff avoid, effort avoidance. **p < 0.001.*

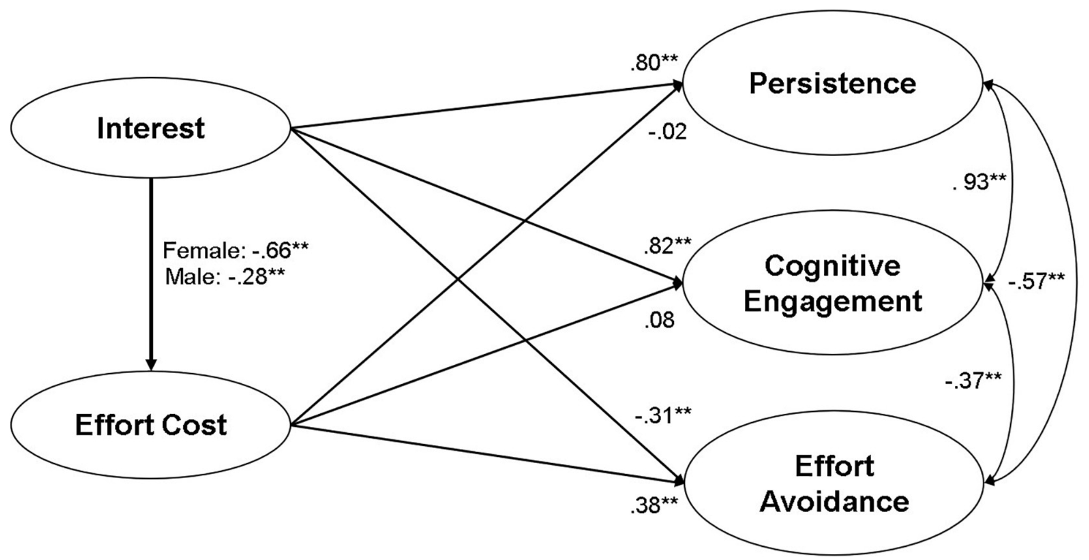


FIGURE 3 | Results from SEM and multi-group analysis. The standardized coefficients for male and female students are separately presented only for the interest-effort cost path that is significantly different between the two. For clarity, indicators, measurement errors, and disturbance terms are not presented. $^{**}p < 0.001$.

Effort cost seems to be more deeply involved in avoidance motivation. Previous research has also observed that effort cost serves as a motivation to avoid the least motivating classes or to leave courses and a STEM major (Eccles et al., 1983; Croxson et al., 2009; Perez et al., 2014; Flake et al., 2015). Consistent with previous findings, effort cost positively predicted effort avoidance only. Findings from this study thus support previous findings that a task-related cost is a precursor of avoidance motivation and has been related to avoidance-oriented maladaptive behaviors and negative emotions, such as test anxiety, negative classroom affect, disorganization, and procrastination (Jiang et al., 2018).

Negative Relationship Between Interest and Effort Cost

When considering the detrimental role of effort cost in motivation, the negative relationship between interest and effort cost is particularly noteworthy. This inverse relationship was supported by both person-centered and variable-centered approaches. We found four different profiles based on five variables (interest, effort cost, persistence, cognitive engagement, and effort avoidance). In all four profiles, students did not have similar levels of interest and effort cost. Rather, levels of interest and effort cost were opposite.

Moreover, as expected, interest appeared as a negative predictor of effort cost perception, although this finding was based on cross-sectional data. It can be conjectured that replenishing energy, automatic attention, or both may be potential mechanisms of interest in predicting perceptions of effort cost. However, the present study did not directly examine this. The negative relationship between interest and effort cost might be due to an unknown third variable. For example, task difficulty can increase effort cost while reducing interest (Eccles et al., 1983; Inzlicht et al., 2018). Therefore, further research is needed to explain the negative relationship between interest and effort costs. Nevertheless, this study is still significant because it reveals a direct connection between interest and effort cost.

Gender Differences

Interestingly, the negative relationship between interest and effort cost is more evident in female students. In previous studies, there was a higher correlation between interest and achievement or knowledge for males than for females (Reeve and Hakel, 2000; Denissen et al., 2007). Denissen et al. (2007) assumed that the low correlation between interest and achievement arises because female students focus on all areas to achieve overall high achievement rather than on the one area in which they are most interested and building a deeper level of knowledge of it. However, given the possibility that interest lowers perception of effort cost, female students' math interest also plays a meaningful role in mathematics-related learning and decision making. Especially, many researchers are trying to find out why female students are leaving the STEM area, including mathematics, and effort cost is supposed to be one of the factors that predict intent to leave (Eccles et al., 1983; Perez et al., 2014). The findings suggest that it is necessary to consider both positive values (e.g., interest) and negative costs (e.g., effort cost) to understand female students' engagement in mathematics.

When it comes to mean differences, there was no gender difference in the mean levels of all variables except for effort avoidance. Previous research has shown that there is no difference between male and female students in actual achievement, but there are gender differences in motivational beliefs in mathematics, such as interest and beliefs about competence (Else-Quest et al., 2010). Gender stereotypes have been identified as one core factor among the possible causes of these gender differences in mathematics (Jacobs, 1991; Spencer et al., 1999). However, it is not only our research that has not found gender differences in math beliefs. Recent German data has also reported that there is no gender difference in expectancy and value beliefs in mathematics (Gaspard et al., 2017). More research is thus needed to confirm whether gender stereotypes are socially weakened and gender differences are decreasing in students' subjective beliefs about mathematics as well as in their actual achievement.

This study even showed that boys are more likely than girls are to show maladaptive patterns of engagement, such as skipping hard part. This is in line with the existing literature that showed boys are more likely to pursue work-avoidance goals in mathematics than girls are (Chouinard and Roy, 2008). It is therefore necessary to accumulate up-to-date data on recent changes in beliefs about mathematics, apart from past findings of gender differences. Nevertheless, there are still some recent studies reporting gender differences in mathematics (Tian et al., 2018). Considering this, research will need to be conducted to examine gender differences in various samples and contexts as to which elements reduce gender differences.

Limitations and Future Directions

Considering the limitations of this study, we suggest future research directions as follows. First, based on the theoretical background, we hypothesized that interest leads to the perception of effort cost, but we measured all variables at one point. In fact, there is little research that draws conclusions about the temporal relationship or causal relationship between interest and effort cost, so further experimental and longitudinal studies should be accumulated. Also, effort cost has emerged as an important factor in academic motivation and choice-related behaviors, especially in avoidance-related behaviors such as procrastination (Perez et al., 2014; Jiang et al., 2018). In addition to interest, more systematic research is needed on what factors lower the perception of effort cost. Basically, it will be influenced by the absolute amount of the task, but it will also be affected by actual characteristics of the task, such as task difficulty, and other subjective perceptions about the task, such as belief in one's competence (Eccles et al., 1983). It is also necessary to further examine whether other types of task value, such as utility value, play the same role in effort cost perception as did interest.

CONCLUSION

The current study has several theoretical and educational implications. It also presents a new perspective on the role of interest in relation to effort cost. First, the present research makes a contribution to interest literature by demonstrating a new role of interest. Although previous studies have shown that interest plays a significant role in promoting academic engagement (Schiefele, 1992; Durik et al., 2006; Walker et al., 2006), our findings suggest that interest is an important predictor of both effort cost perception and effort avoidance. Second, the current study shows that there is a need to pay more attention to the role of effort cost. The classic expectancy-value theory has focused only on task values such as interest and utility value, but not on task costs (e.g., Eccles et al., 1983; Wigfield and Eccles, 2000). Although traditional studies have clearly shown that task values function as a critical factor in predicting achievement-related outcomes, there is little research on cost perception. Recently, researchers have begun to propose the new Expectancy-Value-Cost (EVC) approach which highlights the importance of cost perceptions in learning process (Jiang et al., 2018). The present study further showed that effort cost played a mediating role between interest and effort avoidance, providing additional empirical evidence for supporting the recent EVC approach.

This study also has practical implications for educators. Given the strong relationship between interest, effort cost, and engagement, teachers and parents need to be aware of the importance of interest. Especially, considering the stronger link between interest and effort cost for female students, more attention should be paid to their interest in mathematics. Since girl's math interest can be affected by parents and teachers' gender-related math attitudes and stereotypes through their behaviors and communications (Gunderson et al., 2012), teachers and parents need to be careful about their words and actions so that they would not negatively affect girls' interest in mathematics.

ETHICS STATEMENT

All study participants provided informed consent, and the research protocol was approved by the Institutional Review Board (IRB), Korea University.

AUTHOR CONTRIBUTIONS

JS, SK, and MB designed the research and have contributed to writing the manuscript. JS collected and analyzed the data.

FUNDING

The current research was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2017S1A5A2A03068451).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02146/full#supplementary-material

REFERENCES

Battle, A., and Wigfield, A. (2003). College women's value orientations toward family, career, and graduate school. *J. Vocat. Behav.* 62, 56–75. doi: 10.1016/s0001-8791(02)00037-4

Bong, M., Kim, S., Reeve, J., Lim, H. J., Lee, W., Ahn, H. S., et al. (2012). *The Student Motivation in the Learning Environment Scales (SMILES)*. Seoul: Korea University, Brain and Motivation Research Institute. Available at: http://bmri.korea.ac.kr/english/research/assessment_scales/list.html?id=assessment

Botvinick, M., Huffstetler, A., and McGuire, J. T. (2009). Effort discounting in human nucleus accumbens. *Cogn. Affect Behav. Neurosci.* 9, 16–27. doi: 10.3758/CABN.9.1.16

Browne, M. W., and Cudeck, R. (1993). "Alternative ways of assessing model fit," in *Testing Structural Equation Models*, eds K. A. Bollen, and J. S. Long, (Newbury Park, CA: Sage), 136–162.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Modeling* 14, 464–504. doi: 10.1080/10705510701301834

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Modeling* 9, 233–255.

Chouinard, R., and Roy, N. (2008). Change in high-school students' competence beliefs, utility value and achievement goals in mathematics. *Brit. J. Educ. Psychol.* 78, 31–50. doi: 10.1348/000709907x197993

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E. J., and Rushworth, M. F. S. (2009). Effort-based cost-benefit valuation and the human brain. *J. Neurosci.* 29, 4531–4541. doi: 10.1523/JNEUROSCI.4515-08.2009

Denissen, J. J. A., Zarrett, N. R., and Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: longitudinal coupling between domain-specific achievement, self-concept, and interest. *Child Dev.* 78, 430–447. doi: 10.1111/j.1467-8624.2007.01007.x

Dias, J. G., and Vermunt, J. K. (2006). "Bootstrap methods for measuring classification uncertainty in latent class analysis," in *Proceedings in Computational Statistics*, eds A. Rizzi, and M. Vichi, (Heidelberg: Springer), 31–41. doi: 10.1007/978-3-7908-1709-6_3

Durik, A. M., Vida, M., and Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: a developmental analysis. *J. Educ. Psychol.* 98, 382–393. doi: 10.1037/0022-0663.98.2.382

Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). "Expectancies, values, and academic behaviors," in *Achievement and Achievement Motivation*, ed. J. T. Spence, (San Francisco: Freeman), 75–146.

Eccles, J. S. (2016). Engagement: where to next? *Learn. Instr.* 43, 71–75. doi: 10.1016/j.learninstruc.2016.02.003

Eccles, J. S., and Wigfield, A. (1995). In the mind of the actor: the structure of adolescents' achievement task values and expectancy-related beliefs. *Pers. Soc. Psychol. B* 21, 215–225. doi: 10.1177/0146167295213003

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., and Welsh, M. E. (2015). Measuring cost: the forgotten component of expectancy-value theory. *Contemp. Educ. Psychol.* 41, 232–244. doi: 10.1016/j.cedpsych.2015.03.002

Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* 74, 59–109. doi: 10.3102/00346543074001059

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., et al. (2015). More value through greater differentiation: gender differences in value beliefs about math. *J. Educ. Psychol.* 107, 663–677. doi: 10.1037/edu0000003

Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., and Nagengast, B. (2017). Assessing task values in five subjects during secondary school: measurement structure and mean level differences across grade level, gender, and academic subject. *Contemp. Educ. Psychol.* 48, 67–84. doi: 10.1016/j.cedpsych.2016.09.003

Greene, B. A., and Miller, R. (1996). Influence on achievement goals, perceived ability, and cognitive engagement. *Contemp. Educ. Psychol.* 21, 181–192. doi: 10.1006/ceps.1996.0015

Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles* 66, 153–166. doi: 10.1007/s11199-011-9996-2

Hidi, S. (1995). A re-examination of the role of attention in learning from text. *Educ. Psychol. Rev.* 7, 323–350. doi: 10.1007/bf02212306

Hidi, S., and Ainley, M. (2008). "Interest and self-regulation: relationships between two variables that influence learning," in *Motivation and Self-Regulated Learning: Theory, Research, and Application*, eds D. H. Schunk, and B. J. Zimmerman, (Mahwah, NJ: Erlbaum), 78–109.

Hidi, S., and Baird, W. (1986). Interestingness: a neglected variable in discourse processing. *Cogn. Sci.* 10, 179–194. doi: 10.1207/s15516709cog1002_3

Hidi, S., and Harackiewicz, J. M. (2000). Motivating the academically unmotivated: a critical issue for the 21st century. *Rev. Educ. Res.* 70, 151–179. doi: 10.3102/00346543070002151

Hidi, S., and Renninger, K. A. (2006). The four-phase model of interest development. *Educ. Psychol.* 41, 111–127. doi: 10.1207/s15326985ep4102_4

Hidi, S., Renninger, K. A., and Krapp, A. (2004). "Interest, a motivational variable that combines affective and cognitive functioning," in *Motivation, Emotion, and Cognition: Integrative Perspectives on Intellectual Functioning and Development*, eds D. Y. Dai, and R. J. Sternberg, (New Jersey, NJ: Lawrence Erlbaum Associates), 89–115.

Higgins, E. T. (2006). Value from hedonic experience and engagement. *Psychol. Rev.* 113, 439–460. doi: 10.1037/0033-295x.113.3.439

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6, 1–55. doi: 10.1080/10705519909540118

Hyde, J. S., and Linn, M. C. (2006). Gender similarities in mathematics and science. *Science* 314, 599–600. doi: 10.1126/science.1132154

Inzlicht, M., Shenhav, A., and Olivola, C. Y. (2018). The effort paradox: effort is both costly and valued. *Trends Cogn. Sci.* 22, 337–349. doi: 10.1016/j.tics.2018.01.007

Isen, A. M., and Reeve, J. (2005). The influence of positive affect on intrinsic and extrinsic motivation: facilitating enjoyment of play, responsible work behavior, and self-control. *Motiv. Emotion* 29, 297–325.

Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *J. Educ. Psychol.* 83, 518–527. doi: 10.1037/0022-0663.83.4.518

Jiang, Y., Rosenzweig, E. Q., and Gaspard, H. (2018). An expectancy-value-cost approach in predicting adolescent students' academic motivation and achievement. *Contemp. Educ. Psychol.* 54, 139–152. doi: 10.1016/j.cedpsych.2018.06.005

Kim, S., Jiang, Y. K., and Song, J. (2015). "The effects of interest and utility value on mathematics engagement and achievement," in *Interest in Mathematics and Science Learnin*, eds K. A. Renninger, M. Nieswandt, and S. Hidi, (Washington, DC: American Educational Research Association), 63–78. doi: 10.3102/978-0-935302-42-4_4

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford.

Kool, W., McGuire, J. T., Rosen, Z. B., and Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *J. Exp. Psychol. Gen.* 139, 665–682. doi: 10.1037/a0020198

Krapp, A. (1999). Interest, motivation, and learning: an educational-psychological perspective. *Eur. J. Psychol. Educ.* 14, 23–40. doi: 10.1007/bf03173109

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: an overview. *Theor. Pract.* 41, 212–218. doi: 10.1146/annurev-clinpsy-032813-153710

Kurzban, R., Duckworth, A., Kable, J. W., and Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behav. Brain Sci.* 36, 661–726. doi: 10.1017/S0140525X12003196

Linnenbrink, E. A., and Pintrich, P. R. (2004). "Role of affect in cognitive processing in academic settings," in *Motivation, Emotion, and Cognition: Integrative Perspectives on Intellectual Functioning and Development*, eds D. Y. Dai, and R. J. Sternberg, (New Jersey, NJ: Lawrence Erlbaum Associates), 57–87.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., and Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Dev.* 76, 397–416. doi: 10.1111/j.1467-8624.2005.00853.x

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instr. Sci.* 26, 49–63.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theor. Pract.* 41, 226–232. doi: 10.1207/s15430421tip4104_4

McDaniel, M. A., Waddill, P. J., Finstad, K., and Bourg, T. (2000). The effects of text-based interest on attention and recall. *J. Educ. Psychol.* 92, 492–502. doi: 10.1037/0022-0663.92.3.492

McGuire, J. T., and Botvinick, M. M. (2010). Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc. Natl. Acad. Sci U.S.A.* 107, 7922–7926. doi: 10.1073/pnas.0910662107

McLachlan, G., and Peel, D. (2000). *Finite Mixture Models*. New York, NY: John Wiley & Sons, Inc.

McNeish, D., Stapleton, L. M., and Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychol. Methods* 22, 114–140. doi: 10.1037/met0000078

Miserandino, M. (1996). Children who do well in school: individual differences in perceived competence and autonomy in above-average children. *J. Educ. Psychol.* 88, 203–214. doi: 10.1037//0022-0663.88.2.203

Newmann, F., Wehlage, G. G., and Lamborn, S. D. (1992). "The significance and sources of student engagement," *in Student Engagement and Achievement in American Secondary Schools* ed. F. Newmann, (New York, NY: Teachers College Press), 11–39.

Nylund, K. L., Asparouhov, T., and Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct. Equ. Modeling.* 14, 535–569. doi: 10.1080/10705510701575396

Pajares, F., and Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemp. Educ. Psychol.* 24, 124–139. doi: 10.1006/ceps.1998.0991

Part, R., Perera, H., Bernacki, M. L., and Marchand, G. C. (2018). "Expectancies, values, and costs: reciprocal-effects models," in *Proceedings of the Paper Presented at the Annual Meeting of the American Educational Research Association*, (New York, NY).

Pekrun, R. (1992). The impact of emotions on learning and achievement: towards a theory of cognitive/motivational mediators. *Appl. Psychol. Int. Rev.* 41, 359–376. doi: 10.1111/j.1464-0597.1992.tb00712.x

Perez, T., Cromley, J. G., and Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *J. Educ. Psychol.* 106, 315–329. doi: 10.1037/a0034027

Reeve, C. L., and Hakel, M. D. (2000). Toward an understanding of adult intellectual development: investigating within-individual convergence of interest and knowledge profiles. *J. Appl. Psychol.* 85, 897–908. doi: 10.1037//0021-9010.85.6.897

Reeve, J., Lee, W., and Won, S. (2015). "Interest as emotion, as affect, as schema," in *Interest in Mathematics and Science Learning*, eds K. A. Renninger, M. Nieswandt, and S. Hidi, (Washington, DC: American Educational Research Association), 79–92. doi: 10.3102/978-0-935302-42-4_5

Renninger, K. A., Ewen, L., and Lasher, A. K. (2002). Individual interest as context in expository text and mathematical word problems. *Learn. Instr.* 12, 467–491.

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037//0003-066x.55.1.68

Sansone, C., and Thoman, D. B. (2005). Interest as the missing motivator in self-regulation. *Eur. Psychol.* 10, 175–186. doi: 10.1027/1016-9040.10.3.175

Schiefele, U. (1991). Interest, learning, and motivation. *Educ. Psychol.* 26, 299–323. doi: 10.1207/s15326985ep2603\%264_5

Schiefele, U. (1992). "Topic interest and level of text comprehension," in *The Role of Interest in Learning and Development*, eds K. A. Renninger, S. Hidi, and A. Krapp, (Hillsdale, NJ: Erlbaum), 151–182.

Song, J., Kim, S., and Bong, M. (2012). "Antecedents and consequents of academic goal-construal," in *Procceedings of the Paper Presented at the Annual Meeting of the American Educational Research Association*, (Vancouver, BC).

Song, J., Woo, Y., Bong, M., and Reeve, J. (2017). "The distinct role of three forms of cognitive engagement in an academic setting," in *Proceedings of the Paper Presented at the Annual Convention of the American Psychological Association*, (Washington, DC).

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35, 4–28.

Thoman, D. B., Smith, J. L., and Silvia, P. J. (2011). The resource replenishment function of interest. *Soc. Psychol. Personal. Sci.* 2, 592–599.

Tian, Y., Fang, Y., and Li, J. (2018). The effect of metacognitive knowledge on mathematics performance in self-regulated learning framework: multiple mediation of self-efficacy and motivation. *Front. Psychol.* 9:2518. doi: 10.3389/fpsyg.2018.02518

Vassena, E., Silvetti, M., Boehler, C. N., Achten, E., Fias, W., and Verguts, T. (2014). Overlapping neural systems represent cognitive effort and reward anticipation. *PLoS One* 9:e91008. doi: 10.1371/journal.pone.0091008

Walker, C. O., Greene, B. A., and Mansell, R. A. (2006). Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement. *Learn. Individ. Differ.* 16, 1–12. doi: 10.1016/j.lindif.2005.06.004

Wang, J., Liu, R.-D., Ding, Y., Xu, L., Liu, Y., and Zhen, R. (2017). Teacher's autonomy support and engagement in math: multiple mediating roles of self-efficacy, intrinsic value, and boredom. *Front. Psychol.* 8:1006. doi: 10.3389/fpsyg.2017.0100

Wigfield, A., and Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemp. Educ. Psychol.* 25, 68–81. doi: 10.1006/ceps.1999.1015

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

Interest

I find mathematics interesting.
I lose track of time when I study mathematics.
I want to learn more about mathematics outside of class.
I feel happy when I learn new things related to mathematics.
I want to have a mathematics-related career.

Effort Cost

Dealing with mathematics drains a lot of my energy.
Doing math is exhausting to me.
I often feel completely drained after studying mathematics.
Learning mathematics exhausts me.

Persistence

If a mathematics problem is really hard, I keep working at it.
If I can't understand mathematics contents right the first time, I just keep trying.
If I can't think of the answer to a mathematic question, after minutes it comes to me.
I really concentrate when my teacher presents new materials in mathematics.
When I have trouble with a mathematics problem, I usually get it right in the end.
When I get stuck on a mathematics question, I can usually get it.
I pay attention when I start a new subject in mathematics.

Cognitive Engagement

When I study mathematics, I explore out alternative solutions.
I question the validity of what I have learned in mathematic.
When I study mathematics, I distinguish main points from details.
I outline the learning materials that I learned in mathematics.
I apply what I learned in class while I do exercises.
I try to understand rather than just memorize the way to solve a mathematics problem.
When I learn something new in mathematics, I try to connect it to what I already know.
I try to remember what I have learned in mathematics as many as possible.

Effort Avoidance

I solve only easy problems while I do exercises.
When studying mathematics, I skip all the hard parts.
If I cannot understand the learning materials in mathematics, I just skip it.

frontiers in Psychology

ORIGINAL RESEARCH
published: 16 October 2019
doi: 10.3389/fpsyg.2019.02307

Mathematics Self-Concept in New Zealand Elementary School Students: Evaluating Age-Related Decline

Penelope W. St J. Watson*, Christine M. Rubie-Davies and Kane Meissel

School of Learning, Development and Professional Practice, Faculty of Education and Social Work, The University of Auckland, Auckland, New Zealand

The underrepresentation of females in mathematics-related fields may be explained by gender differences in mathematics self-concept (rather than ability) favoring males. Mathematics self-concept typically declines with student age, differs with student ethnicity, and is sensitive to teacher influence in early schooling. We investigated whether change in mathematics self-concept occurred within the context of a longitudinal intervention to raise and sustain teacher expectations of student achievement. This experimental study was conducted with a large sample of New Zealand primary school students and their teachers. Data were analyzed using longitudinal multilevel modeling with mathematics self-concept as the dependent variable and time (which represents students' increasing age each year), gender, and ethnicity entered as predictors and achievement in mathematics included as a control variable. Interaction terms were also explored to investigate changes over time for different groups. All students demonstrated a small increase in mathematics self-concept over the 3-year period of the current study but mathematics self-concept was consistently greater for boys than girls. Māori, Asian, and Other students' initial mathematics self-concept was higher than that of New Zealand European and Pacific Islanders' (after controlling for achievement differences). However, a statistically significant decline in mathematics self-concept occurred for Māori students alone by the end of the study. The expected age-related reduction over time in student mathematics self-concept appeared to be mitigated in association with the longitudinal study. Nevertheless, the demonstration of a comparatively lower mathematics self-concept remained for girls overall and declined for Māori. Our results reinforce implications for future research into mathematics self-concept as a possible determinant of female student career choices.

Keywords: mathematics self-concept, gender, ethnicity, age, career choice

OPEN ACCESS

Edited by:
M. Gail Jones,
North Carolina State University,
United States

Reviewed by:
Teomara Rutherford,
North Carolina State University,
United States
Sharinaz Hassan,
Curtin University, Australia

****Correspondence:***
Penelope W. St J. Watson
p.watson@auckland.ac.nz

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *30 November 2018*
Accepted: *27 September 2019*
Published: *16 October 2019*

Citation:
Watson PWSJ, Rubie-Davies CM and Meissel K (2019) Mathematics Self-Concept in New Zealand Elementary School Students: Evaluating Age-Related Decline. Front. Psychol. 10:2307. doi: 10.3389/fpsyg.2019.02307

INTRODUCTION

Females remain underrepresented in science, technology, engineering, and mathematics (STEM) fields, which have been traditionally considered masculine domains (Watt, 2010). This *status quo* exists concurrently with a shortage of a skilled STEM labor force (Jacobs, 2005) and counter to evidence that suggests gender diversity in the workplace is beneficial (Woolley et al., 2010). Further, despite demonstrating that their mathematics ability is equal to that of males (OECD, 2015), women choose careers in fields outside STEM and a gendered wage gap, disadvantaging them, continues (Sansone, 2016). Gender differences in elementary school students' mathematics confidence are larger than those for interest and achievement (Ganley and Lubienski, 2016). Moreover, although there are small gender differences in mathematics achievement, gender disparities in student career choice are large (Lindberg et al., 2010). Globally, a persistent gap in mathematics self-concept favoring males has been offered as a central explanation for differences in gendered STEM participation and, importantly, self-concept rather than ability appears to comprise the critical filter in career choice (Schoon, 2015). Mathematics self-concept has also differed by student ethnicity and has declined as student age increases (Wilkins, 2004). Socialization processes and interactions with significant others (e.g., teachers) have been posited as central to shaping students' gendered expectations of success and the value they attribute to specific fields (Schoon, 2015). Further, a self-concept that embraces male mathematics superiority has originated from cultural beliefs and expectations (Correll, 2001). Notably, self-concept may be especially vulnerable to the influence of teachers in the early years of schooling (Petersen and Hyde, 2014). Teacher expectations have been influenced by beliefs (including socialized gender beliefs) about students (Dusek and Joseph, 1985). We explored, therefore, whether an intervention in which teachers were trained in high expectation practices might be associated with mitigating differences in student mathematics self-concept by gender and ethnicity, and might ameliorate its decline with increasing age.

The current research reports on data situated within the context of a 3-year intervention study conducted longitudinally with a large sample of New Zealand elementary school students (n = 1,739) from a range of schools (n = 11). This wider study aimed to raise and sustain teachers' expectations of their students' academic achievement (see Rubie-Davies et al., 2015 for a full account), and was conducted with the student participants and their teachers. The main study provided professional development to all teachers in the project. The pedagogical approaches associated with high expectation practices were targeted with the aim of supporting all participant teachers to develop these practices. These approaches included mixed ability grouping for learning activities (Rubie-Davies, 2008), fostering a warmer socioemotional climate in the classroom, promotion of positivity toward and between students, the development of mastery goals, enhancing collaborative classrooms, and facilitation of student autonomy (Rubie-Davies and Peterson, 2011).

The current study drew on the self-concept data collected as part of the wider professional development project. Specifically, we assessed whether student mathematics self-concept varied for students overall, and by student gender and ethnicity.

Self-Concept

Byrne and Shavelson (1986) defined self-concept as one's own perception of one's abilities and efficacy. Specifically, self-efficacy (one's belief about one's ability to succeed in tasks) could be shaped by mastery and vicarious experiences, verbal persuasion, and one's psychological and affective state (Bandura, 1997). Further, a sense of self-confidence, self-esteem, and acceptance of self have also been subsumed within self-concept (Marsh and Scalas, 2011) and have formed the basis for perceptions of possible selves (Franken, 1994). Self-concept is multi-dimensional and hierarchical in nature, develops early in a child's educational career, and is positively associated with achievement (Chapman et al., 2000). Moreover, self-concept is influenced by experiences gained within one's environment including social comparison and evaluations by significant others (Bong and Skaalvik, 2003), and is continually re-appraised and reinforced by intra-personal inferences (Bong and Clark, 1999). As a sense of identity and awareness of others develops, one's self-concept (including domain-specific self-concept) tends to suffer a decline as a result of peer-comparison (Anderman and Maehr, 1994). Importantly, self-concept may be operationalized differently across cultures (Oettingen and Zosuls, 2006). Thus, examining self-concept within different contexts offers the opportunity to understand variations in student motivation and self-beliefs between cultures (Chiu and Klassen, 2008).

Mathematics Self-Concept

Mathematics self-concept has been defined as one's beliefs about one's competence in mathematics (Ireson and Hallam, 2009). Further, mathematics self-concept has been positively associated with mathematics achievement (Ireson and Hallam, 2009), and student ratings of their skill, enjoyment of, and interest in mathematics (Erdogan and Şengul, 2014). One's perceived ability to achieve well and one's confidence in mathematics have also been associated with mathematics self-concept (Reyes, 1984). Moreover, it has been suggested that mathematics self-concept has been associated with individuals' willingness to engage in quantitative scenarios (Eccles, 1987; Schoon, 2015).

Mathematics Self-Concept and Age

Mathematics self-concept has been shown to decline with schooling age (Wilkins, 2004) paralleling the age-related decline in general self-concept attributed to peer comparison (Anderman and Maehr, 1994). Further, the pattern of age-related decline has been supported in multiple contexts including that of New Zealand (Bonne, 2016).

In a study comprising a large number of students from the wider metropolitan area of Sydney, Australia (Grades 2–9, n = 3,679; Grades 7–11, n = 3,073, and 15 years of age and older, n = 1202), Marsh (1989) reported declines in mathematics self-concept in pre-adolescence and early adolescence (although finding some support for an increase in the construct's levels during late adolescence). Further, of all the subject domains, the decline in academic self-concept was sharpest in mathematics and at the beginning of middle school (Marsh, 1989). Furthermore, Marsh (1989) stated that his study's findings corroborated a consistent decline in mathematics self-concept with age found across numerous previous studies that had employed a range of robust instruments.

Fredricks and Eccles' (2002) study with 514 students from an urban area in the Midwest of the US, ranging from

Grades 1–12 (50% female; predominantly White American) further confirmed earlier findings (e.g., Marsh, 1989; Eccles et al., 1993) that perceptions of mathematics ability declined as grade level progressed. Fredricks and Eccles (2002) offered several explanations for the decline in belief in mathematics competence. In natural developmental patterns, for example, younger children's views of their competence are somewhat more optimistic but as children grow older, comparison with peers incites a more realistic self-concept (Fredricks and Eccles, 2002). In addition, competitiveness increases and the nature of assessments changes through into adolescence (Fredricks and Eccles, 2002). Further, elementary school student evaluation is often more mastery-goal oriented, whereas at middle and secondary school, normative and socially comparative marking systems will more frequently be employed (Eccles et al., 1993).

Further research supported the decline of mathematics self-concept over time with US children. Jacobs et al. (2002) findings (resulting from a longitudinal study of 761 predominantly White American Grade 1–12 students in a large Mid-western city) showed a decline in mathematics self-competence perceptions over the course of schooling. A sharp decline in mathematics self-beliefs about mathematics competence beginning at middle school had been reported in some previous research conducted over shorter periods (e.g., Marsh, 1989). Jacobs et al. (2002), however, revealed a steady decline of mathematics self-beliefs over the 12 years of their longitudinal study. Marsh and Ayotte's (2003) findings in their study of 1,103 French-speaking preadolescent Canadian children in Grades 2–6 further supported the pattern of the decline in math self-concept over time, within an additional context. Moreover, the findings of a longitudinal cohort-sequential study design conducted in Australia, Germany, and the US (Nagy et al., 2010) further endorsed the aforementioned pattern of decline. The authors found that for three cohorts of predominantly white Australian (n = 1,333), German (n = 4,688), and US (n = 2,378) secondary school students, mathematics self-concept declined across time. Erdogan and Şengul (2014) found the same pattern of decline in mathematics self-concept in a further context (Istanbul, Turkey) in a year-long study of 281 primary to secondary school students (Grades 4–6).

Relatively recent New Zealand data confirm a decline in mathematics self-beliefs (including mathematics self-concept) for New Zealand school-age children (Bonne, 2016). Bonne (2016) reported the findings of five New Zealand studies of student mathematics self-beliefs, comprising large numbers of children from early elementary to the end of middle school. From her synthesis of the findings, Bonne (2016) advocated for the power of teachers' influence in improving mathematics self-beliefs.

The results of Wilkins' (2004) large-scale study across multiple national contexts provided overarching evidence that mathematics self-concept declines with age. The author's findings were based on data drawn from the TIMSS study comprising 290,000 students in early adolescence from 41 countries. Further, Wilkins (2004) shed light on the little-explored intersectionality of mathematics self-concept with other demographic characteristics, for example, gender. Boys, for example, have been found to hold higher levels of mathematics self-concept than girls, albeit that levels of the construct diminished as age progressed for both genders (Marsh, 1989; Wilkins, 2004).

Mathematics Self-Concept and Gender

In terms of differences in the rigidity of gender role between cultures, more masculine cultures (e.g., Austria, see Hofstede, 2003) have been associated with an avoidance of careers that are considered gender-role inappropriate (Chiu and Klassen, 2008). Such cultural attitudes could be suggested to limit the engagement of females, for example, in STEM fields. With specific regard to mathematics, girls have been shown to underestimate their ability and have typically expected less success in mathematics than boys (Eccles, 1987; Wang and Kenny, 2014). Notably, female New Zealand elementary and middle school students underestimated their ability even when their mathematics achievement was higher than that of their male peers (Bonne, 2016). In a study of German 8-9-year olds, Dickhäuser and Meyer (2006) found that girls attributed success in mathematics less to high ability and failure in the subject more to low ability than boys. Similarly, in a meta-analysis of cross-national gender differences in mathematics using Trends in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA) data, Else-Quest et al. (2010) determined that boys showed more positive levels of mathematics confidence in almost all the participating countries.

In societies where gender-role rigidity was found to be greater (termed "masculine") girls' perception of themselves as mathematicians seemed to be further complicated (see Hofstede, 2003). Within such societies, girls were likely to value learning mathematics skills less, invest less effort and time in the subject, learn and achieve less mathematics, and have a lower mathematics self-concept than boys (Wigfield et al., 2004). Such scenarios have fostered girls' self-exclusion from mathematics and have resulted in boys' attainment of higher mathematics self-concepts despite their concurrent lower mathematics achievement (Chiu and Klassen, 2008).

The importance of differences in gender norms by cultural context for mathematics self-concept was underlined in further research. A gender gap in mathematics self-concept was found in Nagy et al. (2010) cross-national longitudinal study, and Erdogan and Şengul (2014) year-long study of Turkish students. Nagy et al. (2010) reported that boys' level of mathematics self-concept was consistently greater than that of girls across the cultural contexts of Germany, Australia, and the US, although mathematics self-concept declined across time for both genders. The gender gap in mathematics self-concept was largest, however, for the German cohort when compared to the Australian and US cohorts (Nagy et al., 2010). This difference occurred in parallel to higher levels of gender-norm differentiation (reflective of the gender-role rigidity noted by Hofstede, 2003) found in Germany than in Australia and the US (see Williams and Best, 1990).

Importantly, mathematics self-concept is needed, in addition to content knowledge, to ensure persistence in trajectories leading to STEM fields (Ceci and Williams, 2009) and high levels of aptitude will not guarantee engagement in STEM without a mathematics self-concept that facilitates beliefs of success in the field (Goldman and Penner, 2014). Specifically, mathematics self-concept has been identified as a critical filter to female engagement in STEM fields and low levels of mathematics self-concept has been used to explain women's underrepresentation in those fields (Schoon, 2015).

Mathematics Self-Concept and Ethnicity

Students in individualistic societies may place greater value on individuality engaging in downward comparison with peers and attaining a higher self-concept (Chiu and Klassen, 2008). Conversely, students in collectivist cultures seek upward comparison and consequently have a lower self-concept (termed as modesty bias; see White and Lehman, 2005). Where culture and ethnicity is concerned, however, interesting differences in mathematics self-concept have been found at national level (as opposed to student level) and between cultures.

Comparing national differences in the relationship between mathematics self-concept and achievement, Wilkins (2004) found that at the student level, higher mathematics self-concept was associated with higher mathematics achievement. At a national level, however, countries ranked with higher levels of achievement (e.g., Asian and East European countries) typically demonstrated lower mathematics self-concept. In contrast, countries ranked with comparatively lower achievement (e.g., Western Europe, the US, and Australia) registered higher mathematics self-concept. Students in individualistic societies (e.g., the US, Australia and New Zealand) may boost their mathematics self-concept by downward comparison with peers (Chiu and Klassen, 2008), perhaps explaining how higher self-concept was found in conjunction with lower achievement in mathematics. In contrast, students in collectivist cultures (e.g., East Asian cultures) have been socialized to value modesty and upwardly compare with peers in terms of mathematics self-concept, possibly explaining why, for them, a lower self-concept and higher achievement in mathematics are found simultaneously (White and Lehman, 2005). It could be suggested then that the culture of students may moderate the relationship between mathematics self-concept and performance.

In New Zealand, Caygill et al. (2013) found a positive association between attitude and achievement in mathematics, and mathematics self-confidence (linked to mathematics self-concept, see Reyes, 1984) was the strongest predictor of achievement for students within and between the nation's major ethnicities (New Zealand European, Māori, Pacific Island, and Asian). Interventions to raise the mathematics achievement of New Zealand primary school students by improving their mathematics self-beliefs (see Bonne, 2016) have borne positive results. Research has been called for, however, that engages in conducting such explorations longitudinally (Bonne, 2016).

Factors That Influence Mathematics Self-Concept

Socialization of Stereotypical Gender Beliefs

Socialization comprises interactions with the environment and significant others, including teachers and peers (Marsh, 1990a). Further, an awareness of cultural norms develops (Schoon, 2015) as children form an understanding of roles and behaviors that are appropriate for them through social learning (see Bandura, 2002). Specifically, in the New Zealand context, women have been associated in television advertisements with home-making and glamorous appearance, whereas men have been portrayed as physically and mentally tough, and technically agentic (Michelle, 2012). Such evidence provides support for the suggestion that New Zealand is a gender-essentialist society where expectations of conformity to gender roles is a powerful shaper of choices for young people (Cushman, 2008). Stereotypes can also reinforce relationships between ethnicity and gender. Māori women have been historically stereotyped as unsuitable candidates for mathematics-related fields in post-colonial New Zealand (McKinley, 2008). Further, although New Zealand European women have media dominance among New Zealand ethnic-racial groups, they are most frequently portrayed in roles that negate their agency in scientific and technical fields (Michelle, 2012).

Awareness of stereotypes and popularly held beliefs specific to perceptions of mathematics ability emerge in children between the ages of 7 and 12 years (Steele, 2003). Moreover, elementary school children are particularly receptive to stereotypical information (Gunderson et al., 2012). Thus, primary and middle school children's concept of their mathematical selves has been potentially influenced by a stereotype in which mathematics has been historically associated with masculinity and purports the mathematical superiority of males (Li, 1999; Neugebauer et al., 2011).

Importantly, when stereotypes are activated, they have been known to be central in influencing how information is processed and subsequently affect behavior (Heyder and Kessels, 2016) confirming the underpinning stereotype (Hamilton et al., 1990). Further, gender stereotypes strengthen with age (see Martinot and Désert, 2007). One could suggest, therefore, that implications are presented for the future selves of women and particularly women from ethnic-racial minorities where stereotypes purporting the suitability of males and dominant ethnic-racial groups for STEM fields prevail.

Social Comparison

As mentioned above, social comparison (e.g., comparing performance or attainment with that of peers) influences self-concept (O'Mara et al., 2006). Where students are grouped according to ability, social comparison is significant in shaping students' conceptualization of their own ability because of the opportunity to upwardly or downwardly compare (Burleson et al., 2005). Although New Zealand elementary classes are most frequently non-streamed (untracked), within-class ability grouping is endemic (Rubie-Davies, 2015). New Zealand elementary students have the opportunity then, to compare their ability upward and downward with that of their classroom peers. Self-concept can be boosted as a result of downward comparison but reduced by upward comparison (Burleson et al., 2005). Specifically, comparison with peers who have superior mathematics achievement is likely to promote a relatively negative mathematics self-concept (Bong, 1998).

The Specific Impact of Teachers

Teachers play an important role in shaping student self-concept. Teachers' expectations of their students' success (explained in more detail below), for example, is a central influence on student self-perceptions of ability and competence (Harris and Rosenthal, 1985). Specifically, the ability judgments of significant others (e.g., teachers) have held especial potential to impact mathematics self-concept (Dickhäuser and Stiensmeier-Pelster, 2003) and notably so for elementary school children (Marsh et al., 1998). Further, girls have relied on a perception of teacher evaluations of their mathematics ability to a greater extent than boys in forming attributions of their mathematics ability (Dickhäuser and Meyer, 2006). Erdogan and Şengul (2014) stressed the roles of quality of instruction and the classroom climate as central in influencing the development of students' mathematics self-concept.

Teachers' stereotypical beliefs about mathematics can influence students' perceptions of their own mathematical ability (Tiedemann, 2000). Moreover, teacher expectations have mediated the relationship between such beliefs and student outcomes (Younger and Warrington, 2008). Nevertheless, although negative beliefs (e.g., about the gender-appropriateness of certain subject domains) can result in a non-adaptive interplay between beliefs and performance (Sansone, 2016), self-perceptions are malleable (Harackiewicz et al., 2012). Further, in as much as teachers have conveyed gendered beliefs about mathematics (Gunderson et al., 2012), they also have the power to strengthen students' self-beliefs (Siegle and McCoach, 2007; Bonne, 2016). Importantly, Sansone (2016) suggested that a detrimental self-fulfilling prophecy caused by gender beliefs about mathematics could be disrupted by an improvement in teacher-student relationships and an enhanced climate of safety at school.

Teacher Expectations of Students' Academic Outcomes

Differential teacher expectations have been identified as one of several factors responsible for differences in students' academic outcomes. Specifically, teacher expectations have resulted in their differential treatment of students (Brophy and Good, 1970), and consequent differences in student outcomes including the shaping of self-concept and motivation (Harris and Rosenthal, 1985). Page and Rosenthal (1990) suggested that teacher beliefs about student gender and race could influence teachers' beliefs about student potential in quantitative endeavors. The expression of teaching behaviors associated with increased learning opportunities could thus result from teachers' beliefs of certain groups' superior aptitude and efficacy in mathematics (i.e., Asian and male students, see Page and Rosenthal, 1990). Student beliefs about their own STEM-field ability and competence, a heightened susceptibility to biased teacher expectations, and ultimately student academic outcomes differentiated by gender could, therefore, be results of endorsing the mathematics stereotype (Eccles and Wigfield, 1985).

It could be suggested then that teacher expectations based on stereotypical gender beliefs (founded on inaccuracy) have the potential to become self-fulfilling prophecies for students and promote inequity. Importantly, in New Zealand, teachers' expectations of student academic success have been shown to be influenced by student ethnicity with less success expected for Māori students (Rubie-Davies et al., 2006). Messages conveyed by teachers via such expectations have been known to trigger stereotype threat (see Steele and Aronson, 1995), one outcome of which is dis-identification with specific domains (Woodcock et al., 2012).

Teacher expectations have been identified at both individual and class level, but were found to be most influential when applied to the whole class (Rubie-Davies, 2007). Further, clear distinctions can be drawn between teachers who have high or low expectations for *all* their students (Weinstein, 2002; Rubie-Davies and Peterson, 2011). Rubie-Davies and Peterson (2011) found that certain characteristics were associated with teachers who had high expectations for all the students in their class (e.g., use of mixed ability grouping, choice in learning experiences, positive social climate in the classroom, intrinsic motivation, and well-defined goals). In contrast, the characteristics associated with teachers who had low expectations for all students in their classes (e.g., use of ability groups, teacher-determined learning experiences, negative social climate in the classroom, extrinsic motivation, and uncertainty of learning direction) differed starkly from those of their high expectation colleagues.

The Current Study

A need exists for longitudinal research regarding mathematics self-beliefs and more so given the ability of teachers to alter students' self-beliefs via changes in pedagogy (Bonne, 2016). Whether changes in mathematics self-concept occurred over time was explored in the current study. The aforementioned research was conducted with New Zealand elementary school students during a 3-year longitudinal study evaluating an intervention to raise and sustain teacher expectations of student achievement in mathematics. Positive benefits of the intervention were reported for student academic outcomes (Rubie-Davies and Rosenthal, 2016) and success in modeling the practices of high expectation teachers (McDonald et al., 2014; Rubie-Davies et al., 2015). Further, the consequences of inaccurate teacher expectations for gifted readers (Garrett et al., 2015), the relationship between student ethnicity and teacher expectations (Rubie-Davies et al., 2013), and the influence of teacher gender on teacher expectations in mathematics and reading (Watson et al., 2016, 2017) have already been explored. The current study marked a starting point in exploring whether the intervention was related to changes in student self-beliefs. Specifically, student mathematics self-concept data were analyzed via multilevel modeling.

Previous research had found that teacher expectations could influence student self-beliefs (Harris and Rosenthal, 1985; Younger and Warrington, 2008) and that self-perceptions and expectations of success were malleable (Harackiewicz et al., 2012). On the basis of such research, we hypothesized that within the context of a 3-year longitudinal intervention focused on the development of beliefs and practices associated with teachers' high expectations for all their students, an expected decline in student mathematics self-concept would be ameliorated

(H1), a gender gap in mathematics self-concept favoring boys would be addressed (H2), and equitable mathematics self-concepts between students in different ethnic groups would be promoted (H3).

MATERIALS AND METHODS

The current study draws on data collected across 3 years of a longitudinal study that investigated teacher expectations of elementary school students to determine whether expectations could be raised and sustained and whether or not there were effects on student achievement outcomes (see Rubie-Davies et al., 2015). Socio-demographic factors (i.e., age, gender, and ethnicity) associated with different student self-reported levels of self-concept for mathematics were explored over time (to show any changes associated with an increase in age).

Participants

Following ethical approval from the authors' institutional ethics committee, 11 New Zealand elementary schools comprising a range of socioeconomic (SES) levels were recruited for the study. In New Zealand, school-based SES is based on a 10-point scale with 1 assigned to the poorest schools and 10 to the most affluent. Informed and written consent was gathered from parents/legal guardians in order to invite students to participate in the research. The current study comprised student participants (n_{baseline} = 1,739) drawn from the above schools. As the focus of the study was to report student outcomes, student, rather than school-based participant demographics, are presented (see **Table 1**). The proportions of male and female students are consistent with those found in the New Zealand scholastic population. The range and proportion of ethnicities comprising the current participant sample were similar to that represented in the school-aged population of the large urban area in which the study was conducted (see **Table 1**).

Measures

In addition to the sociodemographic factors explored, two main measures (described below) were used in this study, one to measure student self-concept in mathematics and one to measure student achievement. Descriptive statistics for the achievement and self-concept measures are shown in **Table 2**. Achievement and self-concept were both measured at the beginning and end of each of the 3 years. To account for the differing gaps between each time point, this was recoded into the number of school terms after baseline, with baseline entered as 0. There are four terms per academic year in New Zealand.

TABLE 1 | Participant demographic information (student *n* at baseline = 1,739; school *n* = 11).

Gender	**Male**	**Female**				
	49.8%	50.2%				
Ethnicity	**NZ European**	**Māori**	**Pacific Island**	**Asian**	**Other**	
	47.4%	17.8%	16.2%	15.9%	2.7%	
Age	**7 years**	**8 years**	**9 years**	**10 years**	**11 years**	**12 years**
	8.1%	18.3%	19.2%	19.3%	17.5%	17.5%

TABLE 2 | Descriptive statistics for mathematics self concept and mathematics achievement.

	Time	***n***	**Mean**	**Std. Deviation**
Maths self concept	BOY 1[1]	1739	3.66	1.03
	EOY 1	1015	3.75	0.97
	BOY 2	672	3.68	0.96
	EOY 2	526	3.64	0.98
	BOY 3	266	3.88	0.93
	EOY 3	242	3.79	0.86
	Total	4460	3.70	0.98
Maths achievement	BOY 1	1739	−0.064	0.786
	EOY 1	1015	0.101	0.867
	BOY 2	672	−0.045	0.755
	EOY 2	526	−0.057	0.764
	BOY 3	266	−0.231	0.701
	EOY 3	242	−0.040	0.659
	Total	4460	−0.032	0.791

[1] *BOY, beginning of year; EOY, end of year.*

Self-Concept for Mathematics

Mathematics self-concept was measured using the mathematics self-concept subscale of the Self-Description Questionnaire-1 (SDQ-1; see Marsh, 1990b). Students responded to items (e.g., "Work in mathematics is easy for me") on a Likert scale ranging from 1 to 5, where 1 = false; 2 = mostly false; 3 = sometimes false; sometimes true; 4 = mainly true; 5 = true. This scale was standardized before entry into the multilevel models in order to provide standardized coefficients. The alpha coefficient for the subscale (α = 0.93) gave confidence in using data generated by it in further analyses.

Student Achievement

Student achievement was measured using a standardized tool, e-asTTle mathematics (e-asTTle Project Team, 2009). National norming trials were conducted in 2009–2010 in order to establish the robustness of e-asTTle. All items in the e-asTTle system were calibrated using item response theory (IRT) scoring procedures (Embretson and Reise, 2000) and the standard error of measurement for any e-asTTle test was estimated to be 15 points, with a standard deviation of 100 for each year level (see e-asTTle website)[1]. The test-retest reliability of e-asTTle is reported to be α = 0.96 (Ministry of Education, and NZCER, 2012). Thus, confidence in test consistency was assured. Confidence was also assured in comparing student scores in relation to the normative expectation for their academic year level, regardless of the test level administered to students.

[1] http://e-asttle.tki.org.nz/

Because expected scores differ by year level, the published normative expectation for each year level was subtracted from each student's overall total mathematics score, generating a "centered" score that indicated the distance from the norm, while retaining the metric of measurement. The average achievement in the current sample was approximately equivalent to that of the normative sample at baseline ($M = -6.47$, $SD = 78.65$, $N = 1739$).

This centering process shifts the mean but does not affect the standard deviation. Therefore, the sample variability was somewhat less than for the e-asTTle normative sample. To place e-asTTle on a comparable scale to the other variables, the scores were divided by the standard deviation for the tool, as recommended by Schagen and Hodgen (2009).

Procedures and Missing Data

Data collection for this study occurred at the beginning and end of each academic year over a 3-year period commencing in February 2011. At each data collection point, student mathematics achievement and self-concept were measured using e-asTTle mathematics and the mathematics self-concept subscale of the SDQ-1, as described above. Paper and pencil versions of the questionnaire containing the SDQ-1 were administered by a researcher and research assistant to each class during the school day in their own classroom. Worksheets were provided for any student who opted not to participate, but were not needed in all but three cases. Grade–normed tests had been designed at each curriculum level within e-asTTle mathematics (e.g., Level 2, Level 2/3, Level 3, Level 3/4, Level 4, and Level 4/5), and were chosen by each teacher as was appropriate for the number of students per approximate levels of achievement of their class.

However, not all students provided data at each time point. Sometimes this was due to illness on the day of the data collection, but there were also structural issues that precluded data collection across the full 3-year period in some cases. For example, students in New Zealand often change schools at the end of Year 6 (~11 years old) and most change again in Year 8 (~13 years old). Although HLM procedures are able to incorporate complete case information, whereby a single missing time point is not a major concern, significant attrition over time can still be problematic. To reduce the possibility that our results would simply reflect students dropping out of the sample, we only included students for whom baseline results were available, and who were present for at least a complete year. We also evaluated the student achievement scores of the students included in the final sample, against those who were excluded, and found no statistically significant differences by student achievement ($M_{\text{diff}} = 0.05$ SD, $p = 0.087$). Analyses were carried out using IBM (2016) SPSS v 24.0 and MLwiN v 3.02 (Charlton et al., 2017).

Analytical Procedures

The primary aim of the current study was to investigate whether, within the context of an intervention to raise and sustain teacher expectations of student achievement, an expected decline in student mathematics self-concept with age and gaps in the construct's levels by gender and ethnicity would be observed. To examine this aim, three-level hierarchical linear models (HLM) were built. These models explored the association over time between sociodemographic factors (age, gender, and ethnicity) and standardized self-concept in mathematics (the dependent variable), controlling for measured student achievement in mathematics. HLM is a multilevel regression framework that allows dependency in the data to be considered. Students within a particular classroom, or attending a particular school, tend to have a degree of shared variance, meaning that educational data generally violate the independence assumption (Osborne, 2000). Within the HLM framework, students in the same school can be specified as being nested within schools, for example, with students at level one and schools at level two. As the current study included a time component, observations were treated as level one nested within students at level two, and schools treated as level three. An alternative specification could have included teachers as level three, but when an unconditional model was specified including the teacher level, the variance at this level was negligible (1.4%). Given the extra complexity of adding this level (particularly because students change teachers each year), the teacher level was not included in the final models.

As mathematics achievement had already been "centered" against the normative expectations, achievement was not centered further when added to the HLM. As noted above, these scores were standardized by dividing by the standard deviation for the tool. Ethnicity was entered as a polytomous categorical variable, with Māori, Pasifika (of Pacific Island origin), and Asian students treated as binary dummy variables, and New Zealand European students treated as the reference category because this was the largest group. Student gender was included as a dummy variable with boys treated as the reference category—in this case because we wanted to focus on the relations for female students. Interaction terms (see Aguinis et al., 2013) were also included to explore the change in self-concept over time for each group.

To control for individual differences in achievement at each time point, mathematics achievement was added to the model along with the variable for time. This procedure was carried out after fitting the three-level unconditional random intercepts model with mathematics self-concept as the dependent variable (see **Table 3** for the final model). Age was also included in this model because prior research has indicated that self-concept tends to decline with age. This was followed by the variables of gender and student ethnicity as well as the interactions with time. Random slopes were tested for the effect of student achievement on self-concept but this did not improve the model fit so only random intercepts were used in the final model.

RESULTS

On average, self-concept was relatively high across the 3-year period ($M = 3.7$, see **Table 2**). The results of the unconditional and final multilevel models are shown in **Table 3** below. The change in deviance was 185.41 indicating that the final model was a statistically significantly better fit for the data than the null model. Using the method described by Bryk and Raudenbush (1992) to

TABLE 3 | Multilevel model exploring socio-demographic predictors of standardized mathematics self-concept within a New Zealand elementary student sample.

	Unconditional model			Final model		
Parameter	**β**	***SE***	***p***	**β**	***SE***	***p***
Fixed effects						
Intercept	0.007	0.069	<0.001	−0.101	0.067	0.132
Time				0.029	0.007	<0.001
e-asTTle Maths vs. Norm†				0.119	0.015	<0.001
Age (grand-mean centered)				−0.048	0.015	0.001
Gender (ref. = Male)						
Female				−0.197	0.043	<0.001
Ethnicity (ref. = European)						
Māori				0.209	0.062	<0.001
Pasifika				0.354	0.067	<0.001
Asian				0.448	0.064	<0.001
Other ethnicity				0.567	0.136	<0.001
Interaction terms						
Female student × time				−0.012	0.007	0.087
Māori student × time				−0.023	0.01	0.02
Pasifika student × time				−0.012	0.01	0.23
Asian student × time				−0.016	0.011	0.146
Other ethnicity × time				−0.038	0.023	0.098
Between school variance (σ^2_{v0})	0.056 [6.07%]	0.025	0.034	0.034 [3.7%]	0.016	0.033
Between student variance (σ^2_{u0})	0.542 [58.7%]	0.025	<0.001	0.461 [49.9%]	0.023	<0.001
Repeated measures variance (σ^2_{e})	0.423 [45.8%]	0.012	<0.001	0.428 [46.3%]	0.012	<0.001
−2*log likelihood	11,301.33			11,115.92		

†*e-asTTle scores were divided by 100, so parameter estimates reflect expected difference per 100 e-asTTle points. VPC shown in square brackets.*

approximate the proportion of variance explained by the final model, we found that the model explained approximately 39.3% of the school-level variance, and 14.9% of the student-level variance. There was no reduction in level-one variance.

The results indicated that higher mathematics achievement was statistically significantly associated with higher self-concept for mathematics ($\beta = 0.119$, $p < 0.001$). Several other covariates were associated with self-concept, even after controlling for achievement differences (see **Table 3**).

Mathematics self-concept was negatively, and statistically significantly associated with age at each time point, with older students reporting significantly lower self-concept than younger students (−0.048 per year of age, $p = 0.001$; see **Table 3**). Surprisingly, however, there was a statistically significant increase over time within the current study (0.029 per time point, $p < 0.001$), for the same students, despite a 3-year increase in age. Female students had statistically significantly lower self-concept for mathematics on average (−0.197, $p < 0.001$), and this gap remained throughout the duration of the study. Controlling for achievement, self-concept for mathematics was lowest among New Zealand European students, with all other ethnicities having statistically significantly higher self-concept at baseline. However, the reported self-concept for New Zealand European students (the reference group for ethnicity) increased more over time than that for all other ethnicities (see the interaction terms in **Table 3**); though the relative decline was statistically significant only among Māori students during the 3-year period (−0.023 per time point, $p = 0.02$).

DISCUSSION

We expected that we would see the typical decline in student mathematics self-concept ameliorated within the context of an intervention focused on the development of beliefs and practices associated with teachers' high expectations for all their students (H1). We also hypothesized that the intervention would address a gender gap in mathematics self-concept favoring boys (H2) and would promote equitable mathematics self-concepts in students across different ethnic groups (H3). The well-evidenced age-related decline in mathematics self-concept appeared to be diminished over the course of the 3-year period, supporting (with reservations that are described below) the first hypothesis. Equitable levels of mathematics self-concept were not achieved, however, for gender or across ethnic groups resulting in a lack of support for the second and third hypotheses.

Although older students' mathematics self-concepts were lower at baseline than younger students', with statistical significance, all students, demonstrated a small increase in mathematics self-concept over the course of the longitudinal study. We would have expected the age-related decline reported in prior research (e.g., Wilkins, 2004) to have occurred, but it did not. Thus, we suggest that the pedagogical practices associated with teachers having high expectations of all their students appeared to be positively associated with an amelioration of the expected decline.

Mathematics self-concept is not only shaped by messages received from influential others, for example, teachers

(Dickhäuser and Stiensmeier-Pelster, 2003), but also social comparison with peers (Parker et al., 2014). Peer comparison acts to diminish levels of mathematics self-concept as age advances (Wilkins, 2004). Pedagogical elements of the intervention promoted positive messages from teachers (as influential others) and created classroom climates where social comparison with peers was discouraged. In support of this idea, McDonald et al. (2014) reported that teachers' responses regarding the implementation of the current intervention acknowledged increased student self-belief and self-confidence. Teachers indicated that mixed ability grouping, promoting choice of learning activity, facilitation of a collaborative rather than competitive class climate, teacher positivity, and a buddy-system where students encouraged peer responsibility for the classroom climate fostered ownership of student learning and enabled the celebration of success for each child (McDonald et al., 2014). These changes to teacher beliefs and practices, taught during the intervention, may have created a learning environment that lessened the effect of social comparison as a source of age-related decline in mathematics self-concept.

Although the intervention appeared to be positively associated with student mathematics self-concept overall, neither the gender gap in mathematics self-concept, which favored boys, nor differences in the construct between ethnic groups, which positioned Māori in New Zealand at a particular disadvantage, were altered. Despite finding that mathematics self-concept rose slightly for students overall, girls' mathematics self-concept was comparatively lower than boys', and this gap persisted despite the intervention. The presence of a gender gap in mathematics self-concept supported previous New Zealand findings (e.g., Bonne, 2016) yet suggested that further changes to teacher beliefs and practices (specifically regarding gender) over and above those made during the current study were necessary. Importantly, in terms of gender equity, attention to mathematics achievement will not promote girls' trajectories toward STEM pathways and careers without an accompanying self-concept that bolsters belief of success (Goldman and Penner, 2014).

Stereotypical notions of mathematics ability for girls have been reinforced by teachers with negative outcomes and are strongly related to girls' mathematics self-concept (Ertl et al., 2017). Further, self-concept relies on the part of one's perception of one's self that is activated (Kessels and Hannover, 2008). Given the aforementioned research, it could be suggested that Kessels and Hannovers' (2008) advocacy for reducing gender salience in class climates, should be considered. Thus, raising teacher awareness of gender salience in classrooms seems an important adjunct to challenging teacher beliefs and practices. Such a step may reduce the accessibility of gender-related self-descriptions and, thus, the impact of gender-related stereotypes, with the potential to impact mathematics self-concept (see Kessels and Hannover, 2008). Teachers (and parents) can help girls to build their self-beliefs and confidence in mathematics by realistic evaluation of girls' actual abilities rather than recourse to their own beliefs about girls' potential in the subject (OECD, 2015). Further, girls' mathematics self-beliefs could be bolstered by feedback to combat cultural messages that privilege males in relation to mathematics (Correll, 2001).

The greatest gaps in mathematics self-concept have been found in countries with the most pronounced norms of gender-role rigidity and New Zealand has been identified among these (Hofstede, 2003). Further, New Zealand society has been described as gender-essentialist, that is, clearly demarking masculine and feminine attributes and roles as discreet from and opposite to each other in nature (Cushman, 2008). It seems unsurprising, therefore, that a gender gap in mathematics self-concept persisted in the current data. Training teachers to identify their own gender biases (see OECD, 2015) could further enhance interventions to raise teacher expectations. Such actions could ensure that not only enhanced student achievement but, critically (and with important implications for girls' future capacity for STEM involvement), equitable self-concept in mathematics, were addressed.

Levels of mathematics self-concept for New Zealand European, Asian, and Pacific Island students rose slightly within the context of the intervention (after controlling for achievement differences). Māori students' mathematics self-concept, however, continued to decline across the 3 years of the current study. We might have expected that Māori students, being from a collectivist cultural background (Harrington and Liu, 2002), would have been favorably affected by the promotion of a collaborative rather than competitive class climate, taught to teachers during the intervention. Rothstein-Fisch et al. (1999) found that within a collectivist context, one child's mathematics success became success for the whole group, that the whole group celebrated individual success, and that these student behaviors were promoted in mixed-ability classrooms. Yet, although Asian students' mathematics self-concept grew, that of Māori students diminished.

Previous research (Rubie-Davies et al., 2006) has found that teachers have expected less of Māori students. Importantly, teachers have placed the onus for such expectations on deficits attributed to the Māori students themselves and their cultural background (Bishop et al., 2010). Bishop et al. (2010) described *whanaungatanga* (the warmth and closeness associated with the extended family) reflected in relationships within the classroom, and especially those between the student and teacher, as key to Māori students' success. Notably, Māori students wished to be acknowledged positively by their teachers as Māori (Bishop et al., 2010). The centrality of successful teacher-student relationships for positive student outcomes (Hattie, 2009) was a feature of the current intervention, and here again we would have expected this to have been associated with positive outcomes for all students. In line with the suggestions of Rubie-Davies and Peterson (2016), we suggest that specific attention seems needed to further raise teachers' awareness of culturally responsive pedagogy uniquely tailored to the needs of Māori students.

Limitations and Directions for Future Research

The current study was not conclusively able to deduce that it was the intervention and not confounding elements that were related to the amelioration of the decline in mathematics self-concept over time. Future research could test the effectiveness of pedagogical practices deemed likely to have reduced, for example, the social comparison often held responsible for the age-related decline in the construct. Further, the current sample suffered from significant attrition and comprised only participants from one urban center, which may reduce generalizability of the findings. Although attrition would be hard to address given the schooling structure of New Zealand, other urban locations and rural settings could be investigated in the future. As well, the current intervention did not set out to address mathematics self-concept as such, but the findings suggest that further attention to nuanced teacher beliefs encompassing gender and ethnicity, could be valuable in future iterations. Additional training to help teachers recognize and confront biases they may hold about different groups of students (e.g., in terms of gender, and ethnicity) may aid teacher effectiveness and further enhance student potential (OECD, 2015).

Conclusion

The current paper aimed to explore whether, within the context of an intervention to raise and sustain teacher expectations, the amelioration of an age-related decline and a gender gap in mathematics self-concept would be observed and equitable levels of mathematics self-concept across ethnic groups would be found. The finding that mathematics self-concept over time improved (where a decline had been expected) suggests that comprehensively addressing social comparison issues within class climates may be adaptive in fostering positive levels of the construct. Nevertheless, findings that a gender gap favoring boys persisted and a notable decline occurred for Māori students suggest the need for specific interventions to address gender-related issues and to nurture Māori students' perceptions of themselves as mathematicians. Notably, the findings point to the importance of considering context in the study of mathematics self-concept in order to successfully implement future interventions.

ETHICS STATEMENT

Permission to proceed with the current research was granted by the University of Auckland Human Participants Ethics Committee. Consent to invite teacher and student participation was gained from all participating school Principals. All participating teachers fromalized their consent, and their students (all being under the age of 16) could assent to participation once their parents had given consent that they may be invited to take part in the study. Anonymity and confidentiality regarding all participants and participating schools, were rigorously preserved.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

The current study was made possible by the awarding of a Marsden Fast Start Grant and a grant from the Cognition Institute to CR-D.

REFERENCES

Aguinis, H., Gottfredson, R. K., and Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *J. Manag.* 39, 1490–1528. doi: 10.1177/0149206313478188

Anderman, E. M., and Maehr, M. L. (1994). Motivation and schooling in the middle grades. *Rev. Educ. Res.* 64, 287–309. doi: 10.3102/00346543064002287

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: Freeman.

Bandura, A. (2002). "Social cognitive theory of mass communication," in *Media Effects: Advances in Theory and Research*, eds J. Bryant, and D. Zillmann, (Mahwah, NJ: Erlbaum), 121–154.

Bishop, R., Ladwig, J., and Berryman, M. (2010). The centrality of relationships forpedagogy: the whanaungatanga thesis. *Am. Educ. Res. J.* 51, 184–214. doi: 10.3102/0002831213510019

Bong, M. (1998). Tests of the internal/external frames of reference model with subject-specific academic self-efficacy and frame-specific academic self-concepts. *J. Educ. Psychol.* 90, 102–110. doi: 10.1037/0022-0663.90.1.102

Bong, M., and Clark, R. E. (1999). Comparison between self-concept and self-efficacy inacademic motivation research. *Educ. Psychol.* 34, 139–153. doi: 10.1207/s15326985ep3403_1

Bong, M., and Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: how different are they really? *Educ. Psychol. Rev.* 15, 1–40. doi: 10.1023/A:1021302408382

Bonne, L. (2016). New Zealand Students' mathematics-related beliefs and attitudes: recent evidence. *N. Z. J. Educ. Stud.* 51, 69–82. doi: 10.1007/s40841-016-0035-2

Brophy, J. E., and Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: some behavioral data. *J. Educ. Psychol.* 61, 365–374. doi: 10.1037/h0029908

Bryk, A. S., and Raudenbush, S. W. (1992). *Hierarchical Linear Models for Social and Behavioural Research: Applications and Data Analysis Methods*. Newbury Park, CA: Sage Publications.

Burleson, K. P., Leach, C. W., and Harrington, D. (2005). Upward social comparison and self-concept: inspiration and inferiority among art students in an advanced program. *Br. J. Soc. Psychol.* 44, 109–123. doi: 10.1348/014466604X23509

Byrne, B. M., and Shavelson, R. J. (1986). On the structure of adolescent self-concept. *J. Educ. Psychol.* 78, 474–481. doi: 10.1037/0022-0663.78.6.474

Caygill, R., Kirkham, S., and Marshall, N. (2013). *Year 9 Students' Mathematics Achievement in 2010/11: New Zealand Results From the Trends in International Mathematics and Science Study (TIMSS)*. Wellington: Ministry of Education.

Ceci, S. J., and Williams, W. M. (2009). *The Mathematics of Sex: How Biology and Society Conspire to Limit Talented Women and Girls*. New York, NY: Oxford University Press.

Chapman, J., Tunmer, W. E., and Prochnow, J. (2000). Early reading-related skills and performance, reading self-concept, and the development of academic self-concept: a longitudinal study. *J. Educ. Psychol.* 92, 703–708. doi: 10.1037/0022-0663.92.4.703

Charlton, C., Rasbash, J., Browne, W. J., Healy, M., and Cameron, B. (2017). *MLwiN Version 3.00*. Bristol: University of Bristol.

Chiu, M., and Klassen, R. M. (2008). Relations of mathematics self-concept and its calibration with mathematics achievement: cultural differences among fifteen-year-olds in 34 countries. *Learn. Instr.* 20, 2–17. doi: 10.1016/j.learninstruc.2008.11.002

Correll, S. (2001). Gender and the career choice process: the role of biased self-assessments. *Am. J. f Sociol.* 106, 1691–1730. doi: 10.1086/321299

Cushman, P. (2008). So what exactly do you want? What principals mean when they say 'male role model'. *Gen. nd Educ.* 20, 123–136. doi: 10.1080/09540250701805847

Dickhäuser, O., and Meyer, W.-U. (2006). Gender differences in young children's math ability attributions. *Psychol. Sci.* 48, 3–16.

Dickhäuser, O., and Stiensmeier-Pelster, J. (2003). Wahrgenommene lehrereinschatzungen und das fahigkeitsselbstkonzept von jungen und madchen in der grundschule. [Perceived teachers' ability evaluations and boys' and girls' concepts of their mathematical ability in elementary school]. *Psychologie in Erziehung und Unterricht* 50, 182–190.

Dusek, J. B., and Joseph, G. (1985). "The bases of teacher expectancies," in *Teacher Expectancies*, ed. J. B. Dusek, (Hillsdale, NJ: Lawrence Erlbaum Associates), 229–250.

e-asTTle Project Team, (2009). *e-asTTle Generation 2: e-asTTle Educator Manual*. Wellington: Ministry of Education.

Eccles, J., Midgley, C., Wigfield, A., Buchanan, C., Reuman, D., Flanagan, C., et al. (1993). Development during adolescence: the impact of stage–environment fit on young adolescents' experiences in schools and families. *Am. Psychol.* 48, 90–101. doi: 10.1037/0003-066x.48.2.90

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychol. Women Q.* 11, 135–172. doi: 10.1111/j.1471-6402.1987.tb00781.x

Eccles, J. S., and Wigfield, A. (1985). "Teacher expectations and student motivation," in *Teacher Expectancies*, ed. J. B. Dusek, (Hillsdale, NJ: Lawrence Erlbaum), 185–228.

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Embretson, S. E., and Reise, S. (2000). *Item Response Theory for Psychologists*. Mahwah, NJ: Erlbaum.

Erdogan, F., and Şengul, S. (2014). A study on the elementary school students' mathematics self-concept. *Procedia Soc. Behav. Sci.* 152, 596–601. doi: 10.1016/j.sbspro.2014.09.249

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Franken, R. E. (1994). *Human Motivation*, 3rd Edn. Pacific Grove, CA: Brooks/ColePublishing Co.

Fredricks, J. A., and Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence. *Dev. Psychol.* 38, 519–533. doi: 10.1037//0012-1649.38.4.519

Ganley, C. M., and Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: examining gender patterns and reciprocal relations. *Learn. Individ. Differ.* 47, 182–193. doi: 10.1016/j.lindif.2016.01.002

Garrett, L., Rubie-Davies, C. M., Alansari, M., Peterson, F., Flint, A., Watson, P., et al. (2015). Missing out'? The potential consequences of inaccurate teacher expectations on young gifted readers' achievement outcomes. *Apex N. Z. J. Gift. Educ.* 19, 31–45. doi: 10.21307/apex-2015-005

Goldman, A. D., and Penner, A. M. (2014). Exploring international gender differences in mathematics self-concept. *Int. J. Adoles. Youth* 21, 403–418. doi: 10.1080/02673843.2013.847850

Gunderson, E. A., Ramirez, G., Levine, S. C., and Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles* 66, 153–166. doi: 10.1007/s11199-011-9996-2

Hamilton, D. L., Sherman, S. J., and Ruvolo, C. M. (1990). Stereotype-based expectancies:effects on information processing and social behavior. *J. Soc. Issues* 46, 35–60. doi: 10.1111/j.1540-4560.1990.tb01922.x

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., and Hyde, J. S. (2012). Helping parents to motivate adolescents in mathematics and science: an experimental test of a utility–value intervention. *Psychol. Sci.* 23, 899–906. doi: 10.1177/0956797611435530

Harrington, L., and Liu, J. H. (2002). Self enhancement and attitudes toward high achievers: a bicultural view of the independent and interdependent self. *J. Cross Cult. Psychol.* 33, 33–55.

Harris, M. J., and Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychol. Bull.* 97, 363–386. doi: 10.1037/0033-2909.97.3.363

Hattie, J. A. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London, UK: Routledge.

Heyder, A., and Kessels, U. (2016). Do teachers equate male and masculine with lower academic engagement? How students' gender enactment triggers gender stereotypes at school. *Soc. Psychol. Educ.* 18, 467–485. doi: 10.1007/s11218-015-9303-0

Hofstede, G. (2003). *Culture's Consequences*. Thousand Oaks, CA: Sage.

IBM, (2016). *SPSS for Windows v.24*. New York, NY: IBM.

Ireson, J., and Hallam, S. (2009). Academic self-concepts in adolescence: relations with achievement and ability grouping in schools. *Lear. Instr.* 19, 201–213. doi: 10.1016/j.learninstruc.2008.04.001

Jacobs, J. E. (2005). Twenty-five years of research on gender and ethnic differences in math and science career choices: What have we learned? *New Dir. Child Adolesc. Dev.* 110, 85–94. doi: 10.1002/cd.151

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., and Wigfield, A. (2002). Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Deve.* 73, 509–527. doi: 10.1111/1467-8624.00421

Kessels, U., and Hannover, B. (2008). When being a girl matters less: accessibility of gender-related self-knowledge in single-sex and coeducational classes and its impact on students' physics-related self-concept of ability. *Br. J. Educ. Psychol.* 78, 273–289. doi: 10.1348/000709907X215938

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: a review. *Educ. Res.* 41, 63–76. doi: 10.1080/0013188990410106

Lindberg, S. M., Hyde, J. S., Peterson, J. L., and Linn, M. C. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psycho. Bull.* 136, 1123–1135. doi: 10.1037/a0021276

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: preadolescence to early adulthood. *J. Educ. Psychol.* 81, 417–430. doi: 10.1037//0022-0663.81.3.417

Marsh, H. W. (1990a). A multidimensional, hierarchical model of self-concept: theoretical and empirical justification. *Educ. Psychol. Rev.* 2, 77–172. doi: 10.1007/bf01322177

Marsh, H. W. (1990b). *Self Description Questionnaire - 1 manual*. Campbelltown, NSW: University of Western Sydney.

Marsh, H. W., and Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *J. Educ. Psychol.* 95, 687–706. doi: 10.1037/0022-0663.95.4.687

Marsh, H. W., Craven, R., and Debus, R. (1998). Structure, stability, and development of young children's self-concepts: a multicohort–multioccasion study. *Child Dev.* 69, 1030–1053. doi: 10.2307/1132361

Marsh, H. W., and Scalas, L. F. (2011). "Self-concept in learning: reciprocal effects model between academic self-concept and academic achievement," in *Social and Emotional Aspects of Learning*, ed. S. Jarvela, (Oxford: Elsevier), 191–198.

Martinot, D., and Désert, M. (2007). Awareness of a gender stereotype, personal beliefs and self-perceptions regarding math ability: when boys do not surpass girls. *Soc. Psychol. Educ.* 10, 455–471. doi: 10.1007/s11218-007-9028-9

McDonald, L., Flint, A., Rubie-Davies, C. M., Peterson, E., Watson, P., and Garrett, L. (2014). Teaching high expectation strategies to teachers through an intervention process. *Prof. Dev. Educ.* 42, 290–307. doi: 10.1080/19415257.2014.980009

McKinley, E. (2008). From object to subject: hybrid identities of indigenous women in science. *Cult. Stud. Sci. Educ.* 3, 959–975. doi: 10.1007/s11422-008-9128-7

Michelle, C. (2012). Co-constructions of gender and ethnicity in New Zealand television advertising. *Sex Roles* 66, 21–37. doi: 10.1007/s11199-011-0067-5

Ministry of Education, and NZCER, (2012). *e-asTTle Technical Manual*. Wellington: Ministry of Education.

Nagy, G., Watt, H. M. G., Eccles, J. S., Trautwein, U., Lüdtke, O., and Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: different countries, different trajectories? *J. Res. Adolescence* 20, 482–506. doi: 10.1111/j.1532-7795.2010.00644.x

Neugebauer, M., Helbig, M., and Landmann, A. (2011). Unmasking the myth of the same-sex teacher advantage. *Eur. Sociol. Rev.* 27, 669–689. doi: 10.1093/esr/jcq038

OECD, (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence* PISA. Paris, France: OECD Publishing.

Oettingen, G., and Zosuls, K. M. (2006). "Culture and Self-Efficacy in Adolescents," In *Self-efficacy beliefs of adolescents*, eds F. Pajares, & T. Urdan, (Greenwich, CT: Information Age), 245–265.

O'Mara, A. I., Marsh, H. W., Craven, R. G., and Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educ. Psychol.* 41, 181–206. doi: 10.1027/s15326985ep4103_4

Osborne, J. W. (2000). Advantages of hierarchical linear modeling. *Practical Assess. Res. Eval.* 7, 1–3.

Page, S., and Rosenthal, R. (1990). Sex and expectations of teachers and sex and race of students as determinants of teaching behavior and student performance. *J. Sch. Psychol.* 28, 119–131. doi: 10.1016/0022-4405(90)90003-p

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., and Salah Abduljabbar, A. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educ. Psychol.* 34, 29–48. doi: 10.1080/01443410.2013.797339

Petersen, J., and Hyde, J. S. (2014). Gender-related academic and occupational interests andgoals. *Adv. Child Dev. Behav.* 47, 43–67.

Reyes, L. H. (1984). Affective variables and mathematics education. *Elem. Sch. J.* 84, 558–581. doi: 10.1086/461384

Rothstein-Fisch, C., Greenfield, P. M., and Trumbull, E. (1999). Bridging cultures with classroom strategies. *Educ. Leaders.* 56, 64–67.

Rubie-Davies, C. M. (2007). Classroom interactions: exploring the practices of high-and ow-expectation teachers. *Br. J. Educ. Psychol.* 77, 289–306. doi: 10.1348/000709906X101601

Rubie-Davies, C. M. (2008). "Teacher beliefs and expectations: relationships with student learning," in *Challenging Thinking About Teaching and Learning*, eds C. M. Rubie-Davies, and C. Rawlinson, (Hauppauge, NY: Nova.), 25–39.

Rubie-Davies, C. M. (2015). *Becoming a High Expectation Teacher: Raising the Bar*. Abingdon: Routledge.

Rubie-Davies, C. M., Hattie, J., and Hamilton, R. (2006). Expecting the best for students: teacher expectations and academic outcomes. *Br. J. Educ. Psychol.* 76, 429–444. doi: 10.1348/000709905X53589

Rubie-Davies, C. M., and Peterson, E. R. (2011). "Teacher expectations and beliefs: Influences on the socioemotional climate," in *Educational Psychology: Concepts, Research and Challenges*, ed. C. M. Rubie-Davies, (London: Routledge), 134–149.

Rubie-Davies, C. M., and Peterson, E. R. (2016). Relations between teachers' achievement over- and underestimation, and students' beliefs for Māori and Pākehā students. *Contemp. Educ. Psychol.* 47, 72–83. doi: 10.1016/j.cedpsych.2016.01.001

Rubie-Davies, C. M., Peterson, E. R., Flint, A., Garrett, L., McDonald, L., Watson, P., et al. (2013). Investigating teacher expectations by ethnicity in New Zealand. *Eur. J. Soc. Behav. Sci.* 2, 250–259. doi: 10.15405/FutureAcademy/ejsbs(2301-2218).2012.2.10

Rubie-Davies, C. M., Peterson, E. R., Sibley, C. G., and Rosenthal, R. (2015). A teacher expectation intervention: modelling the practices of high expectation teachers. *Contemp. Educ. Psychol.* 40, 72–85. doi: 10.1016/j.cedpsych.2014.03.003

Rubie-Davies, C. M., and Rosenthal, R. (2016). Intervening in teachers' expectations: a random effects meta-analytic approach to examining the effectiveness of an intervention. *Learn. Individ. Differ.* 50, 83–92. doi: 10.1016/j.lindif.2016.07.014

Sansone, D. (2016). "Teacher characteristics, student beliefs, and the gender gap in STEM fields," in *CeRP Working Papers165, Center for Research on Pensions and Welfare Policies*, (Turin).

Schagen, I., and Hodgen, E. (2009). *How Much Difference Does it make? Notes on Understanding, Using, and Calculating Effect Sizes for Schools*. Wellington: New Zealand Council for Educational Research.

Schoon, I. (2015). Explaining persisting gender inequalities in aspirations and attainment: an integrative developmental approach. *Int. J. Gend. Sci. Technol.* 7, 155–165.

Siegle, D., and McCoach, D. B. (2007). Increasing student mathematics self-efficacy through teacher training. *J. Adv. Acad.* 18, 278–312.

Steele, C. M., and Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *J. Personal. Soc. Psychol.* 69, 797–811. doi: 10.1037//0022-3514.69.5.797

Steele, J. (2003). Children's gender stereotypes about math: the role of stereotype stratification. *J. Appl. Soc. Psychol.* 33, 2587–2606. doi: 10.1111/j.1559-1816.2003.tb02782.x

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *J. Educ. Psychol.* 92, 144–151. doi: 10.1037/0022-0663.92.1.144

Wang, M.-T., and Kenny, J. L. (2014). Gender gap in Science, technology, engineering, and mathematics (STEM): current knowledge, implications for practice, policy, andfuture directions. *Educ. Psychol. Rev.* 29, 119–140. doi: 10.1007/s10648-015-9355-x

Watson, P., Rubie-Davies, C. M., Meissel, K., Peterson, E. R., Flint, A., Garret, L., et al. (2016). Gendered teacher expectations of mathematics achievement in New Zealand: contributing to a kink at the base of the STEM pipeline? *Int. J. Gend. Sci. Technol.* 8, 83–102.

Watson, P. W., St, J., Rubie-Davies, C. M., Meissel, K., Peterson, E. R., Flint, A., et al. (2017). Teacher gender, and expectation of reading achievement in New Zealand elementary school students: essentially a barrier? *Gend. Educ.* 1–20. doi: 10.1080/09540253.2017.1410108

Watt, H. M. G. (2010). "Gender and occupational choice," in *Handbook of Gender Research in Psychology*, eds J. C. Chrisler, and R. McCreary, (New York, NY: Springer), 379–400. doi: 10.1007/978-1-4419-1467-5_16

Weinstein, R. S. (2002). *Reaching Higher: The Power of Expectations in Schooling*. Cambridge, MA: Harvard University Press.

White, K., and Lehman, D. R. (2005). Culture and social comparison seeking: the role of self-motives. *Personal. Soc. Psychol. Bull.* 31, 232–242. doi: 10.1177/0146167204271326

Wigfield, A., Tonks, S., and Eccles, J. S. (2004). "Expectancy value theory in cross-cultural perspective," in *Big Theories Revisited*, eds D. M. McInerney, and S. van Etten, (Charlotte, NC: IAP), 165–198.

Wilkins, J. L. M. (2004). Mathematics and science self-concept: an international investigation. *J. Exp. Educ.* 72, 331–346. doi: 10.3200/jexe.72.4.331-346

Williams, J. E., and Best, D. L. (1990). *Measuring Sex stereotypes: A Multination Study*. Newbury Park: CA: Sage.

Woodcock, A., Hernandez, P. R., Estrada, M., and Schultz, P. W. (2012). The consequences of chronic stereotype threat: domain disidentification and attrition (for some). *J. Personal. Soc. Psychol.* 103, 635–646. doi: 10.1037/a0029120

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science* 330, 686–688. doi: 10.1017/CBO9781107415324.004

Younger, M., and Warrington, M. (2008). The gender agenda in primary teacher education in England: fifteen lost years? *J. Educ. Policy* 23, 429–445. doi: 10.1080/02680930802054396

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 31 October 2019
doi: 10.3389/fpsyg.2019.02386

The Role of Social and Ability Belonging in Men's and Women's pSTEM Persistence

Sarah Banchefsky[1], Karyn L. Lewis[2] and Tiffany A. Ito[3]*

[1] Office of Data Analytics, University of Colorado Boulder, Boulder, CO, United States, [2] Center for School and Student Progress, NWEA, Portland, OR, United States, [3] Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, United States

The benefits of belonging for academic performance and persistence have been examined primarily in terms of subjective perceptions of social belonging, but feeling ability belonging, or fit with one's peers intellectually, is likely also important for academic success. This may particularly be the case in male-dominated fields, where inherent genius and natural talent are viewed as prerequisites for success. We tested the hypothesis that social and ability belonging each explain intentions to persist in physical science, technology, engineering, and math (pSTEM). We further explore whether women experience lower social and ability belonging than men on average in pSTEM and whether belonging more strongly relates to intentions to persist for women. At three time points throughout a semester, we assessed undergraduate pSTEM majors enrolled in a foundational calculus or physics course. Women reported lower pSTEM ability belonging and self-efficacy than men but higher identification with pSTEM. End-of-semester social belonging, ability belonging, and identification predicted intentions to persist in pSTEM, with a stronger relationship between social belonging and intentions to persist in pSTEM for women than men. These findings held after controlling for prior and current academic performance, as well as two conventional psychological predictors of academic success.

Keywords: pSTEM, STEM, gender, social belonging, ability belonging, persistence, self-efficacy, identification

Edited by:
Bernhard Ertl,
Universität der Bundeswehr München, Germany

Reviewed by:
Anne Deiglmayr,
Leipzig University, Germany
Lysann Zander,
Leibniz University Hannover, Germany

***Correspondence:**
Sarah Banchefsky
Sarah.banchefsky@colorado.edu;
banchefs@colorado.edu

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *21 June 2019*
Accepted: *07 October 2019*
Published: *31 October 2019*

Citation:
Banchefsky S, Lewis KL and Ito TA (2019) The Role of Social and Ability Belonging in Men's and Women's pSTEM Persistence. Front. Psychol. 10:2386. doi: 10.3389/fpsyg.2019.02386

INTRODUCTION

> *"It's important to appreciate the background of endless skepticism that every woman in tech faces, and the resulting exhaustion we feel as the legitimacy of our presence is constantly questioned...There is always a jury, and it's always still out."*
>
> *–Cynthia Lee, Ph.D.*

Women remain starkly underrepresented in the physical sciences, technology, engineering, and mathematics (pSTEM) in many countries, including the United States, fields offering careers that are lucrative (Chamberlain and Jayaraman, 2017) and plentiful (President's Council of Advisors on Science and Technology [PCAST], 2012). At the university level in the U.S. in 2012, women obtained just 20% of engineering degrees, 19% of physics degrees, and 18% of computer science degrees (National Science Foundation and National Center for Science and Engineering Statistics, 2015). Moreover, roughly 35% of United States students, despite being well qualified and adequately prepared in math and science, abandon pSTEM fields during the first few semesters of college

(Seymour and Hewitt, 1997; Daempfle, 2003); this percentage is even higher among women (Seymour and Hewitt, 1997; Huang et al., 2000; Blickenstaff, 2005; Vogt et al., 2007; Ohland et al., 2008; Chen, 2013; Ellis et al., 2016). If merely 10% of students who leave STEM majors during higher education could be retained, the United States could achieve its future national workforce needs, which are currently deficient in STEM fields (President's Council of Advisors on Science and Technology [PCAST], 2012).

A number of social and interpersonal factors underlie the gender gap in representation within pSTEM fields, ranging from personal life choices—constrained or freely made—to unwelcoming masculine cultures (for reviews, see Eccles, 1994; Blickenstaff, 2005; Ceci et al., 2009; Cheryan et al., 2017). Of these many factors, recent research suggests a large role for social belonging (Lewis et al., 2016), which entails feeling like a valued, accepted, and legitimate member of a particular environment (Goodenow, 1993; Baumeister and Leary, 1995).[1] The need to belong and form interpersonal attachments with others is a fundamental, ubiquitous human motivation related to both physical and psychological health (Baumeister and Leary, 1995; Cacioppo and Cacioppo, 2014).

Not surprisingly, people who anticipate belonging in pSTEM fields express greater interest in pursuing such fields in the first place (Murphy et al., 2007; Cheryan and Plaut, 2010). Once pursuing a pSTEM education, social belonging is related to persistence (Goodenow, 1993; Freeman et al., 2007; Walton and Cohen, 2007; Hausmann et al., 2009; Walton and Cohen, 2011; Good et al., 2012; Thoman et al., 2014; Eddy and Brownell, 2016) even after accounting for objective ability and confidence that one can successfully complete the tasks required for success (i.e., self-efficacy; Bandura, 1977; Wilson et al., 2015; Lewis et al., 2017).

Schmader and Sedikides (2018) recently argued that sense of "fit" with an environment is multifaceted and dimensions other than social fit also influence an individual's decision to pursue a domain (see also Höhne and Zander, 2019). They also theorize that evaluations of social belonging are relevant only when social interactions are expected. Interacting with instructors and peers is an inherent component of pursuing a course of study. However, there are likely other relevant considerations, including how well one's aptitude meets the perceived requirements of the domain (McPherson et al., 2018b). Consistent with this, Lewis and Hodges (2015) found that the degree of intellectual fit—the subjective sense that one possesses the same abilities, skills, and knowledge as one's peers—predicts motivations and intentions to persist academically among undergraduates enrolled in psychology and linguistics courses (note that they refer to low levels of ability belonging as "ability belonging uncertainty"). pSTEM students were not examined in this past research, but the heavy emphasis within many pSTEM fields on inherent genius and natural talent is well documented (Leslie et al., 2015; Meyer et al., 2015; Ito and McPherson, 2018; Deiglmayr et al., 2019) and likely makes intellectual fit particularly relevant within pSTEM.

Given the important role of belonging in predicting persistence, it is critical to consider not only whether there are gender disparities in pSTEM belonging, and whether these disparities help explain the dearth of women in male-dominated fields, but also in what specific dimensions of belonging gender differences occur. Research indicates that women experience lower social belonging than men in male-dominated fields such as physics and computing (Lewis et al., 2016, 2017). Women's lower social belonging is likely due to a number of factors unique to male-dominated fields: the dearth of women (Murphy et al., 2007; Dasgupta et al., 2015), the lack of relatable role models (Cheryan et al., 2013), subtly unwelcoming or even overtly hostile masculine cultures (Settles et al., 2006; Cheryan et al., 2017), the greater prevalence of sexist jokes (Gonsalves et al., 2016), and non-verbal behavior from men that excludes women from professional conversations (Barthelemy et al., 2016).

The same cues that erode women's sense of social belonging likely also erode their sense of ability belonging—their belief that they have the same abilities and intellectual capacity as their peers. For example, Smith et al. (2013) found that compared to men, women in STEM graduate programs believed they worked harder than the average student in order to achieve the same outcome. In a second study, they demonstrated that the male-dominated status of a given field drove women's concerns about working harder for the same results. Specifically, undergraduates considered a fictional "eco-psychology" graduate program. When eco-psychology was depicted as male-dominated rather than gender-balanced, women anticipated working harder than the typical student to achieve success, which in turn diminished their interest in pursuing the program (Smith et al., 2013).

As far as we are aware, whether women in fact experience lower ability belonging than men in pSTEM fields has not been specifically examined, although women recruited from psychology and linguistics classes did express lower ability belonging than men in two of three studies, even after accounting for their objective ability (GPAs; Lewis and Hodges, 2015). If anything, we suspect that this existing gender gap in ability belonging will be exacerbated in male-dominated fields, where (1) natural ability is valued and viewed as necessary for success (Leslie et al., 2015) and (2) women's natural ability is negatively stereotyped (Spencer et al., 1999; Tiedemann, 2000; Stephens-Davidowitz, 2014). Together, these factors may make ability belonging a particularly relevant consideration in women's decisions about whether or not to persist in pSTEM fields.

In addition to mean gender differences in social and ability belonging, both facets of belonging may be more important factors for women's progression in pSTEM fields than men's. The vulnerability hypothesis (Hughes et al., 2015) states that the individuals most at risk of failure in a particular academic setting will be most affected by their subjective experiences within that setting (see also Johnson et al., 2007; Walton and Cohen, 2007, 2011; Murphy and Zirkel, 2015). Supporting this hypothesis, multiple studies show that social belonging is a stronger predictor of academic outcomes among historically marginalized and negatively stereotyped groups, such as women

[1]There is overlap between social belonging and the social dimension of the self-concept (e.g., Markus and Wurf, 1987; Bracken, 1996). However, because some views of the social dimension of the self-concept focus on perceptions of one's social skills and competence (e.g., Bracken, 1992), whereas we are interested in subjective perceptions of acceptance, we prefer the term *belonging*.

in male-dominated pSTEM fields (Holleran et al., 2011; Good et al., 2012; Walton et al., 2015). Indeed, recent research found not only that women report lower average social belonging than men in computing and physics but also that social belonging more strongly predicted persistence in their major for women than men (Lewis et al., 2017).

Although a small body of research indicates that social belonging is more important to women's persistence than men's in male-dominated majors, it remains to be tested whether ability belonging is likewise a stronger predictor of women's persistence. Just as women may be more prone to scan their environment and daily experiences for examples of whether they socially belong (Cheryan et al., 2009), they may also be more attuned to experiences confirming or denying whether they intellectually belong. In addition, recent research shows that female pSTEM majors more strongly believe that their fields require innate brilliance than their male peers do (Deiglmayr et al., 2019). This greater expectation that brilliance is required may make women more sensitive to their subjective assessments of ability fit than men.

CURRENT RESEARCH

Expanding upon prior research on belonging and in keeping with recent theorizing on the importance of different aspects of fit (Lewis and Hodges, 2015; Schmader and Sedikides, 2018), the current work examines gender differences in social and ability belonging, as well as whether each type of belonging is more tightly tied to women's intentions to persist in pSTEM than men's. We provide a stringent test of the relationship between these variables by accounting for objective academic performance as well as other theoretically important constructs linked to persistence, namely, self-efficacy and identification. In a recent review of possible psychological variables underlying gender disparities in STEM fields, Eddy and Brownell (2016) pointed to (social) belonging, self-efficacy, and identification as three key factors. Together, we call the four variables we measured—social belonging, ability belonging, self-efficacy, and identification—academic self-perceptions (ASPs).

Academic self-efficacy is the belief that one is capable of succeeding in specific academic tasks and goals such as exams and coursework (Bandura, 1977; Multon et al., 1991; Usher and Pajares, 2008). Robustly related to student performance, motivation, and persistence (Lent et al., 1986; Multon et al., 1991), women's lower self-efficacy compared to men is frequently cited as underlying the lack of women in pSTEM fields (Besterfield-Sacre et al., 2001; Stout et al., 2011; Tellhed et al., 2017). Even when objective performance is identical, women tend to report lower self-efficacy than men in pSTEM fields (Correll, 2001, 2004; Spelke, 2005; Else-Quest et al., 2010; Sikora and Pokropek, 2012).

Although correlated, it is worth noting that ability belonging is distinct from self-efficacy (Lewis and Hodges, 2015). Self-efficacy captures confidence that one can succeed in tasks, and does not entail direct social comparison to one's peers. A student could have high self-efficacy (e.g., believe she can perform well on exams) but nevertheless believe she has lower intellectual aptitude than her peers. In contrast, a student could theoretically have low self-efficacy and question his ability to perform well on homework and exams but still believe he is as or more capable than his peers.

Academic identification refers to caring about one's performance in a given domain and basing one's self-worth or self-esteem upon performance in that domain (Spencer et al., 1999; Chemers et al., 2011; Osborne and Jones, 2011; Cundiff et al., 2013). Students highly identified with pSTEM view it as an important aspect of "who they are," an identity from which they draw meaning and self-esteem. Research shows that in pSTEM fields, identification is related to positive academic outcomes including higher performance, motivation, and commitment to doing well (Smith and White, 2001), as well as greater likelihood of choosing a pSTEM career (Hazari et al., 2010). The latter two studies also found that men expressed stronger pSTEM identification than women, suggesting that there may also be gender differences favoring men regarding pSTEM identification.

In summary, research has shown a clear relationship between self-efficacy (e.g., I can succeed in this field), identification (e.g., this field is important to me), and positive outcomes in a field. In addition to these factors, we maintain that belonging—as a basic human need and motivation—should be tied to intentions to persist in a given field, over and above self-efficacy and identification (Lewis et al., 2017; Ito and McPherson, 2018). Specifically, we predicted that over and above prior and current academic performance, women would report lower ability and social belonging in pSTEM than men on average across the semester (Hypothesis 1); that both ability and social belonging would be related to intentions to persist among both women and men even after accounting for self-efficacy and identification (Hypothesis 2); and that the relationship between social and ability belonging and pSTEM persistence would be stronger for women than for men (Hypothesis 3).

MATERIALS AND METHODS

Participants

Participants were undergraduates taking a large gateway physics or calculus course, both of which are required to advance in nearly all pSTEM majors at the university where data were collected. There were three physics sections, all taught by the same male professor, and seven calculus sections, two of which were taught by the same male professor and five of which were each taught by a different male professor. More information on the courses is available in Section 1 in the **Supplementary Material**.

Selecting from introductory courses ensured obtaining a sample of students early in their academic path, when attrition is highest (Daempfle, 2003). Introductory physics and calculus courses were selected in particular because women's representation in physics is among the lowest of any pSTEM field (National Science Foundation and National Center for Science and Engineering Statistics, 2015), and women are 1.5 times more likely to leave the pSTEM pipeline after college calculus compared to men (Ellis et al., 2016), making it particularly relevant to understand gender disparities among students in these classes. Both courses are historically male-dominated (in our sample, 21.24% of physics students and 25.81% of calculus students were women), consist primarily of students majoring in pSTEM, and are the first in a sequence of required courses for pSTEM majors at the university (for example, both courses are required for all engineering majors, who comprised the majority of our sample).

All students enrolled in the selected physics (n = 831) or calculus (n = 648) sections at the beginning of the fall semester were contacted via university email and invited to participate in the study in exchange for compensation (see Procedures for details). Of the 1,479 contacted, 599 responded at Time 1 (40.50%). After data collection, three exclusion criteria were applied. First, because we were interested in the persistence of students pursuing pSTEM fields, students who either self-reported being a non-pSTEM major or were undecided about their major (n = 30) were removed (see Section 2 in the **Supplementary Material**, for qualifying pSTEM majors). Second, we removed students enrolled in honors sections of the courses (n = 31) because these courses were structured differently and had much smaller enrollments relative to the large-lecture courses. Third, students whose institutional records indicated that they withdrew from the course or received an "incomplete" (n = 29) were removed because (1) most did not complete the third survey and were therefore missing data on intentions to persist in pSTEM and (2) intentions to persist in pSTEM fields are inherently constrained for students who have not completed a required gateway pSTEM course. Given that the three exclusion criteria could overlap, the final sample of eligible participants included 516 students.

Of these, 121 (23%) self-identified as women (21.07% of the physics sample and 26.27% of the calculus sample). A majority of students (84.30%) were engineering majors, followed by physics (6.59%), astronomy (2.71%), chemistry (2.32%), biochemistry (1.74%), and mathematics (1.36%). The majority identified as White (70.54%), followed by Asian (9.50%), other (9.30%; the majority wrote in "Asian," but there was also "Indian," "Middle Eastern," and "Arab"), Hispanic (3.68%), and Black (0.39%); 6.20% selected more than one racial category. Institutional data (available for n = 398 or 77.13% of students who permitted access) indicated that the sample was primarily first-year students (90.70%), followed by sophomores (7.79%), juniors (1.26%), and seniors (0.25%).

This study was carried out in accordance with the recommendations of United States Office for Human Research Protections. The protocol was approved by the University of Colorado's Institutional Review Board. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

Materials

These data are part of a broader study that encompassed three surveys and a variety of other measures beyond the focus of this paper (documented in Section 3 in the **Supplementary Material**). The current study focuses on a subset of the measures (all of which can be viewed in **Table 1**) and primarily on Surveys 1 and 3, conducted at the beginning and end of the semester. The response format for items was a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). All scales demonstrated high reliability (see **Table 2**).

TABLE 1 | Academic self-perceptions and intentions to persist in STEM.

Construct	Items
pSTEM social belonging	I feel like I belong in STEM.
	People in STEM accept me.
	I feel like an outsider in STEM. (r)
	I feel a connection with the STEM community.
	People in STEM are a lot like me.
	I fit in well in STEM.
pSTEM ability belonging	I sometimes feel like other students in STEM have skills that I don't have. (r)
	I'm not sure that I'm cut out for STEM. (r)
	I feel similar to the kinds of people who have what it takes to succeed in STEM.
	I'm not certain I fit in intellectually in STEM. (r)
pSTEM identification	It is important to me that I am good at math and science.
	Doing well on math and science tests is important to my self-esteem.
	Is it important to me to perform well on science and math tests.
	Having strong math and science skills is important to me.
pSTEM self-efficacy	I am confident I can. . .
	Complete homework assignments by myself.
	Perform well on exams.
	Demonstrate what I know on exams.
	Learn STEM concepts.
	Complete the course with a B or better.
Intentions to persist in pSTEM	I could see myself going into a career related to STEM.
	I look forward to taking more STEM courses.
	It is my intention to major in a STEM discipline.
	I have no doubt that I will graduate with a degree in a STEM field.
	I have seriously considered changing my major to a non-STEM related field. (r)
	STEM is the right career path for me.

pSTEM, physical sciences, technology, engineering, and mathematics. (r) Indicates reverse-scored.

Demographics

Self-reported demographics included gender (male, female, other), age, ethnicity (Black/African American, Asian American, Hispanic/Mexican American, White/Caucasian, and other), and academic major (open response).

Academic Self-Perceptions

At each time point, six items assessed social belonging (e.g., "I feel like I belong in STEM"; adapted from Walton and Cohen, 2007; Good et al., 2012; Lewis et al., 2016); four assessed ability belonging [e.g., "I'm not sure that I'm cut out for STEM" (reverse-scored); Lewis and Hodges, 2015]; four assessed identification (e.g., "It is important to me that I am good at STEM"; Spencer et al., 1999); and five assessed self-efficacy (e.g., "In STEM,

TABLE 2 | Descriptive statistics for academic self-perceptions and pSTEM intentions.

Construct	# Items	Mean alpha	Mean (Standard Deviation)						Gender difference	
			Overall		Men		Women		*z*-value	Cohen's *d*
Social belonging	6	0.85	3.82	(0.51)	3.81	(0.49)	3.83	(0.54)	0.40	0.04
Ability belonging	4	0.71	3.64	(0.55)	3.68	(0.53)	3.52	(0.59)	−2.76**	0.29
Identification	4	0.78	4.37	(0.46)	4.34	(0.47)	4.44	(0.44)	2.14*	0.21
Self-efficacy	5	0.88	4.12	(0.49)	4.19	(0.47)	3.91	(0.50)	−5.47***	0.58
pSTEM intentions	6	0.90	4.12	(0.72)	4.13	(0.71)	4.06	(0.76)	−0.87	0.10

****p* < 0.001, ***p* < 0.01, **p* < 0.05. The scales ranged from 1 to 5. Alpha is the average of the separate alphas for each scale at Time 1, Time 2, and Time 3. The mean and standard deviation for the academic self-perceptions were grand means across the separate Time 1, Time 2, and Time 3 mean scores (e.g., the mean of belonging at Time 1, 2, and 3). Cohen's d was also computed from the estimated means and standard deviations from the path model using full information maximum likelihood (FIML).*

I am confident that I can demonstrate what I know on exams"; Bandura, 1977; Betz and Hackett, 1983). For brevity in the survey, we did not specify that we were referring to physical sciences, but students in our sample were specifically majoring in physical sciences and not social sciences. We therefore refer to pSTEM throughout this paper. Given that two constructs refer to aspects of belonging, and that past operationalizations of social belonging include items that might be affected by perceptions of ability (e.g., "I feel like I belong in STEM"), we conducted comparative confirmatory factor analyses to assess whether ability and social belonging should be combined into one factor. Compared to the single-factor belonging model, fit was significantly improved when social and ability belonging were treated as separate factors. This was true at Time 1, χ^2 difference (1) = 55.74, Time 2, χ^2 difference (1) = 75.64, and Time 3, χ^2 difference (1) = 59.22, all *p*s < 0.001.

pSTEM Intentions

The key dependent variable was self-reported intentions to persist in pSTEM (see **Table 1**), assessed by six items in the Time 3 survey near the end of the semester (e.g., "It is my intention to major in a STEM discipline"; $\alpha = 0.90$). Intentions are a proximal predictor of behavior (Ajzen, 1985, 1987, 1991, 2011) and consequently a frequently used measure of educational outcomes (e.g., Tinto, 1997; Murphy et al., 2007; Good et al., 2012; Lewis et al., 2017; Ito and McPherson, 2018). Studies also show a strong association between academic intentions to persist and actual persistence (e.g., Bean, 1982; Voorhees, 1987; Cabrera et al., 1993; Davidson et al., 2009; Luke et al., 2015).

Prior and Current Academic Performance

To account for prior academic performance, we accounted for high-school GPA and scores on standardized entrance exams (SAT math, ACT math, ACT science). Of participants who provided access to institutional records, SAT math scores were available for 164 (41.21%), ACT math and science-reasoning for 309 participants (77.64%), and 108 had records for both (27.13%). For each student, each available test score was standardized and, if appropriate, averaged into one index capturing standardized math/science test performance. To account for ongoing objective performance, we calculated the average GPA across all pSTEM courses during the semester in which the surveys were administered (including the course they were enrolled in; available for 98.99% of students who granted permission to access institutional records).

Missing Data

The 516 eligible participants who responded to the Time 1 survey ($n = 280$ in physics and $n = 236$ in calculus) were invited to participate in subsequent surveys. At Time 3, 441 participants responded (85.46% retained from initial enrollment). Women were marginally more likely to be retained at Time 3 than men [90.91% versus 83.80%, $\chi^2(1) = 3.22$, $p = 0.07$].

Little's (1988) Missing Completely at Random (MCAR) test was performed in R to examine missing patterns in the data. The test included the following 15 variables included in the most complex models: intentions to persist in STEM, the four ASPs at Time 1 and Time 3, self-reported gender, course, and the four codes capturing course professor. Institutional record data were not included in the MCAR test because these were not missing at random—rather, participants had the option of giving us access to these records (77.13% did so). The test indicated that the data were not missing completely at random: there were five patterns of missing data, $p < 0.01$ (the null hypothesis being that the data are missing completely at random). This is unsurprising given the longitudinal nature of the data, and moreover, the use of full information maximum likelihood (FIML) estimation accounts for missing data, even if not missing completely at random (Baraldi and Enders, 2010).

Procedures

Participants received an email invitation stating that we were interested in "issues that students in science and math majors at CU Boulder face" and that they were being contacted because they were enrolled in a science or math course. They were informed that their participation would consist of completing up to three online surveys regarding their experiences in their courses. Participants received $10 USD for completing the first survey, $15 USD for the second survey, and $20 USD for the final survey. To encourage complete participation, a $10 USD bonus was offered for completing all three surveys, and students were also entered into a raffle to win an additional prize (ranging from $25 to $50 USD).

The first survey was administered at the beginning of the semester (Time 1; Week 2 of the 16-week semester), the second was administered at the middle of the semester (Time 2; Week 8), and the final survey was administered at the end of the semester (Time 3; Week 16). Each survey was opened the day the invitations were emailed and remained open for 2 weeks. Reminder emails were sent to participants who had not yet completed the survey 1 week after it opened and 1 day before it closed. At the end of each survey, participants were then asked how they would like to receive their payment (Amazon gift card or cash pickup). When data collection was finalized, all participants were emailed a debriefing form.

At Time 1, participants gave informed consent, completed demographics, and were asked whether we could have their student identification number in order to access their institutional records. At all time points, participants first completed the ASP items in a fixed random order. Prior to the measures of interest here, participants were asked about experiences in their particular course (see Section 3 in the **Supplementary Material**). They were then informed, "The following questions ask about your experiences and perceptions of the broader field of science, engineering, and math (STEM) in general," before completing the social belonging, ability belonging, and identification items intermixed in a fixed, random order. The self-efficacy items were presented on the next page and had a slightly different prompt: "Please rate how confident you are that you can do each of the following things." Other items not relevant to the current manuscript were then completed (see Section 3 in the **Supplementary Material**). At Time 3 only, participants reported their intentions to persist in pSTEM after reporting their perceived course utility, followed by the additional measures in Section 3 in the **Supplementary Material**.

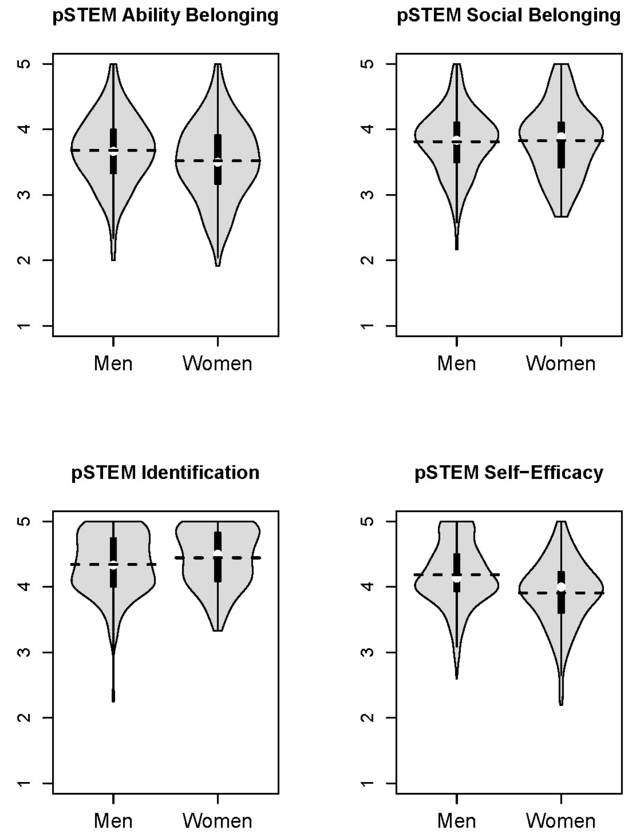


FIGURE 1 | Violin plots depicting the academic self-perceptions on average over the semester for men and women. The boxplot and whiskers are in black; the white circle represents the median, and dashed horizontal lines represent the mean. The distribution of responses, in gray, is reflected on each side of the boxplot.

RESULTS

Preliminary Analyses

On average across the semester (averaging across Time 1, Time 2, and Time 3), students reported relatively high social and ability belonging, identification, and self-efficacy (see **Table 2**). Women reported lower average ability belonging in pSTEM than men. However, women and men did not differ in social belonging in pSTEM. Women also expressed lower average pSTEM self-efficacy, but greater average pSTEM identification, compared to men (see **Table 2** and **Figure 1**). Students expressed strong intentions to persist in pSTEM, which did not differ by gender. Regarding prior performance, women had lower standardized math and science test scores than men, unstandardized $b = -0.12$, $z = -2.02$, $p = 0.04$, but there was no gender difference in high-school GPA, $b = 0.04$, $z = 1.763$, $p = 0.10$. Regarding ongoing academic performance, there was also no gender difference in pSTEM GPA that semester, $z = -0.03$, $p = 0.45$.

Table 3 depicts correlations controlling for participant gender. Overall, the ASPs were themselves highly positively correlated; students who felt greater social belonging also tended to have greater ability belonging, identification, and self-efficacy. Notably, the correlation between social and ability belonging was very similar to that observed in prior research ($r = 0.63$ here and $r = 0.60$ in Lewis and Hodges, 2015). Standardized test scores and high-school GPA were positively correlated with social belonging, ability belonging, and self-efficacy, but were unrelated to identification. Whereas standardized test scores in math and science were positively related to intentions to persist in pSTEM, high-school GPA was notably unrelated to pSTEM intentions. Finally, pSTEM GPA was positively related to all variables—prior performance, each ASP, as well as intentions to persist in pSTEM.

Analytic Strategy

Analyses assessing our three hypotheses were conducted in R version 1.0.136, using path models conducted with the R package lavaan (Rosseel, 2012). All continuous predictors were centered around their mean, and all categorical predictors were contrast-coded. Primary analyses examined gender (women = 1, men = −1), controlling for prior academic performance (i.e., standardized test scores in math and science and high-school GPA) and the students' course professor for the course they were responding about in the survey (five orthogonal contrast-codes, one of which compared physics to calculus). We accounted for indicators of pSTEM achievement to ensure that any observed

TABLE 3 | Correlations and descriptive statistics of prior and ongoing performance, academic self-perceptions, and outcomes, partialing gender.

	1	2	3	4	5	6	7	8
(1) SAT/ACT math/science								
(2) High-school GPA	**0.32**							
(3) Average social belonging	**0.23**	**0.18**						
(4) Average ability belonging	**0.30**	**0.12**	**0.63**					
(5) Average identification	0.01	−0.04	**0.40**	**0.24**				
(6) Average self-efficacy	**0.28**	**0.09**	**0.51**	**0.57**	**0.44**			
(7) pSTEM GPA	**0.48**	**0.32**	**0.34**	**0.35**	**0.14**	**0.40**		
(8) Intentions to persist in pSTEM	**0.14**	−0.06	**0.55**	**0.49**	**0.44**	**0.46**	**0.31**	
Mean	0	3.78	3.82	3.60	4.39	4.05	3.02	4.10
sd	1	0.39	0.50	0.55	0.46	0.48	0.68	0.72

All correlations greater than or equal to an absolute value of 0.10 (in bold) are significant at p < 0.05. SAT/ACT math scores are standardized. sd, standard deviation. Self-perceptions are averaged over all three time points.

relationships and gender differences were not due to differences in prior or ongoing pSTEM performance. The inclusion of strong covariates can reduce the likelihood that observed associations are due to unmeasured confounds. We controlled for professor to account for non-independence in the data, which is likely to occur with grouped data. Here, student responses about their pSTEM experiences and their specific course professor are surely influenced by their professor, and students with the same professor may have more similar responses to each other. Ignoring non-independence biases the results (Judd et al., 2017). We also controlled for students' academic year (e.g., freshmen, sophomores; numerically coded and mean-centered) because initial exploratory analyses indicated that women had a higher class standing than men on average, $t(417) = 3.64$, $p < 0.001$. Specifically, men were more likely to be freshmen (92.77% of men versus 83.17% of women), less likely to be sophomores (6.92% of men versus 11.88% of women), and less likely to be juniors (0.31% of men versus 3.96% of women).

All models used FIML estimation, the preferred analytic method to implicitly handle missing data. FIML provides less biased parameter estimates even if data are not missing completely at random (Baraldi and Enders, 2010). Rather than deleting observations with missing data on predictor variables as would occur using ordinary least squares estimation, FIML estimates the values of the predictors based on the available data. Importantly, this approach improved our statistical power because it retained students who did not provide access to institutional record data ($n = 118$).

Data were collected at three time points; thus, they were longitudinal in nature and could have been examined using latent growth curve modeling (LGCM; Curran et al., 2010). Although LGCM is a powerful technique to analyze longitudinal data, structural equation modeling (of which LGCM is a specific type) requires a large sample size, especially when the model is complex (Wolf et al., 2013); Kline (2011) recommends at least 200 participants per group, and our sample contains only 121 women. Initial attempts to use an LGCM approach were not fruitful—in particular, the models for women demonstrated estimation problems (e.g., negative latent variances, failures to converge). To account for the longitudinal nature of the data and enhance statistical power, we included ASPs (e.g., social belonging) at Time 1 and Time 3 as predictors of pSTEM persistence (for a similar analytic approach, see Good et al., 2012). This model specification enabled us to ask how a change in belonging from baseline (measured at the beginning of the semester) is related to pSTEM persistence. (A complete examination of how each ASP changed throughout the semester is beyond the scope of this paper, but see Section 4 in the **Supplementary Material** and **Supplementary Table 1**, for an analysis of how each changed over the course of the semester. In sum, social and ability belonging remained flat over the course of the semester, and identification and self-efficacy both dropped. Notably, identification and self-efficacy were near ceiling at Time 1. This initial ceiling effect may have contributed to their decline over the course of the semester.)

Finally, because some of the predictors showed high inter-correlations (in particular, social and ability belonging; see **Table 3**), we tested for multicollinearity, which occurs when two or more predictors in a model are highly correlated and can cause statistical estimations to be unreliable (Wooldridge, 2013; Thompson et al., 2017). To do so, we examined the variance inflation factors (VIFs), the extent to which variation in the model is inflated by the presence of correlation among predictor variables (Salmerón Gómez et al., 2016). This was done in a model regressing intentions to persist in pSTEM onto all four ASPs at Time 1 and Time 3 (each mean-centered), high-school GPA (mean-centered), standardized test performance, pSTEM GPA (mean-centered), gender (contrast-coded), academic year (mean-centered), and professor (five orthogonal contrast-codes).

The VIF model revealed that social and ability belonging at Time 3 were redundant enough to warrant being either combined or tested in separate models (see **Supplementary Table 2**) (Wooldridge, 2013). As noted in the Section "Materials and Methods," factor analyses showed better fit when social and ability belonging were treated as separate factors rather than being combined into a single factor. These results, in combination with prior theoretical work establishing that social and ability belonging are two distinct constructs (Lewis and Hodges, 2015), motivated our choice to treat these as separate factors tested in two separate models. Details on the VIF analyses are provided in Section 5 in the **Supplementary Material**.

Do Women Report Lower Social and Ability Belonging Than Men? (Hypothesis 1)

Hypothesis 1 was assessed with path models that accounted for important covariates—academic year, professor, and prior and ongoing performance (i.e., high-school GPA, SAT/ACT math scores, and pSTEM GPA). As seen in **Table 4**, the raw gender differences presented in **Table 2** and **Figure 1** largely persisted when controlling for these four additional variables.

TABLE 4 | Effect of gender and covariates on academic self-perceptions.

Predictors	Unstandardized betas			
Outcomes	**Ability belonging**	**Social belonging**	**Self-efficacy**	**Identification**
Gender (+1 = women, −1 = men)	−0.05+	0.02	**−0.11*****	**0.06***
Academic year	−0.12	−0.11	**−0.13***	−0.04
High-school GPA	−0.05	0.08	−0.10	−0.08
SAT and ACT math and science	**0.09****	0.03	**0.06***	−0.03
pSTEM GPA	**0.22*****	**0.20*****	**0.25*****	**0.12****
Prof code 1 (physics vs. calculus)	0.00	0.00	0.01	0.01
Prof code 2	0.01	0.00	0.00	0.01
Prof code 3	0.00	0.01	0.00	−0.02
Prof code 4	0.06	0.03	0.02	0.03
Prof code 5	0.04	0.01	**−0.07***	0.01
R^2	17.00%	13.30%	24.10%	4.70%

****p* < 0.001, ***p* < 0.01, **p* < 0.05, +*p* < 0.10. Prof, Professor. The first professor code captures whether students were evaluating their physics or calculus class. The other four orthogonal contrast-codes control for which professor the students had for their course. All continuous predictors were mean-centered, and categorical predictors were orthogonally contrast-coded. Significant effects are in bold.*

TABLE 5 | Parameter estimates social belonging—intentions to persist models.

	Unstandardized betas		
Predictors	**Model 1 no ASP**	**Model 2 add belonging**	**Model 3 add other ASP**
Gender (+1 = women, −1 = men)	0.02	−0.02	−0.01
Academic year	−0.17	−0.09	−0.06
High-school GPA	**−0.36****	**−0.37*****	**−0.32****
SAT and ACT math and science	0.04	0.00	0.01
pSTEM GPA	**0.37*****	**0.18****	**0.13***
Prof code 1 (physics vs. calculus)	0.01	0.01	0.00
Prof code 2	−0.02	−0.03	−0.03
Prof code 3	−0.02	−0.03	−0.02
Prof code 4	−0.01	−0.00	0.01
Prof code 5	−0.08	−0.06	−0.03
Social belonging (Time 3)	–	**0.58*****	**0.45*****
Social belonging (Time 1)	–	**0.14***	0.09
Identification (Time 3)	–	–	**0.31*****
Identification (Time 1)	–	–	0.03
Self-efficacy (Time 3)	–	–	0.07
Self-efficacy (Time 1)	–	–	0.07
R^2	15.60%	38.00%	44.50%

****p* < 0.001, ***p* < 0.01, **p* < 0.05. The first professor code captures whether the students were evaluating their physics or calculus class. The other four orthogonal contrast-codes control for which professor the students had for their course. ASP, academic self-perception. All continuous predictors were mean-centered, and categorical predictors were orthogonally contrast-coded. Significant effects are in bold.*

We found partial support for Hypothesis 1, that women would have lower social and ability belonging than men: women expressed the *same* pSTEM social belonging as men, $p = 0.34$, but marginally lower pSTEM ability belonging, b = −0.05, *z*-value = −1.82, $p = 0.068$.

Women also expressed lower pSTEM self-efficacy than men, $b = -0.11$, *z*-value = −4.69, $p < 0.001$, but higher pSTEM identification, $b = 0.06$, $z = 2.50$, $p = 0.012$. In summary, over and above academic preparation and current academic performance, women cared even more about their performance and knowledge in pSTEM than did men (i.e., women's pSTEM identification was greater than men's) but were simultaneously more concerned that they did not have what it takes to succeed in pSTEM both individually (i.e., women's pSTEM self-efficacy was lower than men's) and relative to their peers (i.e., women's pSTEM ability belonging was marginally lower than men's).

Do Social and Ability Belonging Predict pSTEM Persistence? (Hypothesis 2)

Our second hypothesis was that social and ability belonging would predict intentions to persist in pSTEM, even after accounting for self-efficacy and identification, academic preparation, and ongoing academic performance.

We included both belonging at Time 1 and that at Time 3 as predictors of pSTEM persistence (for a similar analytic approach see Good et al., 2012). This model specification enabled us to ask how change in belonging from baseline (measured at the beginning of the semester) is related to pSTEM persistence.

Due to multicollinearity between social and ability belonging, each predictor was tested in a separate model, one for social and one for ability belonging. Results are depicted in **Tables 5**, **6**, each of which shows three models: the first model included only the academic preparation variables, current academic performance, and course professor codes. The second model adds the belonging variable of interest (either social or ability belonging). The comparison between model 1 and model 2 reveals the contribution of belonging to predicting persistence. Finally, a third model adds identification and self-efficacy to provide an assessment of the degree to which belonging continues to predict persistence when other critical aspects of ASPs are included.

Results supported Hypothesis 2: ability and social belonging at the end of the semester were each strongly related to intentions to persist in pSTEM, even after accounting for identification and self-efficacy, as well as prior and ongoing academic performance (see **Tables 5**, **6**, respectively). As seen by comparing columns 1 and 2 in **Tables 5**, **6**, including either social or ability belonging as a predictor in the models explained more than twice the variance in intentions to persist in pSTEM (R^2 = 38.00% and R^2 = 30.00%, respectively) relative to the model without belonging (R^2 = 15.60%). Furthermore, social and ability belonging at Time 3 remained strongly related to intentions to persist after accounting for initial social and ability belonging, as well as end-of-semester self-efficacy (unrelated to intentions to persist) and identification (significantly and positively related to intentions to persist). In sum, changes in the sense of fit within pSTEM, whether socially and intellectually, were significantly related to intentions to persist in pSTEM. Students entered pSTEM with a certain level of belonging, and the change they

TABLE 6 | Parameter estimates ability belonging—intentions to persist models.

Predictors	Unstandardized betas		
	Model 1 no ASP	Model 2 add belonging	Model 3 add other ASP
Gender (+1 = women, −1 = men)	0.02	0.02	0.01
Academic year	−0.17	−0.14	−0.10
High-school GPA	**−0.36****	**−0.35*****	**−0.29****
SAT and ACT math and science	0.04	−0.02	0.00
pSTEM GPA	**0.37*****	**0.21****	**0.13***
Prof code 1 (physics vs. calculus)	0.01	0.01	0.00
Prof code 2	−0.02	−0.02	−0.02
Prof code 3	−0.02	−0.03	−0.02
Prof code 4	−0.01	0.00	0.00
Prof code 5	−0.08	−0.07	−0.03
Ability belonging (Time 3)	–	**0.42*****	**0.34*****
Ability belonging (Time 1)	–	0.11+	0.05
Identification (Time 3)	–	–	**0.39*****
Identification (Time 1)	–	–	0.07
Self-efficacy (Time 3)	–	–	0.07
Self-efficacy (Time 1)	–	–	0.06
R^2	15.60%	30.00%	41.80%

****p < 0.001, **p < 0.01, *p < 0.05, +p < 0.10. The first professor code captures whether students were evaluating their physics or calculus class. The other four orthogonal contrast-codes control for which professor the students had for their course. All continuous predictors were mean-centered, and categorical predictors were orthogonally contrast-coded. Significant effects are in bold.*

experienced in pSTEM belonging over the course of the semester was linked to their intentions to persist in pSTEM. This occurred even after accounting for prior academic performance (high-school GPA and standardized test scores), ongoing academic performance (pSTEM GPA), and changes in self-efficacy and identification over the course of the semester.

For completeness, we also conducted a path model including both predictors. This model indicated that both ability and social belonging were significant predictors of intentions to persist in pSTEM (for more details, see **Supplementary Table 3**). This suggests that social and ability belonging each uniquely explains intentions to persist in pSTEM. In other words, each type of belonging—social and ability—was significantly related to intentions to persist in pSTEM after controlling for the other type of belonging.

Do Ability and Social Belonging Relate to pSTEM Persistence More for Women Than Men? (Hypothesis 3)

We next tested Hypothesis 3, that ability and social belonging would play a stronger role in women's than men's intentions to persist in pSTEM. To do so, gender invariance tests were conducted using the final models used to test Hypothesis 2, as shown in column 3 of **Tables 5**, **6**. This entails comparing the chi-square of a model estimated separately for men and women to a model in which the path of interest (i.e., the relationship between end-of-semester social or ability belonging and intentions to persist in pSTEM) is constrained to be equivalent for the genders (see **Figures 2**, **3**). If the constrained model results in significantly reduced goodness of fit, the tested path is significantly different for women and men.

As seen in **Figure 2**, the relationship between end-of-semester ability belonging and intentions to persist was directionally 1.67 times stronger for women [beta (b) = 0.50, z = 4.74, $p < 0.001$, R^2 for entire model = 50.60%] than men (b = 0.30, z = 4.64, $p < 0.001$, R^2 for entire model = 42.2%). However, an invariance test indicated that the model fit was statistically equivalent when this path was forced to be the same for men and women, χ^2 difference (1) = 2.42, p = 0.12. The low number of women in the sample (n = 121) relative to men (n = 395) may have made it difficult to detect a significant interaction.

Social belonging showed the same pattern of results, but here, the gender difference was significant (see **Figure 3**), and the relationship between social belonging and intentions to persist was roughly twice the size for women than for men. Model fit was significantly worse when the path between end-of-semester social belonging and intentions to persist in pSTEM was forced to be equivalent for men and women. This indicated that the path between end-of-semester social belonging and pSTEM persistence was significantly stronger for women (b = 0.77, z = 5.61, $p < 0.001$, R^2 = 54.20%) than for men (b = 0.40, z = 5.17, $p < 0.001$, R^2 = 46.10%), χ^2 difference (1) = 5.30, p = 0.02. **Figure 4** shows the relation between ability belonging (left panel) and social belonging (right panel) and intentions to persist separately for women and men. As can be seen, the pattern of results is similar, with both ability and social belonging more related to women's intentions to persist in pSTEM than men's intentions to persist, although important for both genders.

Does Self-Efficacy or Identification Show a Gendered Relationship With Persistence?

For completeness, we also tested whether end-of-semester self-efficacy and identification were differentially related to intentions to persist in pSTEM for women and men. For both self-efficacy and identification, two invariance tests were conducted—for each of the models depicted in **Figures 2**, **3**, we compared a model estimated separately by gender to a model in which the path between Time 3 self-efficacy or identification was constrained to be equivalent for men and women (i.e., one invariance test was based on a model controlling for ability belonging, and the other based on the model controlling for social belonging). In the context of controlling for either ability belonging or social belonging, the relationship between self-efficacy and intentions was statistically equivalent for men and women: χ^2 differences (1) < 2.43, ps > 0.12. Similarly for identification, the relationship between end-of-semester identification and intentions to persist did not depend on gender, χ^2 differences (1) < 0.04, ps > 0.84.

DISCUSSION

The current research expands upon prior work in two key ways. First, we examined the influence of both social and ability belonging on pSTEM persistence. Importantly for women and

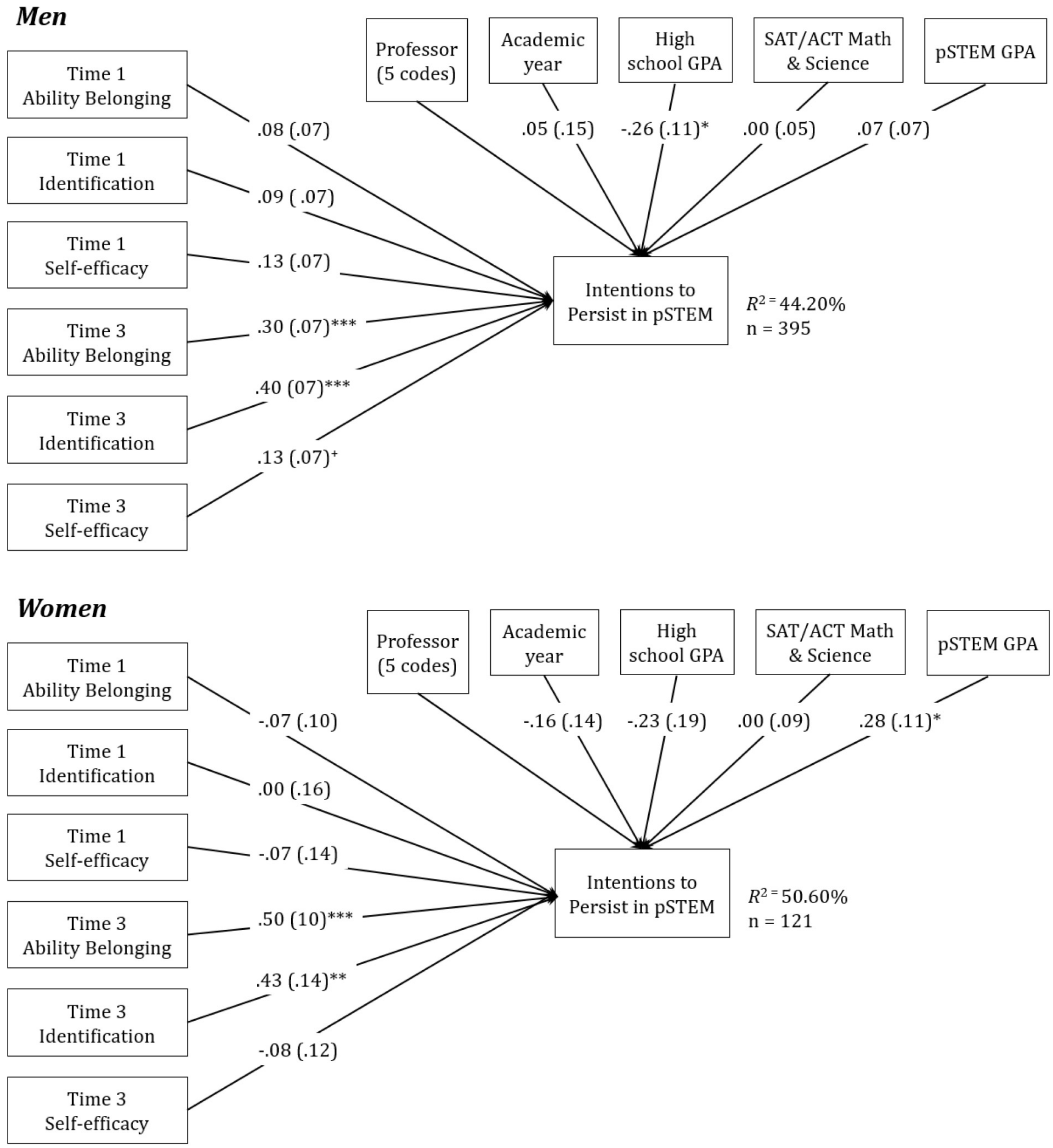


FIGURE 2 | Path models depicting relationships among ability belonging and intentions to persist for men and women, controlling for professor, academic year, and prior and ongoing performance. Unstandardized betas are presented, with Standard Errors in parentheses. For simplicity, paths for professor are not depicted.

men alike, subjective ASPs of one's social belonging, ability belonging, and identification each were uniquely related to intentions to persist in pSTEM, even after controlling for prior and ongoing performance. Indeed, models that included these ASPs predicted far more variation in intentions to persist in pSTEM than models that only included academic preparation. This has important implications for universities and pSTEM programs; although they cannot easily intervene to increase students' math and science preparation, they can certainly aim to foster more welcoming pSTEM environments that normalize academic struggle.

Second, it is the first research we are aware of to examine belonging to a more broadly defined group (i.e., pSTEM) and demonstrate a link between pSTEM belonging and pSTEM persistence. In contrast, most prior research examines belonging to a particular classroom or particular pSTEM field (e.g., mathematics; Good et al., 2012). Our results indicate that students have an experience beyond their individual classrooms and majors and can reflect and report on their ASPs regarding overall "pSTEM culture." This suggests that it may be wise to consider what kind of culture the institution is fostering at a broad level, not just within individual departments. The results

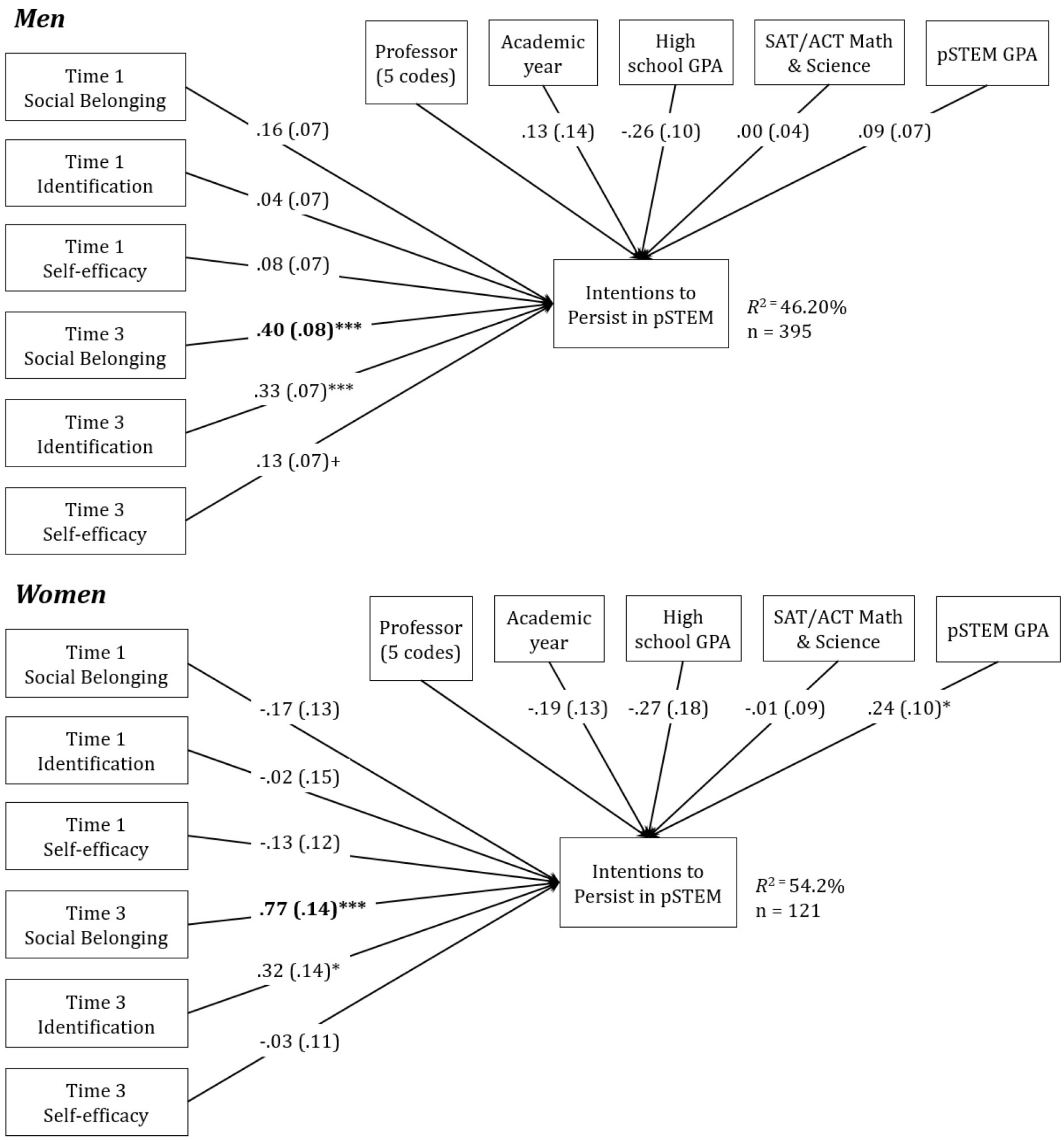


FIGURE 3 | Path models depicting relationships among social belonging and intentions to persist for men and women, controlling for professor, academic year, and prior and ongoing performance. Unstandardized betas are presented, with Standard Errors in parentheses. Paths in bold were significantly different for women and men. For simplicity, paths for professor are not.

also have implications for understanding the gender gap in pSTEM. Even among a selective group of women—within a sample of pSTEM majors who are highly identified with STEM, who have promising high-school GPAs, and many of whom had gained admission to a selective engineering college—women reported feeling more out of step intellectually with their peers than did men. Women also expressed less confidence that they could succeed on tasks important for academic success (i.e., they reported lower self-efficacy). At the same time, women were even more likely to care about their pSTEM ability and performance, reporting greater pSTEM identification than men. Notably, this combination of stronger identification and low self-efficacy may make women particularly susceptible to stereotype threat (Schmader, 2002).

Although women reported lower pSTEM ability belonging than men, no gender difference emerged for pSTEM social belonging. The lack of a gender difference in social belonging was surprising given recent research findings that women in computing felt lower social belonging in their major and that women across multiple pSTEM majors expressed lower social belonging in a physics class (Lewis et al., 2017). One reason we may not replicate this research is that the level of measurement

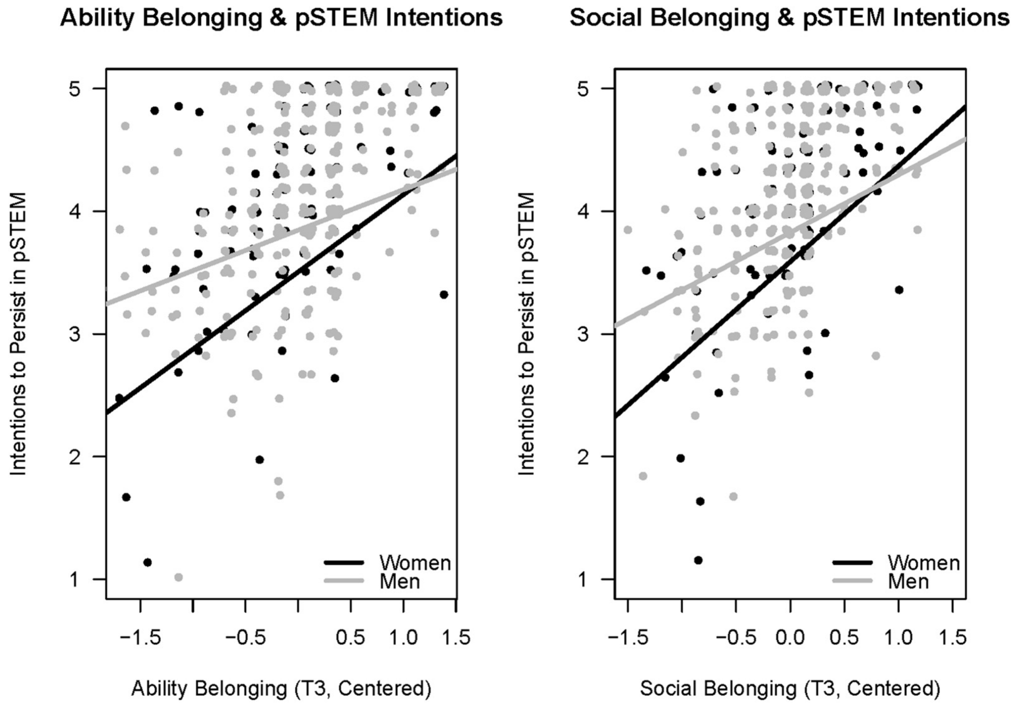


FIGURE 4 | Intentions to persist as a function of belonging (ability on the left, social on the right), by gender. Lines depict partial relationships after controlling for professor, academic year, prior and ongoing performance, baseline belonging, as well as self-efficacy and identification at Time 1 and 3. Although the relationships are significant for both men and women, they are significantly stronger for women in the case of ability belonging.

is different. Whereas the aforementioned research assessed belonging in a particular major or course, we assessed belonging at the broader pSTEM level, and research shows that the level at which belonging is measured is important (Freeman et al., 2007). That different patterns emerged for social and ability belonging further bolsters the importance of assessing these separate subcomponents of belonging in future research, as well differentiating between belonging in the specific classroom versus the broader field overall (pSTEM).

Consistent with past research, social belonging was strongly related to intentions to persist in pSTEM for both men and women, over and above self-efficacy and identification. In line with the vulnerability hypothesis, and of particular importance to explaining gender disparities in pSTEM, we found that social belonging at the end of the semester was more tightly related to women's intentions to persist than men's intentions to persist, even after accounting for powerful covariates. This replicates prior research showing that social belonging is particularly important to women's pSTEM persistence (Lewis et al., 2017). This study allowed us to test whether ability belonging was also particularly important for women's intentions to persist; although the correlation between ability belonging and intentions to persist in pSTEM was directionally stronger for women than men, it was not significantly different. Additionally, although there were mean gender gaps in identification and self-efficacy (with women expressing greater identification but lower self-efficacy), their relationship with intentions to persist in pSTEM was the same for men and women.

This research reinforces the critical role of social belonging in pSTEM pursuits for women and suggests that even when men and women report the same level of social belonging, this level may nevertheless be inadequate for women in terms of translating into actual persistence. A greater sensitivity to belonging experiences suggests that even if women experience the same average level of belonging as men, this may still not be "good enough" for women to be convinced that they belong. Women in pSTEM whose social or ability belonging needs go unmet may consider pursuing other fields where they anticipate experiencing greater social (Murphy and Zirkel, 2015) and ability belonging (Smith et al., 2013)—likely fields where there are more women (Thoman et al., 2014). Furthermore, understanding the extent to which both social and ability belonging matter for women more than men is important for informing interventions aimed at retaining more women in pSTEM. Whereas some interventions focus on bolstering social belonging (Walton et al., 2015), others focus on increasing ability belonging, typically by normalizing hard work (Smith et al., 2013) or the experience of academic struggle (Lin-Siegler et al., 2016).

Limitations

Like most survey research, students opted to partake in the survey, and we cannot know whether results would differ if non-responders were included. Furthermore, Little's (1988) MCAR test indicated that data were not missing completely at random. Given the longitudinal nature of the data, it is not surprising that not all subjects provided full data at each time point, and further, we handled missing data by using analytical techniques robust to missing data (i.e., FIML), even if data were not missing at random. Although we drew students from a total of six different pSTEM professors teaching two different courses, these students do not represent all pSTEM students at the university or at other universities. This research was correlational in nature, and thus, we cannot draw causal inferences about the relationship between belonging and intentions to persist. Future research should investigate whether, for example, interventions aimed at improving social and ability belonging may be particularly beneficial to women in pSTEM relative to men. We also measured intentions to persist in pSTEM pursuits rather than actual persistence. However, intentions measured at the same level of specificity as the behavior of interest are widely viewed as the most proximal predictor of actual behavior (Ajzen, 1985, 1987, 1991, 2011). Intentions are, therefore, often used to assess educational outcomes (e.g., Tinto, 1997; Murphy et al., 2007; Good et al., 2012; Lewis et al., 2017; Ito and McPherson, 2018).

Future Research and Implications for Educational Practices

That qualified men and women who are entirely capable of success in pSTEM fields may nevertheless drop out due to feeling as though they do not "fit" socially or intellectually is a waste of intellectual talent. Attracting and retaining more women in pSTEM would not only supply a deficient workforce but also better address the needs of a diverse population (Blickenstaff, 2005) and potentially enhance the innovation, creativity, and quality of science produced (Hill et al., 2010; Hoever et al., 2012; Nielsen et al., 2017). It is our hope that future research examines how to boost each of these distinct types of belonging. For social belonging, interventions could entail creating inclusive environments that affirm women's sense of social connection with peers. Such environments may strategically place women with female role models (Dennehy and Dasgupta, 2017), remove reminders of masculine stereotypes or culture (Cheryan et al., 2009), or attempt to place more than one woman in small work groups (Dasgupta et al., 2015; Grover et al., 2017). Regarding ability belonging, pSTEM environments should attempt to emphasize effort and hard work over brilliance and innate talent (Smith et al., 2013; Lin-Siegler et al., 2016). Messages that normalize the struggle and journey to find social and ability belonging—particularly among dominant group members—would likely also benefit students who are questioning whether their experience is "normal" (Walton et al., 2015). Notably, this approach may be in direct opposition to the competitive "weed-out" cultures that have been described as commonplace in introductory pSTEM courses (Seymour and Hewitt, 1997; Shapiro and Sax, 2011).

Research shows that anticipated belonging plays a key role in decisions about whether or not to pursue pSTEM (Cheryan et al., 2009) and that it may be a more important criterion for women's pursuit of a field than men's (McPherson et al., 2018a). This suggests that future research is needed to address the relationship between belonging and attraction to pSTEM fields, and to examine whether these relationships also depend on gender. For example, perhaps anticipated belonging in pSTEM is not only lower among women than men but also consistent with women's greater focus on communal goals (Diekman et al., 2010), women may also weigh anticipated belonging more than men when selecting a major or a career (McPherson et al., 2018a).

On the theoretical level, it would also be interesting in future research to further consider the relation of different aspects of fit. Schmader and Sedikides (2018) have recently suggested that multiple aspects of fit all contribute to a sense of authenticity, a gestalt feeling of having one's identity align with the environment, suggesting that social and ability belonging may relate to a superordinate construct of authenticity or general belonging.

Although the present research focuses on the greater attrition of women than men from pSTEM fields, it is important to keep in mind that the gender differences observed here were not of kind but of degree—social and ability belonging were related to pSTEM persistence for women and men alike, suggesting that interventions aimed at boosting either of these should benefit both genders. It is becoming increasingly clear that retaining more talented men and women within pSTEM fields will require creating socially and intellectually welcoming environments in which students feel as though they not only are socially accepted by their peers but also have the same intellectual capacity as their peers.

DATA AVAILABILITY STATEMENT

The datasets for this article are not publicly available, because they contain educational records. Requests to access the datasets should be directed to TI at tiffany.ito@colorado.edu.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by CU Boulder Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TI and KL designed and implemented the study. KL oversaw the data collection. SB and KL performed the data analyses. SB, TI, and KL wrote the manuscript.

FUNDING

This research was supported by the National Science Foundation grant HRD-550840. Publication of this article was funded by the University of Colorado Boulder Libraries Open Access Fund.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02386/full#supplementary-material

REFERENCES

Ajzen, I. (1985). "From intentions to actions: a theory of planned behavior," in *Action—Control: From Cognition to Behavior*, eds J. Kuhi, and J. Beckmann, (Heidelberg: Springer), 11–39. doi: 10.1007/978-3-642-69746-3_2

Ajzen, I. (1987). "Attitudes, traits, and actions: dispositional prediction of behavior in personality and social psychology," in *Advances in Experimental Social Psychology*, Vol. 20, ed. L. Berkowitz, (New York, NY: Academic Press), 1–63. doi: 10.1016/s0065-2601(08)60411-6

Ajzen, I. (1991). The theory of planned behavior. *Org. Behav. Hum. Decis. Process.* 50, 179–211.

Ajzen, I. (2011). "Behavioral interventions: design and evaluation guided by the theory of planned behavior," in *Social Psychology for Program and Policy Evaluation*, eds M. M. Mark, S. I. Donaldson, and B. C. Campbell, (New York, NY: Guilford), 74–100.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychol. Rev.* 84, 191–215. doi: 10.1037//0033-295x.84.2.191

Baraldi, A. N., and Enders, C. K. (2010). An introduction to modern missing data analyses. *J. Sch. Psychol.* 48, 5–37. doi: 10.1016/j.jsp.2009.10.001

Barthelemy, R. S., McCormick, M., and Henderson, C. (2016). Gender discrimination in physics and astronomy: graduate student experiences of sexism and gender microaggressions. *Phys. Rev. Phys. Educ. Res.* 12, 1–14.

Baumeister, R. F., and Leary, M. R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol. Bull.* 117, 497–529. doi: 10.1037//0033-2909.117.3.497

Bean, J. P. (1982). Student attrition, intentions, and confidence: interaction effects in a path model. *Res. High. Educ.* 17, 291–320. doi: 10.1007/bf00977899

Besterfield-Sacre, M., Moreno, M., Shuman, L. J., and Atman, C. J. (2001). Gender and ethnicity differences in freshmen engineering student attitudes: a cross-institutional study. *J. Eng. Educ.* 90, 477–489. doi: 10.1002/j.2168-9830.2001.tb00629.x

Betz, N. E., and Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *J. Vocat. Behav.* 23, 329–345. doi: 10.1016/0001-8791(83)90046-5

Blickenstaff, J. (2005). Women and science careers: leaky pipeline or gender filter? *Gender Educ.* 17, 369–386. doi: 10.1080/09540250500145072

Bracken, B. A. (1992). *Multidimensional Self Concept Scale*. Austin, TX: PRO-ED.

Bracken, B. A. (1996). "Clinical applications of a context-dependent, multidimensional model of self-concept," in *Handbook of Self-Concept: Developmental, Social, and Clinical Considerations*, ed. B. A. Bracken, (New York, NY: John Wiley & Sons, Inc.), 463–504.

Cabrera, A. F., Nora, A., and Castañeda, M. B. (1993). College persistence: structural equations modeling test of an integrated model of student retention. *J. High. Educ.* 64, 123–139. doi: 10.1080/00221546.1993.11778419

Cacioppo, J. T., and Cacioppo, S. (2014). Social relationships and health: the toxic effects of perceived social isolation. *Soc. Pers. Psychol. Comp.* 8, 58–72. doi: 10.1111/spc3.12087

Ceci, S. J., Williams, W. M., and Barnett, S. M. (2009). Women's under-representation in science: sociocultural and biological considerations. *Psychol. Bull.* 135, 218–261. doi: 10.1037/a0014412

Chamberlain, A., and Jayaraman, J. (2017). The pipeline problem: how college majors contribute to the gender pay gap. *Glassdoor* 1–31. Available at: https://www.glassdoor.com/research/app/uploads/sites/2/2017/04/FULL-STUDY-PDF-Gender-Pay-Gap2FCollege-Major.pdf

Chemers, M. M., Zurbriggen, E. L., Syed, M., Goza, B. K., and Bearman, S. (2011). The role of efficacy and identity in science career commitment among underrepresented minority students. *J. Soc. Issue.* 67, 469–491. doi: 10.1111/j.1540-4560.2011.01710.x

Chen, X. (2013). *STEM Attrition: College Students' Paths into and out of STEM Fields (NCES 2014-001)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences.

Cheryan, S., Drury, B. J., and Vichayapai, M. (2013). Enduring influence of stereotypical computer science role models on women's academic aspirations. *Psychol. Women Q.* 37, 72–79. doi: 10.1177/0361684312459328

Cheryan, S., and Plaut, V. C. (2010). Explaining underrepresentation: a theory of precluded interest. *Sex Roles* 63, 475–488. doi: 10.1007/s11199-010-9835-x

Cheryan, S., Plaut, V. C., Davies, P. G., and Steele, C. M. (2009). Ambient belonging: how stereotypical cues impact gender participation in computer science. *J. Pers. Soc. Psychol.* 97, 1045–1060. doi: 10.1037/a0016239

Cheryan, S., Ziegler, S. A., Montoya, A. K., and Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychol. Bull.* 143, 1–35. doi: 10.1037/bul0000052

Correll, S. J. (2001). Gender and the career choice process: the role of biased self-assessments. *Am. J. Sociol.* 106, 1691–1730. doi: 10.1086/321299

Correll, S. J. (2004). Constraints into preferences: gender, status, and emerging career aspirations. *Am. Sociol. Rev.* 69, 93–113. doi: 10.1177/000312240406900106

Cundiff, J. L., Vescio, T. K., Loken, E., and Lo, L. (2013). Do gender–science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Soc. Psychol. Educ.* 16, 541–554. doi: 10.1007/s11218-013-9232-8

Curran, P. J., Obeidat, K., and Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *J. Cogn. Dev.* 11, 121–136. doi: 10.1080/15248371003699969

Daempfle, P. A. (2003). An analysis of the high attrition rates among first year college science, math, and engineering majors. *J. College Stud. Retent.* 5, 37–52. doi: 10.2190/dwqt-tya4-t20w-rcwh

Dasgupta, N., Scircle, M. M., and Hunsinger, M. (2015). Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4988–4993. doi: 10.1073/pnas.1422822112

Davidson, W. B., Beck, H. P., and Milligan, M. (2009). The college persistence questionnaire: development and validation of an instrument that predicts student attrition. *J. College Stud. Dev.* 50, 373–390. doi: 10.1353/csd.0.0079

Deiglmayr, A., Stern, E., and Schubert, R. (2019). Beliefs in "Brilliance" and belonging uncertainty in male and female STEM students. *Front. Psychol.* 10:1114. doi: 10.3389/fpsyg.2019.01114

Dennehy, T. C., and Dasgupta, N. (2017). Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proc. Natl. Acad. Sci. U.S.A.* 114, 5964–5969. doi: 10.1073/pnas.1613117114

Diekman, A. B., Brown, E. R., Johnston, A. M., and Clark, E. K. (2010). Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol. Sci.* 21, 1051–1057. doi: 10.1177/0956797610377342

Eccles, J. S. (1994). Understanding women's educational and occupational choices. *Psychol. Women Q.* 18, 585–609. doi: 10.1111/j.1471-6402.1994.tb01049.x

Eddy, S. L., and Brownell, S. E. (2016). Beneath the numbers: a review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. *Phys. Rev. Phys. Educ. Res.* 12:020106.

Ellis, J., Fosdick, B. K., and Rasmussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: lack of mathematical confidence a potential culprit. *PLoS One* 11:e0157447. doi: 10.1371/journal.pone.0157447

Else-Quest, N. M., Hyde, J. S., and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychol. Bull.* 136, 103–127. doi: 10.1037/a0018053

Freeman, T. M., Anderman, L. H., and Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *J. Exp. Educ.* 75, 203–220. doi: 10.3200/jexe.75.3.203-220

Gonsalves, A. J., Danielsson, A., and Pettersson, H. (2016). Masculinities and experimental practices in physics: the view from three case studies. *Phys. Rev. Phys. Educ. Res.* 12:020120.

Good, C., Rattan, A., and Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *J. Pers. Soc. Psychol.* 102, 700–717. doi: 10.1037/a0026659

Goodenow, C. (1993). Classroom belonging among early adolescent students: relationships to motivation and achievement. *J. Early Adolesc.* 13, 21–43. doi: 10.1177/0272431693013001002

Grover, S. S., Ito, T. A., and Park, B. (2017). The effects of gender composition on women's experience in math work groups. *J. Pers. Soc. Psychol.* 112, 877–900. doi: 10.1037/pspi0000090

Hausmann, L. R., Ye, F., Schofield, J. W., and Woods, R. L. (2009). Sense of belonging and persistence in White and African American first-year students. *Res. High. Educ.* 50, 649–669. doi: 10.1007/s11162-009-9137-8

Hazari, Z., Sonnert, G., Sadler, P. M., and Shanahan, M. C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: a gender study. *J. Res. Sci. Teach.* 47, 978–1003.

Hill, C., Corbett, C., and St. Rose, A. (2010). *Why so few? Women in Science, Technology, Engineering, and Mathematics*. Washington, DC: AAUW.

Hoever, I. J., Van Knippenberg, D., Van Ginkel, W. P., and Barkema, H. G. (2012). Fostering team creativity: perspective taking as key to unlocking diversity's potential. *J. Appl. Psychol.* 97, 982–996. doi: 10.1037/a0029159

Höhne, E., and Zander, L. (2019). Sources of male and female students' belonging uncertainty in the computer sciences. *Front. Psychol.* 10:1740. doi: 10.3389/fpsyg.2019.01740

Holleran, S. E., Whitehead, J., Schmader, T., and Mehl, M. R. (2011). Talking shop and shooting the breeze: a study of workplace conversation and job disengagement among STEM faculty. *Soc. Psychol. Pers. Sci.* 2, 65–71. doi: 10.1177/1948550610379921

Huang, G., Taddese, N., and Walter, E. (2000). *Entry and Persistence of Women and Minorities in College Science and Engineering Education (NCES 2000-601)*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Hughes, J. N., Im, M. H., and Allee, P. J. (2015). Effect of school belonging trajectories in grades 6–8 on achievement: gender and ethnic differences. *J. Sch. Psychol.* 53, 493–507. doi: 10.1016/j.jsp.2015.08.001

Ito, T. A., and McPherson, E. (2018). Factors influencing high school students' interest in pSTEM. *Front. Psychol.* 9:1535. doi: 10.3389/fpsyg.2018.01535

Johnson, D. R., Soldner, M., Leonard, J. B., Alvarez, P., Inkelas, K. K., Rowan-Kenyon, H. T., et al. (2007). Examining sense of belonging among first-year undergraduates from different racial/ethnic groups. *J. College Stud. Dev.* 48, 525–542. doi: 10.1353/csd.2007.0054

Judd, C. M., McClelland, G. H., and Ryan, C. S. (2017). *Data Analysis: A Model Comparison Approach to Regression, ANOVA, and Beyond*. Abingdon: Routledge.

Kline, R. B. (2011). *Principles and Practice of Structural Equation Modeling*. New York, NY: Guildford.

Lent, R. W., Brown, S. D., and Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *J. Couns. Psychol.* 33, 265–269. doi: 10.1037//0022-0167.33.3.265

Leslie, S. J., Cimpian, A., Meyer, M., and Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347, 262–265. doi: 10.1126/science.1261375

Lewis, K. L., and Hodges, S. D. (2015). Expanding the concept of belonging in academic domains: development and validation of the ability uncertainty scale. *Learn. Individ. Differ.* 37, 197–202. doi: 10.1016/j.lindif.2014.12.002

Lewis, K. L., Stout, J. G., Finkelstein, N. D., Pollock, S. J., Miyake, A., Cohen, G. L., et al. (2017). Fitting in to move forward: using a belonging framework to understand gender disparities in persistence in the physical sciences, technology, engineering, and mathematics (pSTEM). *Psychol. Women Q.* 41, 420–436. doi: 10.1177/0361684317720186

Lewis, K. L., Stout, J. G., Pollock, S. J., Finkelstein, N. D., and Ito, T. A. (2016). Fitting in or opting out: a review of key social-psychological factors influencing a sense of belonging for women in physics. *Phys. Rev. Phys. Educ. Res.* 12, 1–10.

Lin-Siegler, X., Ahn, J. N., Chen, J., Fang, F. F. A., and Luna-Lucero, M. (2016). Even einstein struggled: effects of learning about great scientists' struggles on high school students' motivation to learn science. *J. Educ. Psychol.* 108, 314–328. doi: 10.1037/edu0000092

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *J. Am. Stat. Assoc.* 83, 1198–1202. doi: 10.1080/01621459.1988.10478722

Luke, C., Redekop, F., and Burgin, C. (2015). Psychological factors in community college student retention. *Community College J. Res. Pract.* 39, 222–234. doi: 10.1080/10668926.2013.803940

Markus, H., and Wurf, E. (1987). The dynamic self-concept: a social psychological perspective. *Ann. Rev. Psychol.* 38, 299–337. doi: 10.1146/annurev.psych.38.1.299

McPherson, E., Banchefsky, S., and Park, B. (2018a). Using social psychological theory to understand choice of a pSTEM academic major. *Educ. Psychol.* 38, 1278–1301. doi: 10.1080/01443410.2018.1489526

McPherson, E., Park, B., and Ito, T. A. (2018b). The role of prototype matching in science pursuits: perceptions of scientists that are inaccurate and diverge from self-perceptions predict reduced interest in a science career. *Pers. Soc. Psychol. Bull.* 44, 881–898. doi: 10.1177/0146167217754069

Meyer, M., Cimpian, A., and Leslie, S. J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Front. Psychol.* 6:235. doi: 10.3389/fpsyg.2015.00235

Multon, K. D., Brown, S. D., and Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: a meta-analytic investigation. *J. Couns. Psychol.* 38, 30–38. doi: 10.1037//0022-0167.38.1.30

Murphy, M. C., Steele, C. M., and Gross, J. J. (2007). Signaling threat: how situational cues affect women in math, science, and engineering settings. *Psychol. Sci.* 18, 879–885. doi: 10.1111/j.1467-9280.2007.01995.x

Murphy, M. C., and Zirkel, S. (2015). Race and belonging in school: how anticipated and experienced belonging affect choice, persistence, and performance. *Teach. College Record* 117, 1–40.

National Science Foundation and National Center for Science and Engineering Statistics (2015). *Integrated Postsecondary Education Data System, Completions Survey. Table 5-2. Bachelor's Degrees Awarded, by Field and Sex: 2004–14 (Special Tabulations of U.S. Department of Education, National Center for Edu- cation Statistics)*. Available at: https://www.nsf.gov/statistics/2015/nsf15311/tables.cfm (accessed January 8, 2017).

Nielsen, M. W., Alegria, S., Börjeson, L., Etzkowitz, H., Falk-Krzesinski, H. J., Joshi, A., et al. (2017). Opinion: gender diversity leads to better science. *Proc. Natl. Acad. Sci. U.S.A.* 114, 1740–1742. doi: 10.1073/pnas.1700616114

Ohland, M. W., Sheppard, S. D., Lichtenstein, G., Eris, O., Chachra, D., and Layton, R. A. (2008). Persistence, engagement, and migration in engineering programs. *J. Eng. Educ.* 97, 259–278. doi: 10.1002/j.2168-9830.2008.tb00978.x

Osborne, J. W., and Jones, B. D. (2011). Identification with academics and motivation to achieve in school: how the structure of the self influences academic outcomes. *Educ. Psychol. Rev.* 23, 131–158. doi: 10.1007/s10648-011-9151-1

President's Council of Advisors on Science, and Technology [PCAST] (2012). *Engage to Excel: Producing One Million Additional College Graduates With Degrees in Science, Technology, Engineering, and Mathematics*. Washington, DC: Author.

Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *J. Stat. Softw.* 48, 1–36. doi: 10.3389/fpsyg.2014.01521

Salmerón Gómez, R., García Pérez, J., López Martín, M. D. M., and García, C. G. (2016). Collinearity diagnostic applied in ridge estimation through the variance inflation factor. *J. Appl. Stat.* 43, 1831–1849. doi: 10.1080/02664763.2015.1120712

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *J. Exp. Soc. Psychol.* 38, 194–201. doi: 10.1177/0146167208318404

Schmader, T., and Sedikides, C. (2018). State authenticity as fit to environment: the implications of social identity for fit, authenticity, and self-segregation. *Pers. Soc. Psychol. Rev.* 22, 228–259. doi: 10.1177/1088868317734080

Settles, I. H., Cortina, L. M., Malley, J., and Stewart, A. J. (2006). The climate for women in academic science: the good, the bad, and the changeable. *Psychol. Women Q.* 30, 47–58. doi: 10.1111/j.1471-6402.2006.00261.x

Seymour, E., and Hewitt, N. (1997). *Talking About Leaving: Why Undergraduates Leave the Sciences*. Boulder, CO: Westview Press.

Shapiro, C. A., and Sax, L. J. (2011). Major selection and persistence for women in STEM. *New Direct. Institut. Res.* 2011, 5–18. doi: 10.1002/ir.404

Sikora, J., and Pokropek, A. (2012). Gender segregation of adolescent science career plans in 50 countries. *Sci. Educ.* 96, 234–264. doi: 10.1002/sce.20479

Smith, J. L., Lewis, K. L., Hawthorne, L., and Hodges, S. D. (2013). When trying hard isn't natural women's belonging with and motivation for male-dominated stem fields as a function of effort expenditure concerns. *Pers. Soc. Psychol. Bull.* 39, 131–143. doi: 10.1177/0146167212468332

Smith, J. L., and White, P. H. (2001). Development of the domain identification measure: a tool for investigating stereotype threat effects.

Educ. Psychol. Measure. 61, 1040–1057. doi: 10.1177/00131640121971635

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *Am. Psychol.* 60, 950–958. doi: 10.1037/0003-066x.60.9.950

Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women's math performance. *J. Exp. Soc. Psychol.* 35, 4–28.

Stephens-Davidowitz, S. (2014). *Google, tell me. Is my son a genius? The New York Times*. Available at: https://www.nytimes.com/2014/01/19/opinion/sunday/google-tell-me-is-my-son-a-genius.html?mcubz=3 (accessed January 8, 2017).

Stout, J. G., Dasgupta, N., Hunsinger, M., and McManus, M. A. (2011). STEMing the tide: using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *J. Pers. Soc. Psychol.* 100, 255–270. doi: 10.1037/a0021385

Tellhed, U., Bäckström, M., and Björklund, F. (2017). Will I fit in and do well? The importance of social belongingness and self-efficacy for explaining gender differences in interest in STEM and HEED majors. *Sex Roles* 77, 86–96. doi: 10.1007/s11199-016-0694-y

Thoman, D. B., Arizaga, J. A., Smith, J. L., Story, T. S., and Soncuya, G. (2014). The grass is greener in non-science, technology, engineering, and math classes: examining the role of competing belonging to undergraduate women's vulnerability to being pulled away from science. *Psychol. Women Q.* 38, 246–258. doi: 10.1177/0361684313499899

Thompson, C. G., Kim, R. S., Aloe, A. M., and Becker, B. J. (2017). Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. *Basic Appl. Soc. Psychol.* 39, 81–90. doi: 10.1080/01973533.2016.1277529

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *J. Educ. Psychol.* 92, 144–151. doi: 10.1037//0022-0663.92.1.144

Tinto, V. (1997). Classrooms as communities: exploring the educational character of student persistence. *J. High. Educ.* 68, 599–623. doi: 10.2307/2959965

Usher, E. L., and Pajares, F. (2008). Sources of self-efficacy in school: critical review of the literature and future directions. *Rev. Educ. Res.* 78, 751–796. doi: 10.3102/0034654308321456

Vogt, C. M., Hocevar, D., and Hagedorn, L. S. (2007). A social cognitive construct validation: determining women's and men's success in engineering programs. *J. High. Educ.* 78, 337–364. doi: 10.1353/jhe.2007.0019

Voorhees, R. A. (1987). Toward building models of community college persistence: a logit analysis. *Res. High. Educ.* 26, 115–129. doi: 10.1007/bf00992024

Walton, G. M., and Cohen, G. L. (2007). A question of belonging: race, social fit, and achievement. *J. Pers. Soc. Psychol.* 92, 82–96. doi: 10.1037/0022-3514.92.1.82

Walton, G. M., and Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* 331, 1447–1451. doi: 10.1126/science.1198364

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., and Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *J. Educ. Psychol.* 107, 468–485. doi: 10.1037/a0037461

Wilson, D., Jones, D., Bocell, F., Crawford, J., Kim, M. J., Veilleux, N., et al. (2015). Belonging and academic engagement among undergraduate STEM students: a multi-institutional study. *Res. High. Educ.* 56, 750–776. doi: 10.1007/s11162-015-9367-x

Wolf, E. J., Harrington, K. M., Clark, S. L., and Miller, M. W. (2013). Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. *Educ. Psychol. Measure.* 73, 913–934. doi: 10.1177/0013164413495237

Wooldridge, J. (2013). *Introductory Econometrics*, 5th Edn. Boston, MA: Cengage Learning.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

OPINION
published: 16 August 2019
doi: 10.3389/fpsyg.2019.01844

It Takes More Than One Swallow to Make a Summer: Measures to Foster Girls' and Women's Pathways Into STEM

Silke Luttenberger[1], Petra Steinlechner[2], Bernhard Ertl[3] and Manuela Paechter[2]*

[1] Institute in Early Childhood and Primary Teacher Education, University College of Teacher Education Styria, Graz, Austria, [2] Educational Psychology, Institute of Psychology, University of Graz, Graz, Austria, [3] Learning and Teaching with Media, Department of Education, Universität der Bundeswehr München, Neubiberg, Germany

Keywords: gender, STEM, self-efficacy, interest, career choice, career counseling, STEM teaching

OPEN ACCESS

Edited by:
Douglas F. Kauffman,
Medical University of the Americas—Nevis, United States

Reviewed by:
Michael D. Toland,
University of Kentucky, United States

***Correspondence:**
Silke Luttenberger
silke.luttenberger@phst.at

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: 02 January 2019
Accepted: 26 July 2019
Published: 16 August 2019

Citation:
Luttenberger S, Steinlechner P, Ertl B and Paechter M (2019) It Takes More Than One Swallow to Make a Summer: Measures to Foster Girls' and Women's Pathways Into STEM. Front. Psychol. 10:1844. doi: 10.3389/fpsyg.2019.01844

THEORY TIES RESEARCH AND PRACTICAL INTERVENTIONS TOGETHER

For decades, the proportion of women in STEM professions (Science, Technology, Engineering, Mathematics) has remained at approximately one fourth in the European Union – an alarmingly low number (Center of Excellence Women and Science, 2014). With labor markets continuing to communicate an increasing need in STEM workforces, this low number signals unfulfilled talent that is otherwise greatly needed in many critical fields. Effective interventions are needed (Walsh and Heppner, 2006) to foster girls' and women's pathways into STEM. Yet, when it comes to the implementation of interventions and their effectiveness, current efforts leave a lot to be desired. The present article describes how girls and women can be encouraged to consider STEM professions as real options.

There are many situations where women lose interest or fail to build up interest in STEM over their formative years from early childhood to school and tertiary education. The Social Cognitive Career Theory (Lent et al., 1994; Lent and Brown, 2019) stresses key variables for the development and realization of career interest and goals (**Figure 1**). Important personal factors are STEM self-efficacy and outcome expectations for entering a STEM career. These factors are related to STEM interest, which in turn may lead to STEM career goals. "Building self-efficacy for math and science and fostering positive and realistic outcome expectations would lead to realistic and investigative interests, that would, in turn, lead to STEM career goals and preparation for, and entry into, a STEM occupation" (Fouad and Santana, 2017, p. 27).

However, pathways into STEM careers are not only related to personal factors. Structural or social factors may work as barriers and filter out girls and women from STEM careers (Watt et al., 2006; Turner et al., 2019).

The present article investigates closer how girls and women can be supported in the formation of interest in a field, to a certain career goal, to a specific choice of action.

STEM PATHWAYS: FROM INTEREST TO A CAREER GOAL AND CHOICE OF ACTION

Interest

Positive STEM experiences in school are a key to the development of interest and career goals in STEM (Fouad et al., 2010; Ertl et al., 2017; Luttenberger et al., 2019). Ideally, they should raise interest as well as self-efficacy in STEM (compare **Figure 1**).

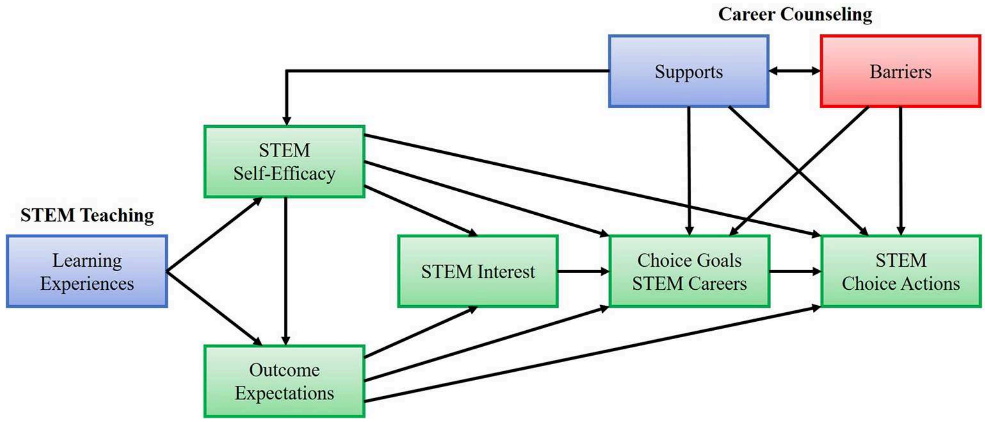


FIGURE 1 | Social Cognitive Career Theory: Interest and Choice Model.

Intervention studies concerning STEM teaching found positive effects on students' interest and motivation with measures such as inquiry-based STEM teaching (UNESCO, 2017), improving teaching using STEM pedagogy (Gaspard et al., 2015), bringing real-life applications into the classroom (Taskinen et al., 2013), hands-on activities (Lee and Erdogan, 2007), or design-based learning with laboratory and workshop experiences (Vongkulluksn et al., 2018). Furthermore, informal learning environments such as STEM summer camps were an effective means to spark middle-school students' interest in STEM (Mohr-Schroeder et al., 2014). In a meta-analysis by Furtak et al. (2012), inquiry-based STEM teaching was found to be highly effective for learning and the development of interest.

It seems to be important that interventions to foster interest and raise career aspirations start early, i.e., in primary education (even though to date, programs have been mainly implemented starting in middle school). Primary school students were the focus of two intervention studies: Here, design-based learning via Makerspaces (Vongkulluksn et al., 2018) was able to spark students' interest in STEM activities. However, even though self-efficacy and interest remained moderately high, they declined over the Makerspace intervention semester. One of the few studies which combined indirect (learning experiences and STEM) and direct interventions (career counseling) was carried out by Panayiotou and Eteokleous-Grigorio (2017) with a robotic course as didactic intervention. Although it found increases in STEM interest, positive attitudes, and motivation, no effects on career aspirations were identified.

Altogether, STEM teaching has to be designed carefully to raise not only the interest in STEM, but also in STEM careers (Fouad et al., 2010). Of note is the general lack of intervention studies which combine STEM teaching and career counseling.

Career Goals

Career counseling is assumed to have a direct impact on students' career choices because students will only strive for professions they are aware of (Herr et al., 2004).

Students generally require more information about STEM professions, and have to actively search for job-related information. In an intervention study by Turner and Lapan (2002), a computerized training for middle school students was developed to foster interest and self-efficacy in STEM professions for females. A strength of this study lies in its discussion of the training results for career counseling. It was possible to identify girls who are in fact interested in a STEM career but lack social support. Yet, only short-term gains were measured after one week in this study; its long-term effects remain unclear. Painter et al. (2006) focused on raising students' interest in scientific careers in STEM. They found positive effects of contextualized science materials and interactions with scientists among 7th and 10th graders (about 10% reported having previously interacted with scientists at school).

All in all, career counseling is essential for informing students about career choices. It should aim at the development of realistic expectations about STEM professions that match individual interest.

Specific Choice of Action

Teachers play a crucial role in supporting students on the pathway from career goals to choice of actions. Career teachers' lack of knowledge often prevents the choice of STEM professions (Cleaves, 2005). The higher teachers' encouragement and help when needed, the higher the motivation of girls and women to explore STEM careers (Blustein et al., 2013). In general, girls and women experience less support to develop and pursue STEM-related career goals. Also parental beliefs and stereotypes in particular can support or hinder career choices in STEM (Ertl et al., 2017).

Social Cognitive Career Theory (Lent et al., 1994) points at the importance of role models. A lack of female role models (family members, peers etc.) can decrease the sense of belonging in STEM (Blickenstaff, 2005). In an intervention study by Robnett et al. (2018) instrumental and socio-emotional mentoring were able to foster women's sense of belonging to a STEM community.

STARTING EARLY: THE IMPORTANCE OF EARLY CAREER-RELATED LEARNING EXPERIENCES FOR STEM PATHWAYS

Most interventions focus on students in upper secondary education from age fourteen onwards (DeWitt and Archer, 2015), even though career choices are unlikely to change dramatically by this age. Career aspirations mostly have already formed by the age of 13. After this, it is increasingly difficult to interest students in STEM (Lindahl, 2007). Therefore, the critical age period during which aspirations are formed is during primary and lower secondary education (Lindahl, 2007; DeWitt and Archer, 2015).

In the first phase of career orientation, interest in both STEM learning and STEM aspirations are still unstable (Ardies et al., 2015). Interest in STEM typically shows a downward trend from primary school on (Taskinen et al., 2013). Interventions in primary education focus mainly on STEM teaching and learning (Panayiotou and Eteokleous-Grigorio, 2017; Vongkulluksn et al., 2018) or informal learning experiences by using authentic STEM workplaces (Roberts et al., 2018). They show that learning experiences are needed at an early age to support the transition from career interest to choice of goals. Girls who aspire to a STEM career as early as primary school are more likely to choose a STEM profession (Schoon, 2001). However, the relationship between interest and career aspirations has seldom been the focus of interventions. This is why there is a need for studies providing advice on how to foster not only interest in STEM, but STEM career aspirations as well (Panayiotou and Eteokleous-Grigorio, 2017).

A problem with many intervention studies is that they often appeal only to those students who are already interested in STEM, and not to those who are skeptical about these fields. Interventions should aim at all girls and students, at interested students, as well as not-so-interested ones. They should start at an early age, aim to raise and sustain interest, and transform it into career goals and choices of action. Real-life experiences with STEM, e.g., hands-on experiences, apprenticeships, career counseling, and role models can expand girls' knowledge about STEM and professions while maintaining effective levels of interest (UNESCO, 2017).

FOCUSING NOT ONLY ON PERSONAL FACTORS BUT ALSO ON REMOVING EXTERNAL BARRIERS

Social Cognitive Career Theory provides an empirical basis for interventions to foster STEM interest and goals (Lent et al., 2018). Interventions and attempts in education to foster the proportion of women in STEM mostly focus on personal factors, e.g., self-efficacy, outcome expectations, interest, or STEM belonging. Social Cognitive Career Theory also stresses the importance of social factors. Turner and Lapan (2002) showed that parents may either support their daughters on a pathway into STEM or put barriers along it. In the same way, not only parents, but other family members, peers, teachers as well as future employers may offer support or create barriers. The identification of barriers and support is important for transferring interest into choices of action and promoting females' participation in STEM (Fouad et al., 2010). Contextual and social factors play different valuable roles, indirectly and directly, in fostering women's STEM interest and goals (Lent et al., 2018).

CONCLUSION

There are a multitude of factors that influence the career paths of girls and women. Fostering girls' pathways into STEM requires continuous and multiple interventions that start at an early age and address personal as well as social factors. As Social Cognitive Career Theory points out, they should take into account key variables in the development and realization of career wishes, the formation of interest in a field, and the formation of career goals to coincide with specific choices of action.

AUTHOR CONTRIBUTIONS

SL, PS, BE, and MP have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Ardies, J., De Maeyer, S., and Gijbels, D. (2015). A longitudinal study on boys' and girls' career aspirations and interest in technology. *Res. Sci. Technol. Educ.* 33, 366–386. doi: 10.1080/02635143.2015.1060412

Blickenstaff, J. (2005). Women and science careers: leaky pipeline or gender filter? *Gender Educ.* 17, 369–386. doi: 10.1080/09540250500145072

Blustein, D. L., Barnett, M., Mark, S., Depot, M., Lovering, M., Lee, Y., et al. (2013). Examing urban students' constructions of a STEM/career development intervention over time. *J. Career Dev.* 40, 40–67. doi: 10.1177/0894845312441680

Center of Excellence Women and Science (2014). *Studentinnenanteil in Mathematik/Naturwissenschaften und Ingenieurwissenschaften (ISCED 5-6) im internationalen Vergleich* (2011). [Proportion of female students in mathematics/sciences and engineering (ISCED 5-6) in an international comparison]. Retrieved from: http://www.gesis.org/cews/fileadmin/cews/www/statistiken/08_d.gif

Cleaves, A. (2005). The formation of science choices in secondary school. *Int. J. Sci. Educ.* 27, 471–486. doi: 10.1080/0950069042000323746

DeWitt, J., and Archer, L. (2015). Who aspires to a science career? A comparison of survey responses from primary and secondary school students. *Int. J. Sci. Educ.* 37, 2170–2192. doi: 10.1080/09500693.2015.1071899

Ertl, B., Luttenberger, S., and Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* 8:703. doi: 10.3389/fpsyg.2017.00703

Fouad, N. A., Hackett, G., Smith, P. L., Kantamneni, N., Fitzpatrick, M., Haag, S., et al. (2010). Barriers and supports for continuing in mathematics and science: gender and educational level differences. *J. Vocat. Behav.* 77, 361–373. doi: 10.1016/j.jvb.2010.06.004

Fouad, N. A., and Santana, M. C. (2017). SCCT and underrepresented populations in STEM fields: moving the needle. *J. Career Assess.* 25, 24–39. doi: 10.1177/1069072716658324

Furtak, E. M., Seidel, T., Iverson, H., and Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: a meta-analysis. *Rev. Educ. Res.* 82, 300–329. doi: 10.3102/0034654312457206

Gaspard, H., Dicke, A. L., Flunger, B., Brisson, B. M., Hafner, I., Nagengast, B., et al. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Dev. Psychol.* 51, 1226–1240. doi: 10.1037/dev0000028

Herr, E. L., Cramer, S. H., and Niles, S. G. (2004). *Career Guidance and Counseling Through the Lifespan: Systematic Approaches*, Vol. 6. Boston, MA: Pearson Education.

Lee, M. K., and Erdogan, I. (2007). The effect of science-technology-society teaching on students' attitudes toward science and certain aspects of creativity. *Int. J. Sci. Educ.* 29, 1315–1327. doi: 10.1080/09500690600972974

Lent, R. W., and Brown, S. D. (2019). Social cognitive career theory at 25: empirical status, choice, and performance models. *J. Vocat. Behav*. 115:103316. doi: 10.1016/j.jvb.2019.06.004

Lent, R. W., Brown, S. D., and Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav*. 45, 79–122. doi: 10.1006/jvbe.1994.1027

Lent, R. W., Sheu, H.-B., Miller, M. J., Cusick, M. E., Penn, L. T., and Truong, N. N. (2018). Predictors of science, technology, engineering, and mathematics choice options: a meta-analytic path analysis of the social-cognitive choice model by gender and race/ethnicity. *J. Counsel. Psychol.* 65, 17–35. doi: 10.1037/cou0000243

Lindahl, B. (2007). "A longitudinal study of student's' attitudes towards science and choice of career," in *Paper presented at the 80th NARST International Conference,* New Orleans, LA.

Luttenberger, S., Paechter, M., and Ertl, B. (2019). Self-concept and support experienced in school as key variables for the motivation of women enrolled in STEM subjects with a low and moderate proportion of females. *Front. Psychol.* 10:1242. doi: 10.3389/fpsyg.2019.01242

Mohr-Schroeder, M. J., Jackson, C., Miller, M., Walcott, B., Little, D. L., Speler, L., et al. (2014). Developing middle school students' interests in STEM via summer learning experiences: see blue STEM camp. *School Sci. Math.* 114, 291–301. doi: 10.1111/ssm.12079

Painter, J., Jones, M. G., Tretter, T. R., and Kubasko, D. (2006). Pulling back the curtain: uncovering and changing students' perception of scientists. *School Sci. Math.* 106, 181–190. doi: 10.1111/j.1949-8594.2006.tb18074.x

Panayiotou, M., and Eteokleous-Grigorio, N. (2017). "Robotics to increase students' STEM attitudes," in *Education and New Developments 2017,* ed M. Carmo (Lisbon: InScience Press), 216–219.

Roberts, O. T., Jackson, C., Mohr-Schroeder, M. J., Bush, S. B., Maiorca, C., Cavalcanti, M., et al. (2018). Students' perceptions of STEM learning after participating in a summer informal learning experience. *Int. J. STEM Educ.* 5:35. doi: 10.1186/s40594-018-0133-4

Robnett, R. D., Nelson, P. A., Zurbriggen, E. L., Crosby, F. J., and Chembers, M. M. (2018). Research mentoring and science identity: insights from undergraduates and their mentors. *Int. J. STEM Educ.* 5:41. doi: 10.1037/0022-0663.95.4.667

Schoon, I. (2001). Teenage job aspirations and career attainment in adulthood: a 17-year follow-up study of teenagers who aspired to become scientists, health professionals, or engineers. *Int. J. Behav. Dev*. 25, 124–132. doi: 10.1080/01650250042000186

Taskinen, P. H., Schütte, K., and Prenzel, M. (2013). Adolescents' motivation to select an academic science-related career: the role of school factors, individual interest, and science self-concept. *Educ. Res. Eval.* 19, 717–733. doi: 10.1080/13803611.2013.853620

Turner, S. L., Joeng, J. R., Sims, M. D., Dade, S. N., and Reid, M. F. (2019). SES, gender, and STEM career interests, goals, and actions: a test of SCCT. *J. Career Assess.* 27, 134–150. doi: 10.1177/1069072717748665

Turner, S. L., and Lapan, R. T. (2002). Career self-efficacy and perceptions of parent support in adolescent career development. *Career Dev. Quart.* 51, 44–55. doi: 10.1002/j.2161-0045.2002.tb00591.x

UNESCO (2017). *Cracking the Code: Girls' and Women's Education in Science, Technology, Engineering and Mathematics (STEM)*. Paris.

Vongkulluksn, V. W., Matewos, A. M., Sinatra, G. M., and Marsh, J. A. (2018). Motivational factors in makerspaces: a mixed methods study of elementary school students' situational interest, self-efficacy, and achievement emotions. *Int. J. STEM Educ.* 5:43. doi: 10.2307/1131888

Walsh, W. B., and Heppner, M. J. (2006). *Handbook of Career Counseling for Women*, 2nd ed. Mahwah, NJ: Lawrence Erlbaum.

Watt, H. M. G., Eccles, J. S., and Durik, A. M. (2006). The leaky mathematics pipeline for girls: a motivational analysis of high school enrolments in Australia and the USA. *Equal Opportunities Int.* 25, 642–659. doi: 10.1108/02610150610719119

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.