frontiers | Research Topics

Women in STEM Education

Edited by
Lisbet Rønningsbakk and Karen Blackmore

Published in
Frontiers in Education
Frontiers in Psychology

ISSN 1664-8714
ISBN 978-2-8325-3977-4
DOI 10.3389/978-2-8325-3977-4

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: frontiersin.org/about/contact

Women in STEM Education

Topic editors
Lisbet Rønningsbakk — UiT The Arctic University of Norway, Norway
Karen Blackmore — University of Worcester, United Kingdom

Citation
Rønningsbakk, L., Blackmore, K., eds. (2023). *Women in STEM Education*.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3977-4

Table of contents

frontiers | Frontiers in Education

TYPE Editorial
PUBLISHED 06 November 2023
DOI 10.3389/feduc.2023.1319874

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Lianghuo Fan,
East China Normal University, China

*CORRESPONDENCE
Lisbet Rønningsbakk
lisbet.ronningsbakk@uit.no

RECEIVED 11 October 2023
ACCEPTED 16 October 2023
PUBLISHED 06 November 2023

CITATION
Blackmore K and Rønningsbakk L (2023) Editorial: Women in STEM Education. *Front. Educ.* 8:1319874. doi: 10.3389/feduc.2023.1319874

Editorial: Women in STEM Education

Karen Blackmore[1] and Lisbet Rønningsbakk[2]*

[1]Department of Primary Education, Institute of Education, University of Worcester, Worcester, United Kingdom, [2]Department of Education, UiT The Arctic University of Norway, Tromsø, Norway

KEYWORDS

science technology engineering mathematics (STEM), pedagogy, inclusion, women, learning

Editorial on the Research Topic
Women in STEM Education

We are delighted to present this special edition of Women in STEM Education as a positive platform to examine insights from a range of fields including engineering, mathematics and information technology. The articles fall into two broad groups; those pertaining to the experiences and perspectives of women in STEM fields and those exploring enabling pedagogical approaches for all learners.

Learning experiences and perspectives of women in STEM fields

The first section presents papers which address perspectives of how women experience learning and education in STEM fields. The articles represent collaboration from a range of different countries and phases of education from young learners to adult students, with a strong emphasis on the importance of creating inclusive and positive STEM learning environments and communities, to meet the global needs of STEM competences.

Kim and Kim's study from South Korea documents the chilly climate perceived by female undergraduate engineering students. Using a concept-mapping method, they conceptualized and identified aspects that contribute to the chilly climate, such as inherent exclusion and alienation within the culture, lack of gender sensitivity, male-centered study situations, and indirect prejudice toward women. The paper highlights the expectations that female students experience a negative culture formed by social myths which arguably inhibits female participation in STEM.

Fernandez et al., have focused on differences in the time use in laboratory-based activities for university students. They used surveys where students reported the duration of undertaking activities and made observations of the actual time used. They found that female students spent less time in laboratories and interacting with the equipment, but took more time to observe and take notes than their male peers. However both groups were equally content with their time use. The authors suggest that females adopt roles in the laboratory which are formed by their expectations that they are less technically active and more observant than male students, a perspective that requires consideration by universities to ensure their obligations to an inclusive pedagogical stance.

Ladewig et al. questioning of gender stereotypes in the German Physics Olympiad, a science competition for German students, affords yet another lens through which to view gender differences. Using social threat theory to identify the reasons why women are under-represented and drop out earlier than men in the competition, they assumed that negative stereotypes about women's abilities in physics and the predominant male environment worked as a threat to the female gender-identity. An intervention study in the form of two weekend assemblies for the participants, where they focused on belonging and values, however, showed no significant differences between genders. The researchers suggest that female students who have a genuine interest for physics are resilient toward the negative impacts from threats against their social and gender identities and conclude that genuine interest is a strong factor for female participation within the STEM fields.

Balasubramanian et al.'s case study shows how the retention of women in STEM careers in the US can be met by programmatic innovations. Various innovations were tried out over a period of 8 years, both curricular (colloquium, laboratory) and co-curricular (community building, junior review and conference participation). They concluded that a combination of several innovations resulted in an increased number of successful female major students, increasing degree completion by 200% (over 10 years) and resulting in an average graduation rate far above national standard. Another important outcome was that 80 % of the women physics majors maintained careers within the field greater than 5 years post-graduation.

Lane et al., were also concerned about the students' participation in STEM-subjects for future competence needs. Their study aimed to identify students' attitudes toward STEM education and careers for post-primary level and beyond, in an Irish context. Their survey study revealed that female students have more positive attitudes toward physics while male students are more positive about mathematics. The authors suggest this is an interesting aspect for discussing the role of STEM subjects within the post-primary curriculum.

Effective STEM pedagogies for all

The collaborative review paper by Zhu et al. illustrates the efficacy of providing learners with both integrated and informal opportunities to develop their spatial skills from kindergarten to adulthood. It makes the case that spatial skill development can enhance complex STEM problem-solving and *vice versa*. The authors also highlight how extra-curricular involvement is a positive predictor for girls' interest and confidence in mathematics, a vital enabler for further engagement in STEM careers. Despite an abundance of literature suggesting that boys outperform girls in spatial skills there appears to be no biological reason why this should be the case, rather it appears this factor may be due girls' perceptions of societal norms. This argument gains credence when it is considered that there is evidence that when girls are given constructive feedback on their spatial ability, they improve their performance. The timing and manner of feedback are important factors, with the suggestion that accurate peer feedback can also be facilitative.

Collaborative problem-solving facilitated by peer support, features extensively in our paper (Blackmore and Rønningsbakk) exploring children's perspectives on using mobile technology during science inquiry. The technology was seen to act as a conduit for sharing approaches and results of a range of STEM investigations. Children were exceptionally positive about developing both knowledge and skills by interactions with each other, supplemented by appropriate teacher feedback and guidance. This technology enhanced learning approach also conferred the advantage of capturing in-the-moment science phenomenon in primary classrooms, arguably a significant enabler for maintaining children's innate curiosity of the world around them.

Heim and Wang explored an alternative pedagogical perspective in their paper focusing on the affordances of cross-curricular learning to support computer science programming. They endeavored to examine linkages of children's perceptions of computational thinking during mathematics and food and health study. Whilst the majority of children were not able to make explicit links between unplugged programming and following recipes, there were indications that some children understood that food production involves a step-by-step set of instructions. This approach resonates with the recommendations made by Zhu et al. in terms of providing children with rich authentic experiences, that mirror STEM problem-solving and affording opportunities for learning to be transferred from one context to another.

The learning of student teachers is also an important element to be considered when exploring effective STEM pedagogies. In the paper by Leming and Johanson examining the perceptions of digital critical competence by student-teachers, they found a range of different perspectives including an awareness for the necessity of source criticism and links with a functioning democracy. This is heartening since these professionals will act as facilitators for 21st century skill acquisition. Such findings have implications for initial teacher educators since student teachers need to be given time and opportunity to develop a secure understanding of processes to identify authentic and rigorous learning resources.

In summary, we hope this special edition will provide insights into enabling STEM learning for all and inform practice in a range of contexts to support future collaboration and innovation.

Author contributions

KB: Writing—original draft, Writing—review & editing. LR: Writing—original draft, Writing—review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We would like to thank all the authors that contributed to the Special Edition of Women in STEM for their research insights and commitment to enhancing STEM learning. In addition, we are most grateful to the reviewers for their timely constructive feedback and support of the authors and editors. Lastly we would like to thank the chief editor for their exemplary commitment to this special edition.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 29 September 2022
DOI 10.3389/fpsyg.2022.959972

OPEN ACCESS

EDITED BY
Karen Blackmore,
University of Worcester,
United Kingdom

REVIEWED BY
Sabina Ličen,
University of Primorska,
Slovenia
Maude Modimothebe Dikobe,
University of Botswana,
Botswana

*CORRESPONDENCE
Ciara Lane
shaciaral@gmail.com

SPECIALTY SECTION
This article was submitted to
Gender, Sex and Sexualities,
a section of the journal
Frontiers in Psychology

RECEIVED 02 June 2022
ACCEPTED 29 August 2022
PUBLISHED 29 September 2022

CITATION
Lane C, Kaya-Capocci S, Kelly R, O'Connell T and Goos M (2022) Fascinating or dull? Female students' attitudes towards STEM subjects and careers.
Front. Psychol. 13:959972.
doi: 10.3389/fpsyg.2022.959972

Fascinating or dull? Female students' attitudes towards STEM subjects and careers

Ciara Lane[1*], Sila Kaya-Capocci[2], Regina Kelly[3], Tracey O'Connell[1] and Merrilyn Goos[1]

[1]EPI*STEM, National Centre for STEM Education, School of Education, University of Limerick, Limerick, Ireland, [2]Faculty of Education, Agri Ibrahim Cecen University, Agri, Turkey, [3]School of Education, University of Limerick, Limerick, Ireland

Internationally, the need to advance science, technology, engineering and mathematics (STEM) education is recognized as being vital for meeting social and economic challenges and developing a scientifically, mathematically, and technologically literate citizenry. In many countries, however, there are gender differences in the participation and achievement of girls and women in STEM education and STEM careers, usually to the disadvantage of females. This paper aims to identify challenges to female students' participation in STEM both at post-primary (secondary school) level and beyond in the Irish context. The research questions we aim to address in this paper are: (1) what are student attitudes towards science, technology, engineering and mathematics as measured through interest and perceived ability in STEM, students' valuing of STEM and students' commitment to STEM? and (2) what gender differences occur regarding students' attitudes to science, technology, engineering and mathematics? A survey was completed by 308 post-primary students in Ireland as part of a one-year research project titled "STEMChAT: Women as catalysts for change in STEM education." Data analysis compiled descriptive statistics, including response frequencies and percentages and median and interquartile range values, and compared gender differences in survey responses using the Kruskal–Wallis H Test. Results indicated that female students had significantly more positive attitudes to science compared to males while in comparison, males had significantly more positive responses to mathematics compared to females. Challenges regarding access to and understanding of STEM in the context of post-primary education are discussed.

KEYWORDS

STEM, attitude, gender, awareness, ability, value, commitment

Introduction

Internationally, there has been an increased emphasis on science, technology, engineering and mathematics (STEM) education due to its significant impact on social, environmental and/or economic development (Kelley and Knowles, 2016; Martin-Paez et al., 2019). Even though STEM education is a highly studied topic, there are still disputes about its meaning. While some researchers explain STEM education with a simple description of the four STEM disciplines (science, technology, engineering and mathematics), others view it as an educational approach at the intersections of any number of the four disciplines. For example, different researchers view STEM education as:

- Science or Maths,
- both Science and Maths,
- Science by incorporating Technology,
- Engineering or Maths,
- a quartet of separate STEM disciplines,
- Science and Maths are connected by a Technology or Engineering program,
- coordination across STEM disciplines,
- combining two or three STEM disciplines,
- complementary overlapping across STEM disciplines,
- transdisciplinary STEM course or program.

Bybee (2013)

Additionally, some studies adopting a more complex understanding view STEM education as merging all four STEM disciplines in an integrated manner (McLoughlin et al., 2020). This study was conducted in Ireland, where the meaning of STEM education is multi-faceted and can include the teaching of the four STEM disciplines in isolation as well as encouragement for cross-disciplinary approaches, especially in primary schools (Department of Education and Skills [DES], 2017). However, subjects taught in Irish post-primary schools are not integrated; rather, students study subjects from the constituent STEM disciplines in a discrete manner.

In many countries it is recognised that providing effective STEM education at primary, secondary and tertiary levels is vital to increase the number and quality of STEM graduates (Marginson et al., 2013; Honey et al., 2014; DES, 2017; The Scottish Government, 2017). In an increasingly global society, it is important for all students to develop 'STEM literacy' to meet social, personal, economic and environmental challenges (Mohr-Schroeder et al., 2020) and, thus, STEM education has come to the fore of national and global policies in recent decades. However, in Ireland, as in many other countries worldwide, gender differences in the participation and achievement of girls and women in STEM education and the STEM workforce are palpable, usually to the disadvantage of females.

Abbreviations: DES, Department of Education and Skills; EU, European Union; OECD, Organisation for Economic Co-operation and Development; SFI, Science Foundation Ireland; STEM, Science Technology Engineering Mathematics; UNESCO, United Nations Educational Scientific and Cultural Organization.

In 2018, Ireland had the highest rate of STEM graduates in the EU at 35.2 per 1,000 persons aged 20–29 (Central Statistics Office, 2021). However, Ireland also had the largest gender differential in the EU, with 47.3 male STEM graduates per 1,000 compared to 23.0 female STEM graduates. The gender gap problem is often portrayed as a "leaky pipeline," with low female participation in second-level STEM subjects leading to similarly low participation rates in third-level STEM programs. In Ireland, there are many informal activities available to students, both male and female, that are designed to boost participation in STEM education and STEM careers. These include the BT Young Scientist Awards (BT Young Scientist and Technology Exhibition, 2021), CoderDojo (CoderDojo Foundation, n.d.), SciFest (2021), and RDS STEM Learning (RDS, 2021). However, the benefits of these informal initiatives are not fully realised unless education systems and schools also provide equal opportunities for boys and girls to access and benefit from quality STEM education.

This paper draws on survey data collected as part of a Science Foundation Ireland (SFI) funded research project titled "STEMChAT: Women as catalysts for change in STEM education" aiming to encourage female post-primary students to pursue STEM subjects and careers. The survey was conducted with post-primary students from 12 schools before STEMChAT activities were introduced (school-based workshops involving conversations with female undergraduate STEM students about university courses and careers). As such, findings reflect participants' pre-existing attitudes to STEM. The conceptual framework for our study is informed by the UNESCO (2017) Ecological Framework, which depicts the multiple and overlapping factors that may influence girls' and women's participation, achievement and progression in STEM studies and careers. These factors are described at four interactive levels: individual; family and peer; school; and society. At the individual level, cognitive traits such as spatial or linguistic skills may be influential, along with psychological factors that include self-efficacy, interest and motivation. Family and peer-related factors highlight the role of parental beliefs, expectations, educational and occupational backgrounds; the household environment and resources; and peer relationships. School-level factors include the learning environment, teacher characteristics, teaching strategies, the curriculum and learning materials, and assessment procedures and tasks. Societal and cultural norms can reinforce or challenge gender stereotypes, and at this societal level of the Ecological Framework the mass media and social media are significant influences on the socialisation of children and young people. Also, at the societal level, formal policies and legislation can also promote gender equality and the advancement of women in the STEM fields.

While it is difficult to investigate the Ecological Framework's multiple levels and inter-related factors in a single study, attention paid to any influential factor or level (e.g., the psychological factors at the individual level) must also take into account the interactions of the other levels and factors (e.g., family, peer, school and societal factors). Although this paper focuses on specific psychological factors at the individual level (students' attitudes), we are cognizant of the various other factors at other levels in which our study is contextualized. The research questions we aim to address in this paper are:

1. What are student attitudes towards science, technology, engineering and mathematics as measured through:

- awareness of STEM (initial interest)
- perceived ability in STEM
- students' valuing of STEM
- students' commitment to STEM (long-term interest)?

2. What gender differences occur regarding students' attitudes to science, technology, engineering and mathematics?

Our paper firstly introduces STEM education in the Irish context and then discusses relevant literature with regard to students' attitudes to STEM, including gender differences. A quantitative analysis is conducted on the survey data collected from post-primary students. The results are discussed in terms of access to STEM subjects and students' attitudes and gender differences, and the limitations of the study are presented. The paper concludes with the contributions to the STEM education field by identifying the challenges to female students' participation in STEM and providing potential research areas to address these challenges.

STEM education in Ireland

The *STEM Education Policy Statement 2017–2026* released by the Department of Education and Skills (2017) in Ireland reveals a vision for providing "the highest quality STEM education experience for learners that nurtures curiosity, inquiry, problem-solving, creativity, ethical behaviour, confidence and persistence, along with the excitement of collaborative innovation" (p. 12). The policy statement underlines the built-in educational benefits of inspiring young people's curiosity about the natural world while also highlighting the importance of producing STEM graduates to drive Ireland's economy. The policy statement discusses the necessity of high-quality STEM education for all students, not only those interested in STEM-related careers, due to its importance in creating STEM-literate citizens who can make well-informed decisions about global issues affecting future generations. According to the *STEM Education Policy Statement,* engagement and participation of learners with STEM is the first of four pillars of the STEM policy plan for advancement. Therefore, it is highly important to investigate and address problems of low participation in STEM disciplines in Irish post-primary schools.

In Ireland, post-primary education comprises Junior Cycle (the first 3 years) and Senior Cycle (the final 2 years), with an optional Transition Year between Junior and Senior Cycle. At the end of Senior Cycle, students sit the Leaving Certificate examination, the results of which determine their entry to third level courses and careers. Students typically study a minimum of six subjects for Senior Cycle, with English, Irish and mathematics taken by the majority of students, due to these being effectively compulsory (being required for entry to most third level courses). Apart from mathematics, other STEM subject choices for Senior Cycle include applied mathematics, biology, chemistry, physics, physics & chemistry, agricultural science, construction studies, design & communication graphics, and engineering.

Students' subject choice at Senior Cycle is naturally affected by subjects they completed at Junior Cycle, subjects offered in the school, availability of teachers, and the students' attitudes towards these subjects, among others. In Ireland, there are dramatic gender imbalances at the post-primary level in favour of male students in physical science and technology subjects and in favour of female students in biology. DES (2017) shows the significant gender differences in the selection of science subjects at Senior Cycle, with the ratios of male students to female students greater than 3:1 for physics and approximately 2:3 for biology. Moreover, the number of female students is considerably lower than male students in STEM courses in higher education.

Access to subjects for female students at post-primary level may be affecting interest and opportunity to study these subjects. Access to subjects in post-primary schools can be difficult at times depending on school budgets and resources in subjects such as engineering or construction studies where tools and space may not be available to equip a practical workshop classroom. In 2019, 325 females sat the Leaving Certificate engineering examination while 4,440 males sat the same exam. Attending single-sex schools, particularly single-sex girls' schools, affects students' access to subjects. Many all-girls post-primary schools would not have workshops for practical subjects like construction studies. Single-sex education is common in Irish schools: in 2017, the *Irish Times* newspaper reported that one-third of schools in Ireland catered for either girls only or boys only (Ahlstrom, 2017). This is a structural feature of the education landscape that can reinforce negative gender stereotypes, such as the perception that STEM subjects and careers are more suitable for boys (report by Accenture, 2014).

Gender disparities in STEM education in Ireland are compounded by subject hierarchies and sub-cultures that exist within the post-primary school curriculum. McGarr and Lynch (2017), in their analysis of the STEM agenda in the Irish education system, highlighted the hierarchical ordering of subjects that often reflects the social and educational capital available to those who choose these subjects. The pursuit and performance of students in technology and engineering subjects versus that in mathematics and science subjects is seen as reflective of student ability in these subjects. STEM subjects have generally been presented as one interrelated entity; however, they play different roles in Irish post-primary schools. The vocationally focused subjects of technology and engineering have traditionally served the needs of lower socio-economic groups, while mathematics and science have been viewed as higher-status subjects that prepare middle-class students for university degrees and more privileged occupations. McGarr and Lynch highlighted the absence of technology and engineering from the broader STEM agenda in post-primary schools, which they claimed is largely due to the under-resourced and out-of-date subjects on offer. For example, the curriculum for engineering has been in place without change or update for over two decades. These researchers argued that the academically oriented subjects of science and mathematics monopolize the STEM agenda, while the traditional role of the vocational subjects in preparing students for post-school employment is now downplayed due to the massification of higher education.

Further research on issues in STEM education internationally is discussed in the next section, particularly in relation to attitudes and gender differences.

Issues in STEM education

Many issues relating to STEM education are raised by international researchers, some of which involve students' attitudes towards STEM education, particularly regarding gender differences at high school (post-primary) level (Brown et al., 2017). To limit our review of the literature, attitudes towards STEM are characterized as STEM interest, STEM values, and STEM perceived ability according to Mahoney's (2010) Students' Attitudes Towards STEM Survey as used in our study, and gender differences are specified as STEM stereotypes. It is acknowledged that these constructs overlap in many research studies, as the following brief review demonstrates.

Students' attitudes towards STEM

In the last decade, several studies investigated particular variables that drive or limit interest in STEM (Valla and Ceci, 2014; Falk et al., 2017; Means et al., 2021). In general, females have been found to have lower interest levels in STEM compared to males (Falk et al., 2017). Social inclusion factors are noted as a particular reason for lower interest. Means et al. (2021) reported on a large-scale meta-analysis of the relationship between attendance at an inclusive STEM high school and a range of academic and motivational outcomes. Rather than selecting students on the basis of prior academic achievement, inclusive STEM high schools provide opportunities for under-represented youth to develop STEM interest and talent. The meta-analysis found that students who attended an inclusive STEM high school were more likely than students in non-STEM comparison schools to report high interest in undertaking a graduate degree and in entering a STEM career, and this effect was also found for low-income and female students. Turner et al. (2019) reported on the importance of efficacy in relation to STEM interest; peer support was noted as a contributing factor to efficacy. These studies call on the need to focus on equity-oriented interventions that increase the social belongingness of students in STEM domains where there is unequal participation by gender in order to increase interest in STEM.

Value beliefs have gained increasing attention in the psychology domain in recent years, with value-related beliefs noted as a strong predictor of career aspirations in STEM (Wang et al., 2013; Wegemer and Eccles, 2019). Van Tuijl and Van der Molen (2016) conducted a review of literature regarding the study choice factors that are most influential on children from the age of 8–16. They conclude that the undesirable affective value associated with many STEM fields is detrimental to the career aspirations of young people, particularly those who do not align with the stereotypical image of STEM careers. Tzu-Ling (2019) used multiple regression to investigate the difference between males' and females' career aspirations using the variables of task value, self-efficacy and family support. Tzu-Ling reported that task value is a variable that can significantly predict STEM career aspirations, regardless of gender, whereas self-efficacy could significantly predict STEM career aspiration for male students only. Eccles and Wigfield (2020) argued that more research is needed to explore how complex interactions between culture, gender, and ethnicity influence the development of individuals' subjective task values. These studies highlight the need to develop the cultural and affective value of STEM and STEM tasks as a means to counter the negative values associated with some STEM fields.

Historically, researchers report that females' lack of interest in STEM was attributed to their lack of ability (Jungert et al., 2019; Sobieraj and Krämer, 2019). More recently, research in this area concerns the difference between males' and females' perceived abilities (Hand et al., 2017; Sobieraj and Krämer, 2019). Brown et al. (2017) reported instances where, although post-primary school students achieve the same grades in STEM subjects, females' perceived ability was lower than that of males. Similarly, Sobieraj and Krämer investigated differences between male and female self-efficacy and perceived competence, finding that female students had lower self-perceptions of their abilities in STEM compared to male students. Hand et al. (2017) argued that the subtle biases of high school teachers have a detrimental effect on female self-efficacy in mathematics and science: such biases were evident when teachers expected girls and boys to perform differently in STEM subjects based on their perceptions of masculine and feminine traits. Ertl et al. (2017) reported that stereotypes have a damaging impact on female students' self-concept, even when they have academic success. Ertl et al. suggested the reason may derive from the stereotypical belief that female achievements are due to diligence rather than ability. Kessels (2015) noted that when the stereotype is associated with a particular STEM career that does not align with a student's self-concept, this constrains their career choice. Van Aalderen-Smeets and van der Molen (2018) hypothesised that changing students' implicit theories of intelligence might improve their STEM-related self-efficacy beliefs and possibly the likelihood of choosing STEM-related study or careers. This is particularly important for female students, who are thought to hold entity beliefs more so than males; in other words, girls are more likely to believe that intelligence is fixed rather than malleable.

Overall, the literature identifies that females have lower levels of interest in STEM and lower perceptions of their abilities in STEM than males. This difference may be caused by social factors which influence students' interest in STEM.

Gender differences in STEM stereotypes

Blažev et al.'s (2017) study with school pupils in Croatia shows that male students and those who had previous success in STEM subjects are more likely to hold stereotypical beliefs about STEM. Several factors have been proposed to positively impact stereotypical beliefs, such as the presence of females in a class (Gunderson et al., 2012; Master et al., 2014; Riegle-Crumb et al., 2017). Riegle-Crumb et al. (2017) conducted a study regarding the presence of females in high school classes. They reported that female peers had a positive impact in reducing male peers' stereotypical beliefs. The presence of female teachers seemed to have a similar impact: Master et al. (2014) found that female teachers reduced female students' concerns about being negatively stereotyped in classroom situations. In contrast, Gunderson et al. (2012) reported that gender-biased stereotypes about females' mathematics capabilities are cultivated, rather than ameliorated, by teachers. Exposure to role models is often promoted as a way of overcoming negative stereotypical beliefs about STEM. Gladstone and Cimpian's (2021) systematic review of the literature in this field yielded four recommendations for maximising the effectiveness of role models in STEM for motivating students from diverse gender and ethnic backgrounds: (1) portray role models as being competent and successful, while avoiding extreme levels of success that might instead be alienating; (2) portray role models as being meaningfully similar to students; (3) prioritise exposure to role models from groups that are traditionally underrepresented in STEM; (4) portray role models' success as being attainable. Luo et al. (2021) investigated upper primary students' stereotypical beliefs about STEM careers and found that these beliefs negatively predicted STEM self-efficacy and career-related outcome expectations. Their findings suggest that interventions targeting STEM career aspirations need to target STEM stereotypes, self-efficacy, and outcome expectations simultaneously.

Methodology

The data presented in this paper were collected as part of the STEMChAT project between 2019 and 2020 in Ireland. In line with the UNESCO (2017) Ecological Framework, while this paper focuses on specific psychological factors at the individual level as described, we are cognizant of the various other factors at other levels in which our study is contextualized. For example, at the family and peer level, our participants were sampled from schools located in socially disadvantaged areas, meaning that many of the students were likely from families with a lower socio-economic status which has been linked to lower academic achievement and expectations as well as possible adherence to more conventional gender role beliefs (Tenenbaum and Leaper, 2003; Organization for Economic Co-operation and Development [OECD], 2016). At the school level, participants attended 12 different post-primary schools with consequential exposure to different learning environments, including teacher quality and instructional practices, resources, assessments and school environments (UNESCO, 2017). At the societal level, deeply embedded societal and cultural norms regarding 'traditional' gendered subject choice at school and perceived gender 'appropriate' careers permeate more recent gender equality and inclusive policies in relation to STEM education and the STEM workforce (DES, 2017). As the project aimed to encourage female post-primary students to pursue STEM subjects and careers, the participating schools were selected according to their accessibility, social disadvantage of the area, and enrolment of female students.

Participants

The participants of the study were post-primary students (mainly Transition Year students) in Ireland. Transition Year is a one-year program that students can volunteer to complete between Junior Cycle and Senior Cycle. In some Irish schools, Transition Year is mandatory. Each school designs its own Transition Year programme; therefore, programme content can vary between schools. A total of 308 students completed the survey including 218 females (71%) and 89 males (29%). One student did not disclose gender. Participants were aged between 14 and 18 years with a mean age of 16 years. Participants were sampled from schools participating in the STEMChAT project which led to a sampling bias in favour of females; this is discussed later as a potential limitation of the study.

Study context and design

The surveys were completed by students from 12 post-primary schools ($n_{\text{female school}} = 4$, $n_{\text{male school}} = 2$, $n_{\text{mixed school}} = 6$) with 4 schools offering subjects from each of the STEM disciplines (science, technology, engineering and mathematics), 2 schools offering science, technology and mathematics subjects, and 6 schools offering only science and mathematics subjects. It is worth noting here that none of the single-sex girls' schools were offering technology or engineering subjects.

In the Junior Cycle (lower secondary) curriculum, mathematics and science are taught as stand-alone subjects, and there is a suite of "technology" subjects comprising applied technology, engineering, wood technology, and graphics. There is a wider range of STEM subjects offered in the Senior Cycle curriculum, reflecting increased specialisation at this level of schooling. In their analysis of the treatment of STEM subjects within the Irish post-primary school context, McGarr and Lynch (2017) observed that mathematics and science occupy a higher status in the hierarchy of school subjects, in part because they have clearly defined subject boundaries and draw on long-established bodies of academic knowledge. In contrast, the traditionally vocational engineering and technology subjects hold a lower social status, draw on subject knowledge from a range of disciplines and are thus more loosely framed. This classification proved to be significant for our study, since the naming of the engineering and technology subjects made it difficult for students to identify them as belonging to the STEM categories of "engineering" and especially "technology."

Survey instrument

Our study is concerned with individual level factors referred to in the UNESCO (2017) Ecological Framework, specifically interest, perceived ability, value and commitment to STEM. These individual psychological factors are captured for both male and female students in our study, enabling gender comparison. We used a survey to identify students' perceptions about STEM careers and students' attitudes towards STEM. The items of the survey were drawn from two pre-existing validated surveys; Students' Attitudes Towards STEM Survey (Mahoney, 2010) and STEM Semantics Survey (Tyler-Wood et al., 2010).

The Students' Attitudes Towards STEM Survey (Mahoney, 2010) involved 24 items aiming to investigate high-school students' awareness of (initial interest in) STEM, perceived ability in STEM, perceptions of the value of STEM, and commitment to (long-term interest in) STEM. For each item, students were asked to choose either science, technology, engineering or mathematics and indicate their response. For example, the first item, in the category of awareness of STEM, was "I do not like […]"; students responded by identifying one of the STEM disciplines that they did not like. Thus, for each item, students chose one STEM discipline that matched the attitude portrayed by that item: they were not required to indicate their attitudes towards every STEM discipline on every survey item.

The STEM semantics survey (Tyler-Wood et al., 2010) included semantic differential scales comprising five adjective pairs that reflect perceptions of science, technology, engineering and mathematics, respectively. A fifth scale, using the same adjective pairs, elicited perceptions of STEM career interests. Thus, the survey consisted of 25 items (five adjective pairs x five target areas). Students selected a response on a 1–7 scale to indicate how they felt about each content area. The wording of items was reviewed in terms of age-appropriateness and cultural differences. Only one word "mundane" was changed to "dull" to ensure students would understand the intended meaning.

Data collection

Data collection took place in the Spring semester of 2018–2019 academic year (January–May) and in the Autumn semester of 2019–2020 academic year (September–December). Consent was obtained from the school principals, teachers who were providing their class time, students who volunteered to participate, and the guardians of participating students. The survey was completed at the beginning of a session for the STEMChAT project and students were asked to complete the survey to identify their pre-existing attitudes towards STEM.

Data analysis

The survey data were analyzed through the Statistical Package for the Social Sciences (SPSS version 26). Data analysis comprised descriptive statistics, including response frequencies and percentages for both attitudinal items from Mahoney (2010) and STEM Semantics Survey items from Tyler-Wood et al. (2010), and median and interquartile range values for the STEM Semantic Survey items. The Kruskal–Wallis H test for differences between genders was also employed for the STEM Semantic Survey items but not for attitudinal items as the test can only be applied to continuous or ordinal variables. Because the attitudinal items used a nominal scale, it was neither possible nor meaningful to measure reliability *via* internal consistency. Reliability of the semantic differential scale was measured by calculating Cronbach's alpha for each of the five component sub-scales (replicating the reliability analysis conducted by Tyler-Wood et al., 2010). Each sub-scale had a high level of internal consistency as determined by the Cronbach's alpha results shown in Table 1. The internal consistency of the scale could not be improved by removing any of the items.

Findings

Context of the study

Participants were asked whether their school offered each of the STEM subjects and which STEM subjects they studied. Responses indicated that all schools offered science and mathematics subjects and only 4 schools offered engineering, but there was some confusion amongst participants as to whether technology was offered as a subject in their school (subjects actually offered by each school were confirmed by the authors and are provided in the Methodology section). This may be indicative of students' confusion about what technology means: for example, they may have interpreted technology as meaning ICT or computer science as opposed to construction studies or design and communication graphics, even though ICT/computer science was not a subject in the school curriculum at the time of this study. Frequency of responses for males and females studying each of the STEM disciplines are shown in Table 2. Percentages by gender were calculated based on total females or males. For example, 32 females correspond to 14.68% of female participants.

Apart from two students, all participants studied mathematics and the majority studied science, with a slightly greater percentage

TABLE 1 Internal consistency reliabilities for STEM semantics scales.

Scale	Number of items	Alpha
Science	5	0.841
Mathematics	5	0.833
Engineering	5	0.886
Technology	5	0.887
STEM career	5	0.810

of females compared to males studying science. Males were almost twice as likely as females to be studying technology and engineering, but the number of participants studying these subjects was considerably lower than mathematics and science. It should be recalled that mathematics is effectively compulsory in Irish post-primary schools and science is compulsory in the majority of Irish post-primary schools at Junior Cycle. When considering students' attitudes and perceptions as reported in the following section, we need to bear in mind the relatively low numbers of students who have experienced technology and engineering which will feasibly have impacted their responses to the survey items. In particular, when participants were asked to choose one of science, technology, engineering or mathematics when responding to Mahoney's (2010) Attitudes toward STEM items, some participants may have been chosen only from those subjects they have personally experienced in school or may have indifferent attitudes towards subjects (such as technology/engineering subjects) in which they have had little or no experience. Nevertheless, the different response patterns, indicating that students may be less familiar with or interested in technology and engineering subjects, constitute a significant finding from the study and raise questions about the relative status and visibility of the constituent STEM disciplines in Irish post-primary schools.

TABLE 2 STEM subjects studied by gender.

	Science, *N* (%)	Technology, *N* (%)	Engineering, *N* (%)	Mathematics, *N* (%)
Females	206 (94.5%)	32 (14.7%)	25 (11.5%)	218 (100.0%)
Males	79 (88.8%)	26 (29.2%)	18 (20.2%)	87 (97.8%)
Total	286 (92.9%)	58 (18.8%)	43 (14.0%)	306 (99.4%)

Students' attitudes towards STEM

Participants' attitudes to STEM were measured using Mahoney's (2010) instrument, where participants select either science, technology, engineering or mathematics for each of the items. Frequency of responses for each item across the four options was obtained for all participants and for females and males. Frequencies for the items intended to capture participants' initial interest in STEM are shown in Table 3. Negatively worded items are labelled (N) and positively worded items are labelled (P). The response option with the highest frequency is shaded. For each item there were some participants who did not respond or

TABLE 3 Initial interest in STEM.

Item	Science, *N* (%)	Technology, *N* (%)	Engineering, *N* (%)	Mathematics, *N* (%)	Not given/multiple responses, *N* (%)
I do not like […] (N)					
Female	31 (14.2%)	34 (15.6%)	57 (26.1%)	89 (40.8%)	7 (3.2%)
Male	26 (29.2%)	8 (9.0%)	13 (14.6%)	33 (37.1%)	9 (2.0%)
All	58 (18.8%)	42 (13.6%)	70 (22.7%)	122 (39.6%)	16 (5.2%)
I enjoy learning about […] (P)					1 ((0.5%)
Female	147 (67.4%)	27 (12.4%)	12 (5.5%)	31 (14.2%)	
Male	32 (36.0%)	20 (22.5%)	18 (20.2%)	16 (18.0%)	3 (3.4%)
All	179 (58.1%)	47 (15.3%)	30 (9.7%)	47 (15.3%)	5 (1.6%)
I am curious about […] (P)					
Female	80 (36.7%)	66 (30.3%)	59 (27.1%)	11 (5.0%)	2 (0.9%)
Male	23 (25.8%)	31 (34.8%)	24 (27.0%)	7 (7.9%)	4 (4.5%)
All	103 (33.4%)	97 (31.5%)	83 (26.9%)	18 (5.8%)	7 (2.2%)
I am not interested in […] (N)					
Female	19 (8.7%)	62 (28.4%)	78 (35.8%)	47 (21.6%)	12 (5.5%)
Male	29 (32.6%)	13 (14.6%)	15 (16.9%)	27 (30.3%)	5 (5.6%)
All	48 (15.6%)	75 (24.4%)	93 (30.2%)	74 (24.0%)	18 (5.9%)
I like […] (P)					
Female	89 (40.8%)	36 (16.5%)	14 (6.4%)	71 (32.6%)	8 (3.7%)
Male	23 (25.8%)	18 (20.2%)	22 (24.7%)	24 (27.0%)	2 (2.2%)
All	112 (36.4%)	54 (17.5%)	36 (11.7%)	95 (30.8%)	11 (3.6%)
[…] is appealing to me (P)					
Female	96 (44.0%)	47 (21.6%)	34 (15.6%)	35 (16.1%)	6 (2.8%)
Male	25 (28.1%)	14 (15.7%)	34 (38.2%)	12 (13.5%)	4 (4.5%)
All	121 (39.3%)	61 (19.8%)	68 (22.1%)	47 (15.3%)	11 (3.5%)

TABLE 4 Perceived ability in STEM.

Item	Science, *N* (%)	Technology, *N* (%)	Engineering, *N* (%)	Mathematics, *N* (%)	Not given/multiple responses, *N* (%)
[…] is difficult for me (N)					
Female	35 (16.1%)	22 (10.1%)	26 (11.9%)	129 (59.2%)	6 (2.7%)
Male	26 (29.2%)	6 (6.7%)	5 (5.6%)	48 (53.9%)	4 (4.5%)
All	61 (19.8%)	28 (9.1%)	31 (10.1%)	177 (57.5%)	11 (3.5%)
I do well in […] (P)					
Female	115 (52.8%)	10 (4.6%)	8 (3.7%)	75 (34.4%)	10 (4.6%)
Male	25 (28.1%)	10 (11.2%)	13 (14.6%)	36 (40.4%)	5 (5.6%)
All	140 (45.5%)	20 (6.5%)	21 (6.8%)	111 (36.0%)	16 (5.1%)
I am not confident about my work in […] (N)					
Female	50 (22.9%)	30 (13.8%)	27 (12.4%)	103 (47.2%)	8 (3.7%)
Male	29 (32.6%)	12 (13.5%)	6 (6.7%)	34 (38.2%)	8 (9.0%)
All	79 (25.6%)	42 (13.6%)	33 (10.7%)	137 (44.5%)	17 (5.5%)
I have a hard time in […] (N)					
Female	43 (19.7%)	28 (12.8%)	29 (13.3%)	109 (50.0%)	9 (4.1%)
Male	30 (33.7%)	3 (3.4%)	5 (5.6%)	44 (49.4%)	7 (7.9%)
All	74 (24.0%)	31 (10.1%)	34 (11.0%)	153 (49.7%)	16 (5.2%)
Assigned work in […] is easy for me (P)					
Female	118 (54.1%)	19 (8.7%)	10 (4.6%)	63 (28.9%)	8 (3.7%)
Male	26 (29.2%)	13 (14.6%)	12 (13.5%)	36 (40.4%)	2 (2.2%)
All	144 (46.8%)	32 (10.4%)	22 (7.1%)	100 (32.5%)	10 (3.2%)
I cannot figure out […] (N)					
Female	33 (15.1%)	46 (21.1%)	48 (22.0%)	84 (38.5%)	7 (3.2%)
Male	23 (25.8%)	15 (16.9%)	11 (12.4%)	29 (32.6%)	11 (12.4%)
All	57 (18.5%)	61 (19.8%)	59 (19.2%)	113 (36.7%)	18 (5.8%)

who gave multiple responses and these participants are shown in the last column of the table.

Participants' responses regarding initial interest in STEM show the most interest in science. In particular, majority of students stated that they enjoyed learning about science, although it is clear from the gender breakdown that more females are interested in science compared to males. Mathematics was the subject least liked by both male and female participants with almost 40% of participants stating that they do not like mathematics. There might be different reasons for this finding, such as students' mathematics self-efficacy, stereotypical beliefs about mathematics being for intelligent people, and students perceiving mathematics as numbers rather than understanding its role in real life. Bearing in mind that only 18.8% of participants were studying technology subjects and 14.0% engineering, it was noteworthy to see the high proportion (31.5%) who identified technology as the STEM discipline they were curious about and engineering as that having the least interest for them (30.2%).

Frequencies for the items measuring participants' perceived ability in STEM are shown in Table 4. Negatively worded items are labelled (N) and positively worded items are labelled (P). The response option with the highest frequency is shaded for each item.

Supporting the finding about students' attitude, the results showed that students' perceived ability in mathematics was low. A clear majority of participants stated that mathematics was difficult for them, that they were not confident about their work in mathematics, that they have a hard time in and cannot figure out mathematics. Similar to the initial interest items, science was the most frequent response option for positively worded items for females, although mathematics was the subject in which male participants most frequently reported that they do well. Unsurprisingly, science and mathematics were the two most frequently chosen response options for most of the items in relation to participants' perceived ability, likely due to the fact that most students do not study technology or engineering. The data also showed that while the female participants' perceived ability in technology and engineering was very low, a much higher percentage of male participants said they did well in these subjects and found the work easy.

Frequencies for the items measuring participants' perceptions of the value of STEM are shown in Table 5. Negatively worded items are labelled (N) and positively worded items are labelled (P). The response option with the highest frequency is shaded for each item.

Responses to the attitudinal items regarding value of STEM demonstrate high value of both science and mathematics as

TABLE 5 Value of STEM.

Item	Science, *N* (%)	Technology, *N* (%)	Engineering, *N* (%)	Mathematics, *N* (%)	Not given/multiple responses, *N* (%)
[…] is important to me (P)					
Female	95 (43.6%)	24 (11.0%)	9 (4.1%)	75 (34.4%)	15 (6.9%)
Male	26 (29.2%)	13 (14.6%)	15 (16.9%)	31 (34.8%)	4 (4.5%)
All	121 (39.3%)	37 (12.0%)	24 (7.8%)	107 (34.7%)	19 (6.1%)
I feel there is a need for […] (P)					
Female	79 (36.2%)	38 (17.4%)	20 (9.2%)	73 (33.5%)	8 (3.7%)
Male	25 (28.1%)	12 (13.5%)	17 (19.1%)	31 (34.8%)	4 (4.5%)
All	104 (33.8%)	50 (16.2%)	37 (12.0%)	105 (34.1%)	12 (3.9%)
I do not need […] (N)					
Female	24 (11.0%)	67 (30.7%)	89 (40.8%)	21 (9.6%)	17 (7.8%)
Male	25 (28.1%)	20 (22.5%)	29 (32.6%)	6 (6.7%)	9 (10.1%)
All	49 (15.9%)	87 (28.2%)	118 (38.3%)	27 (7.8%)	27 (8.8%)
It is valuable for me to learn […] (P)					
Female	93 (42.7%)	13 (6.0%)	8 (3.7%)	96 (44.0%)	8 (3.6%)
Male	20 (22.5%)	12 (13.5%)	8 (9.0%)	46 (51.7%)	3 (3.3%)
All	114 (37.0%)	25 (8.1%)	16 (5.2%)	142 (46.1%)	11 (3.5%)
[…] is good for me (P)					
Female	95 (43.6%)	17 (7.8%)	7 (3.2%)	88 (40.4%)	11 (5.1%)
Male	37 (41.6%)	9 (10.1%)	12 (13.5%)	28 (31.5%)	3 (3.4%)
All	133 (43.2%)	28 (8.4%)	19 (6.2%)	116 (37.7%)	14 (4.5%)
I do not care about […] (N)					
Female	19 (8.7%)	67 (30.7%)	79 (36.2%)	41 (18.8%)	12 (5.5%)
Male	26 (29.2%)	18 (20.2%)	27 (30.3%)	9 (10.1%)	9 (10.1%)
All	45 (14.6%)	85 (27.6%)	106 (34.4%)	50 (16.2%)	22 (7.2%)

important and valuable to learn, although again male participants responded more positively towards mathematics than female participants. The responses indicate that although students may have lower interest in mathematics and have poor perceptions of their mathematical ability, they recognize the value of learning mathematics. There was a strong perception amongst participants that engineering, and to a lesser extent technology, were subjects they did not need and did not care about. As many participants had not experienced either of these subject areas, their responses highlight the difficulty of promoting positive attitudes to the full range of STEM subjects when there is not universal and equitable access to these subjects in all schools.

The final group of attitudinal items from Mahoney (2010) measuring commitment or long-term interest in STEM are shown in Table 6. Negatively worded items are labelled (N) and positively worded items are labelled (P). The response option with the highest frequency is shaded for each item.

Responses to the items regarding long-term interest in STEM show similarity with initial interest as females' responses were overwhelmingly positive towards science with more than half indicating that they would continue to enjoy science and are interested in alternative programs. However, there were distinctive gender differences in students' interest in science. Males were least interested in a career involving science, were more committed to learning mathematics, and curious to learn more about technology. Technology was the second most frequently selected option in relation to interest in alternative programs for both males and females. Engineering was overwhelmingly the field in which participants, especially females, were not interested in pursuing a career. As for the previous set of items on the value of STEM, lack of familiarity with and exposure to engineering subjects at school may have influenced the responses.

Perceptions of STEM disciplines and careers

The second part of our survey included Tyler-Wood et al.'s (2010) STEM Semantics Scale where participants select a rating between 1 and 7 for five adjective pairs. Three of the adjective pairs are ordered positive–negative and two adjective pairs are ordered negative–positive. Therefore, for the positive–negative pairs, a lower score indicates a positive response while a higher score indicates a negative response. For the negative–positive pairs, a lower score indicates a negative response while a higher score indicates a positive response. Median scores and Interquartile range values for each adjective pair was calculated for each of the STEM disciplines as well as STEM careers as shown in Table 7.

TABLE 6 Long-term interest in STEM.

Item	Science, *N* (%)	Technology, *N* (%)	Engineering, *N* (%)	Mathematics, *N* (%)	Not given/multiple responses, *N* (%)
I will continue to enjoy […] (P)					
Female	128 (58.7%)	27 (12.4%)	12 (5.5%)	45 (20.6%)	6 (2.8%)
Male	24 (27.0%)	18 (20.2%)	23 (25.8%)	19 (21.3%)	5 (5.6%)
All	152 (49.4%)	45 (14.6%)	35 (11.4%)	64 (20.8%)	12 (3.8%)
I am not interested in a career involving […] (N)					
Female	25 (11.5%)	37 (17.0%)	91 (41.7%)	47 (21.6%)	18 (8.2%)
Male	36 (40.4%)	13 (14.6%)	22 (24.7%)	11 (12.4%)	7 (7.8%)
All	61 (19.8%)	50 (16.2%)	113 (36.7%)	58 (18.8%)	26 (8.4%)
I am interested in alternative programs in […] (P)					
Female	110 (50.5%)	37 (17.0%)	23 (10.6%)	33 (15.1%)	15 (6.9%)
Male	32 (36.0%)	20 (22.5%)	18 (20.2%)	13 (14.6%)	6 (6.7%)
All	143 (46.4%)	57 (18.5%)	41 (13.3%)	46 (14.9%)	21 (6.8%)
I would like to learn more about […] (P)					
Female	94 (43.1%)	60 (27.5%)	42 (19.3%)	13 (6.0%)	9 (4.1%)
Male	25 (28.1%)	27 (30.3%)	23 (25.8%)	12 (13.5%)	2 (2.2%)
All	119 (38.6%)	87 (28.2%)	65 (21.1%)	26 (8.4%)	11 (3.5%)
I do not wish to continue my education in […] (N)					
Female	36 (16.5%)	49 (22.5%)	47 (21.6%)	59 (27.1%)	27 (12.4%)
Male	31 (34.8%)	14 (15.7%)	15 (16.9%)	16 (18.0%)	13 (14.6%)
All	67 (21.8%)	63 (20.5%)	62 (20.1%)	75 (24.4%)	41 (13.3%)
I am committed to learning […] (P)					
Female	100 (45.9%)	9 (4.1%)	9 (4.1%)	83 (38.1%)	17 (7.8%)
Male	22 (24.7%)	13 (14.6%)	15 (16.9%)	34 (38.2%)	5 (5.6%)
All	122 (39.6%)	22 (7.1%)	24 (7.8%)	118 (38.3%)	22 (7.1%)

TABLE 7 Median and interquartile range for STEM semantics survey.

Adjective pair	Science (IQR Q1:Q3)	Mathematics (IQR Q1:Q3)	Engineering (IQR Q1:Q3)	Technology (IQR Q1:Q3)	STEM career (IQR Q1:Q3)
Fascinating – Dull	2.0 (1.0:4.0)	4.0 (3.0:6.0)	4.0 (2.0:6.0)	4.0 (2.0:5.0)	3.0 (2.0:4.0)
Appealing – Unappealing	3.0 (2.0:4.0)	4.0 (3.0:6.0)	4.0 (2.0:6.0)	3.0 (2.0:5.0)	3.0 (1.0:4.0)
Exciting – Unexciting	3.0 (2.0:4.0)	4.0 (3.0:6.0)	4.0 (3.0:6.0)	4.0 (3.0:5.0)	3.0 (2.0:4.0)
Means Nothing – Means a Lot	6.0 (4.0:7.0)	5.0 (4.0:6.75)	4.0 (3.0:5.0)	5.0 (3.0:6.0)	5.0 (4.0:7.0)
Boring – Interesting	5.0 (4.0:7.0)	4.0 (2.0:6.0)	4.0 (3.0:6.0)	4.0 (3.0:6.0)	5.0 (4.0:7.0)

With regards to perceptions of science, the median scores for the five items indicate that participants generally responded positively with science being perceived as fascinating, somewhat appealing, somewhat exciting, meaning a lot and somewhat interesting. Looking at students' attitude, perceived ability, value and long-term interest, this result may not be surprising. Median scores for the five items in relation to mathematics, engineering and technology were generally more neutral compared to perceptions of science so the interquartile range is used to determine whether responses tend towards the positive or negative perception. Participants' perceptions rated mathematics as neutral but tending towards dull, unappealing, unexciting, somewhat meaningful, and interesting. Students perceived engineering neutral for three adjective pairs but tending towards unexciting and interesting. Participants perceived technology as neutral but tending towards dull, somewhat appealing, somewhat meaningful and interesting. Median scores for the five items regarding perceptions of STEM careers were slightly positive with a STEM career perceived as somewhat fascinating, somewhat appealing, somewhat exciting, somewhat meaningful and somewhat interesting.

To compare differences between genders for the STEM Semantics Scale, we used the Kruskal–Wallis H-Test since there was a considerable difference between the number of participants in the comparative groups ($n_{\text{female}} = 218$; $n_{\text{male}} = 89$). As the distributions of females' and males' scores were not identical, we cannot compare medians and instead report on mean ranks. Results for differences between genders are shown in Appendix 1.

With regards to participants' perceptions of science, there was a statistically significant difference ($p < 0.05$) between genders for all five items as follows:

- In rating science as fascinating – dull, $\chi^2(1) = 6.002$, p = 0.014 with a mean rank score of 143.02 for females and 169.42 for males.
- In rating science as appealing – unappealing, $\chi^2(1) = 14.044$, p = 0.000 with a mean rank score of 137.11 for females and 177.23 for males.
- In rating science as exciting – unexciting, $\chi^2(1) = 5.901$, p = 0.015 with a mean rank score of 141.41 for females and 167.63 for males.
- In rating science as means nothing – means a lot, $\chi^2(1) = 5.729$, p = 0.017 with a mean rank score of 155.93 for females and 130.35 for males.
- In rating science as boring – interesting, $\chi^2(1) = 6.243$, p = 0.012 with a mean rank score of 155.73 for females and 128.91 for males.

These results indicate that females in this study had significantly more positive perceptions of science compared to males for all items, with the greatest difference between the genders found in their rating of science in terms of appealing – unappealing.

Comparing females' and males' perceptions of mathematics, there was a statistically significant difference ($p < 0.05$) between genders in rating mathematics as boring – interesting, $\chi^2(1) = 4.638$, p = 0.031 with a mean rank score of 143.25 for females and 166.72 for males. There was no statistically significant difference ($p < 0.05$) between genders in rating other mathematics items. These results indicate that males have a significantly greater interest in mathematics compared to females in our study, but while male participants had more positive responses to mathematics compared to females for other items the differences between the genders' responses were not statistically significant.

In relation to participants' perceptions of engineering, a statistically significant difference ($p < 0.05$) between genders was found for four of the scale items as follows:

- In rating engineering as fascinating – dull, $\chi^2(1) = 12.133$, p = 0.000 with a mean rank score of 158.97 for females and 121.34 for males.
- In rating engineering as appealing – unappealing, $\chi^2(1) = 15.977$, p = 0.000 with a mean rank score of 160.84 for females and 117.35 for males.
- In rating engineering as exciting – unexciting, $\chi^2(1) = 11.003$, p = 0.001 with a mean rank score of 157.45 for females and 121.85 for males.
- In rating engineering as boring – interesting, $\chi^2(1) = 5.255$, p = 0.022 with a mean rank score of 139.90 for females and 164.66 for males.

These results indicate that males in our study had significantly more positive perceptions of engineering compared to females for four items, with the greatest difference between the genders was found in their rating of engineering in terms of appealing – unappealing.

There was a statistically significant difference ($p < 0.05$) between genders found for three of the scale items in relation to technology as follows:

- In rating technology as fascinating – dull, $\chi^2(1) = 6.485$, p = 0.011 with a mean rank score of 157.05 for females and 129.56 for males.
- In rating technology as appealing – unappealing, $\chi^2(1) = 7.718$, p = 0.005 with a mean rank score of 157.08 for females and 126.83 for males.
- In rating technology as exciting – unexciting, $\chi^2(1) = 9.670$, p = 0.002 with a mean rank score of 157.68 for females and 124.08 for males.

These results indicate that male participants in our study had significantly more positive responses to technology compared to females for these three items, with the greatest difference between the genders found in their rating of technology in terms of being exciting – unexciting. No statistically significant difference emerged between the genders' responses to the items means nothing – means a lot or boring – interesting.

There was no statistical difference ($p < 0.05$) between genders in rating any of the career items.

Discussion

In this section, we discuss the findings to address our research questions: 1. What are student attitudes towards science, technology, engineering and mathematics as measured through: awareness (initial interest); perceived ability; value; and commitment (long-term interest)? 2. What gender differences occur regarding students' attitudes to science, technology, engineering and mathematics? In addressing these questions, our discussion is centred around three key issues and/or challenges: access to STEM, attitudes to STEM and gender differences. We also discuss limitations of our study.

Access to STEM subjects

Our findings highlight the differential access to STEM subjects for post-primary students in Ireland, particularly in relation to technology and engineering subjects. Students attending single-sex schools in this study did not have access to engineering subjects at their school and many students were confused as to whether they had access to technology subjects due to the lack of clarity in labelling of these subjects in the school curriculum. Post-primary schools in Ireland fall into three categories: voluntary secondary schools, vocational schools, and community and comprehensive schools (Eurydice, 2019). Single-sex schools are secondary schools while co-educational schools include all three school types. Traditionally, secondary schools chiefly offered perceived 'academic' subjects while vocational schools offered more 'practical' subject choices and this traditional segregation continues to some extent today. As a result, students attending secondary schools, and consequentially students attending single-sex schools, have limited access to subjects from the technology and engineering suite. Interpreting this situation in terms of the UNESCO (2017) Ecological Framework, we see how school and societal factors impact individual factors in terms of females' participation, achievement and progression in STEM studies and careers. Many students do not have a choice in relation to the post-primary school they attend, meaning that access to all STEM subjects is essentially a 'postcode lottery'. Students, or indeed their parents, who have a choice and prefer single-sex educational settings may or may not be aware of the limiting consequences of their choice. If the gender gap problem is considered as a "leaky pipeline," with low female participation in second-level STEM subjects leading to similarly low participation rates in third level STEM programs, lack of access to STEM subjects must be seen as one of the sources of the leaks and a key challenge to students, and for the purpose of our interest particularly, females' participation in STEM education and careers. "Education systems and schools play a central role in determining girls' interest in STEM subjects and in providing equal opportunities to access and benefit from quality STEM education" (UNESCO, 2017, p. 11), but clearly there is significant scope for improving equal opportunities in the current Irish post-primary system.

Furthermore, there is a need for greater clarity about what STEM means in relation to post-primary education. Our study found considerable confusion among students about what STEM means in relation to school subjects, particularly technology. This could be observed in the fact some students were unsure whether technology was offered at their school. At Junior Cycle in Ireland (which all participants in our study had completed), there are two technology subjects – construction studies and design and communication graphics. However, even at a state curricular level, there appears to be a lack of clarity about what technology constitutes. The National Council for Curriculum and Assessment (NCCA) includes engineering in the technology suite of subjects, which highlights a systemic lack of clarity concerning the technology and engineering aspects of STEM in the Irish school system. McGarr and Lynch (2017) discussed the absence of technology and engineering from the broader STEM agenda in post-primary schools, pointing not only to the under-resourced and out of date subjects on offer but also to the social history of these vocationally-oriented subjects and their resultant lower status in the school curriculum. Our findings indicate that this 'absence' manifests in lack of understanding and awareness of these subjects among students. This leads us to suggest that, in Ireland, the STEM education agenda might be better represented as S(t)eM because of the limited offering of engineering subjects and the confusion that exists about the meaning of "technology." Adopting engineering design as a catalyst to STEM education and developing a more profound and clear understanding of technology is therefore paramount to STEM education (Kelley and Knowles, 2016). With different interpretations of STEM amongst researchers (Bybee, 2013) and between contexts (research, policy, school, society, media etc.), we argue for the need for greater transparency and consistency about what STEM means, especially in relation to school subjects.

Student attitudes to STEM

Students in our study generally held positive attitudes towards science (enjoyment, curiosity, liking, importance, value, doing well, commitment to learning) and less positive attitudes towards mathematics (not liking, difficult, not confident, do not understand) while still endorsing the value of mathematics and a commitment to learning it. Students in our study also expressed a lack of interest in engineering, and no strong views on technology, possibly because they did not understand what this discipline meant in a school context which presents a challenge to future participation. Engineering was perceived as having the least value among the STEM subjects, which again could be linked to the lack of access in schools compared to science and mathematics and the sub-cultures surrounding STEM subjects in post-primary education in Ireland (McGarr and Lynch, 2017). This finding highlights again the interrelation of school and societal factors with the individual's STEM ecosystem (UNESCO, 2017). Ireland's *STEM Education Policy Statement 2017–2026* advocates for "the highest quality STEM education experience for learners that nurtures curiosity, inquiry, problem-solving, creativity, ethical behaviour, confidence and persistence, along with the excitement of collaborative innovation" (DES, 2017, p. 12). The report highlights the importance of inspiring young people's curiosity and the related educational benefits. Our study indicates that students are distinctly lacking this curiosity about mathematics, more so than any other STEM subject. Our study also found demonstrably lower confidence in mathematics compared to other STEM subjects, although this is caveated by the low participation in technology and engineering. We argue the need to identify and implement ways to foster students' curiosity and confidence regarding mathematics as mathematics underpins all STEM learning. Additionally, there is an obvious need to introduce engineering and technology subjects to students to be able to determine their real attitude towards these subjects. Recent research has highlighted the importance of value-related beliefs as a strong predictor of career aspirations in STEM

(Wang et al., 2013; Wegemer and Eccles, 2019). In our study, technology and engineering were perceived by students as having less value than science and mathematics, and this 'lower value' is also reflected in the challenges of reduced access to these subjects and the lack of clarity surrounding these subjects in the Irish post-primary system.

Gender differences

In our study, females had significantly more positive perceptions and attitudes to science compared to males. Science was perceived by female students as the most valuable STEM subject to learn, more appealing, interesting and enjoyable and with higher levels of confidence in science compared to other STEM subjects. Our survey did not differentiate between biology, chemistry, physics, and agricultural science as individual science subjects, which may have elicited different responses. In light of the significant gender differences in science subject enrolments at Senior Cycle in Ireland, the question remains as to how to harness this interest in science reported by females in our study and apply it to the individual scientific disciplines such as physics and chemistry, as well as to other STEM subjects. What exactly is it about science at Junior Cycle that attracts female students and yet does not translate into studying the physical sciences at Senior Cycle and beyond? There is future work to be done in answering these questions.

Female students in our study reported less positive attitudes to mathematics compared to male students in terms of liking, interest and value. Mathematics was perceived by all students as the most difficult STEM subject, although female students reported less positive perceptions about their abilities compared to males. This finding ties in with the findings of Brown et al. (2017), who reported that in instances where post-primary school students achieve the same grades, females' perceived ability in STEM subjects is lower than males. This lower perception of ability by females can be particularly potent in relation to future ability and may be a factor in females' avoidance of future studies or careers in these disciplines. Female students in our study also reported lower enjoyment of engineering compared to male students, while females also had significantly less positive perceptions of technology compared to males. This is likely linked to these female students having little exposure to technology and engineering at school, although that challenge also exists for male students.

A positive finding in our study was that no significant difference emerged between male and female students' perceptions of STEM careers. One possible explanation is that the participants were not yet at Senior Cycle and therefore had not decided on a career/college path. This may indicate the potential for interceding at this stage of post-primary education, prior to subject choice at Senior Cycle, with a view to enhancing awareness of STEM courses and careers and encouraging future participation and engagement with STEM. The Accenture (2014) report highlights that parents may have low career knowledge, which makes it difficult for them to offer career advice to their children, who in turn have trouble making decisions and understanding careers. Future work is required to design, implement and evaluate interventions aimed at students, particularly females, and potentially their parents to encourage informed decisions about future STEM subject/career choice.

Limitations of the study

There are some limitations of our study which we recognise and outline here. Firstly, a sampling bias in favour of females occurred as only schools that participated in the STEMChAT project were sampled and the project purposefully targeted female students in particular. Therefore, some of the schools that participated were all-girls schools. However, there were sufficient numbers of male students to allow for statistically valid gender comparison. The inclusion of single-sex girls' schools also highlighted the inequitable access to technology and engineering subjects experienced by female students. Secondly, the schools sampled in this study were all located in one province of Ireland. We do not claim that findings are representative of all students in Ireland. Rather, our findings provide some insight into these particular students' attitudes towards STEM while bringing to light some challenges that exist regarding subject access, attitudes and gender differences which need further research on a national scale. Thirdly, although STEM was defined as science, technology, engineering and mathematics, what was meant by each subject was not interrogated in the survey. For example, in relation to the term science, some students may have been answering the survey based on a particular aspect of science (e.g., physics) or interpreting the term as encompassing many different aspects of science. There was evident confusion among some students about what technology means, as previously discussed. Future data collection should seek to clarify and unpack each of the STEM disciplines.

Conclusion and recommendations

This paper aimed to identify students' attitudes to STEM as well as any gender differences with regards to these attitudes, and ultimately to determine challenges to female students' participation in STEM both at the post-primary (secondary) level and beyond. We surveyed 308 post-primary students in Ireland as part of a one-year SFI-funded research project "STEMChAT: Women as catalysts for change in STEM education." The results pointed to three key findings and challenges relating to (1) access to STEM subjects, (2) students' attitudes, and (3) gender differences, and we offer some recommendations in relation to these.

The results in this study highlight the challenge of equal access to STEM subjects in different types of post-primary schools. For example, many secondary schools, particularly the single-sex ones, have limited access to technology and engineering subjects. If our aim is to enhance the STEM subject and career involvement, the current Irish post-primary school system should be reformed to provide equal opportunities to all students, particularly girls, in order to access STEM subjects. Additionally, considerable confusion emerged among students about what STEM means regarding school subjects, particularly technology and engineering, which might be related to the access (or lack thereof) to these subjects in schools. STEM education stays as a contested term lacking a unified definition (Bybee, 2013; McLoughlin et al., 2020); we suggest developing greater transparency and consistency about what STEM means, especially in relation to school subjects.

Students in our study generally held positive attitudes towards science and less positive attitudes towards mathematics, while they expressed a lack of interest in engineering and no strong views on technology. Six schools (including 4 girls' schools) did not offer access to technology and engineering subjects, which conceivably might have affected these results. We suggest identifying and implementing ways to foster students' curiosity and confidence regarding mathematics and providing increased exposure to engineering and technology subjects to enable students to determine more informed attitudes towards these subjects.

Our study found that female students had significantly more positive attitudes towards science compared to males while, in comparison, males had a significantly more positive response to mathematics compared to females. Further research is needed to determine how to harness female students' positive attitudes towards science and cultivate this in the other STEM subjects. Additionally, no significant difference was found in this study between male and female students' perceptions of STEM careers. We suggest that an opportunity exists during Transition Year, prior to Senior Cycle subject choice, to enhance students' awareness of STEM subjects, third-level courses and careers in order to encourage future participation and engagement with STEM.

Overall, this study contributes to the STEM education field by identifying challenges to female students' participation in STEM in the Irish context and recommending future research to further understand and overcome them. In so doing, our paper contributes to the international discussion and agenda to provide "equal opportunities to access and benefit from quality STEM education" (UNESCO, 2017, p. 11).

Data availability statement

The datasets presented in this article are not readily available because of participant confidentiality. Requests to access the datasets should be directed to CL: shaciaral@gmail.com.

Ethics statement

The studies involving human participants were reviewed and approved by The University of Limerick Research Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

CL contributed to data collection, led data analysis and interpretation, manuscript design, and writing and revision. SK-C and RK contributed to survey design, data collection, and manuscript writing and revision. TO'C contributed to data collection and manuscript writing and revision. MG contributed to survey design and manuscript revision. All authors read the final manuscript. All authors contributed to the article and approved the submitted version.

Funding

The STEMChAT project was funded by a Science Foundation Ireland Discover Program Grant (Proposal ID 18/DP/5926). The funding body was not involved in the design, data collection, analysis, interpretation or writing of this manuscript.

Acknowledgments

The authors wish to thank the schools, teachers, community partners and participants who engaged with the STEMChAT research team and enabled data collection for this study. We also thank Michael Carey for providing assistance with statistical analysis.

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.

Publisher's note

References

Accenture (2014). *Powering economic growth: Attracting more young women into science and technology.* Available at: http://www.role.uk.com/uploads/5/2/8/8/52884055/accenture-stem-powering-economic-growth.pdf (Accessed May 24, 2021).

Ahlstrom, D. (2017). CAO 2017: Will STEM resurgence continue? *The Irish Times.* Available at: https://www.irishtimes.com/business/innovation/cao-2017-will-stem-resurgence-continue-1.2949475 (Accessed May 5, 2021).

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

Brown, R., Ernst, J., DeLuca, B., and Kelly, D. (2017). Engaging females in STEM. *Technol. Eng. Teach.* 77, 29–31.

BT Young Scientist and Technology Exhibition (2021). Available at: https://btyoungscientist.com/about/ (Accessed December 24, 2021).

Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities*. Arlington, VA: National Science Teachers Association (NSTA Press).

Central Statistics Office (2021). *Measuring Ireland's progress, 2019*. CSO Statistical Publication. Cork, Ireland. Available at: https://www.cso.ie/en/releasesandpublications/ep/p-mip/measuringirelandsprogress2019/introduction/ (Accessed April 14, 2021).

CoderDojo Foundation (n.d.) *CoderDojo*. Available at: https://coderdojo.com/ (Accessed December 24, 2021).

Department of Education and Skills [DES] (2017). *STEM Education Policy Statement 2017–2026*. Dublin: Department of Education and Skills. Available at: https://www.gov.ie/en/publication/0e94b-stem-education-policy-statement-20172026/ (Accessed January 11, 2022).

Eccles, J., and Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: a developmental, social cognitive, and sociocultural perspective on motivation. *Contemp. Educ. Psychol.* 61:101859. doi: 10.1016/j.cedpsych.2020.101859

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

Eurydice (2019). Organisation of upper secondary school. European Commission: Eurydice. Available at: https://eacea.ec.europa.eu/national-policies/eurydice/content/organisation-upper-secondary-education-5_en (Accessed May 4, 2021).

Falk, N. A., Rottinghaus, P. J., Casanova, T. N., Borgen, F. H., and Betz, N. E. (2017). Expanding women's participation in STEM: insights from parallel measures of self-efficacy and interests. *J. Career Assess.* 25, 571–584. doi: 10.1177/1069072716665822

Gladstone, K., and Cimpian, A. (2021). Which role models are effective for which students? A systematic review and four recommendations for maximizing the effectiveness of role models in STEM. *Int. J. STEM Educ.* 8:59. doi: 10.1186/s40594-021-00315-x

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 J. Res.* 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

Honey, M., Pearson, G., and Schweingruber, A. (2014). *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*. Washington, DC: National Academies Press.

Jungert, T., Hubbard, K., Dedic, H., and Rosenfield, S. (2019). Systemizing and the gender gap: examining academic achievement and perseverance in STEM. *Eur. J. Psychol. Educ.* 34, 479–500. doi: 10.1007/s10212-018-0390-0

Kelley, T. R., and Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3:11. doi: 10.1186/s40594-016-0046-z

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.

Luo, T., So, W. W. M., Wan, Z. H., and Li, W. C. (2021). STEM stereotypes predict students' STEM career interest via self-efficacy and outcome expectations. *Int. J. STEM Educ.* 8:36. doi: 10.1186/s40594-021-00295-y

Mahoney, M. P. (2010). Students' attitudes towards STEM: Development of an instrument for high school STEM-based programs. *J. Technol. Stud.* 36, 24–34. doi: 10.21061/jots.v36i1.a.4

Marginson, S., Tytler, R., Freeman, B., and Roberts, K. (2013). *STEM: Country Comparisons. International Comparisons of Science, Technology, Engineering and Mathematics (STEM) Education*. Melbourne: Australian Council of Learned Academies. Available at: https://acola.org.au/wp/PDF/SAF02Consultants/SAF02_STEM_FINAL.pdf (Accessed January 11, 2022).

Martin-Paez, T., Aguilera, D., Perales-Palacios, F. J., and Vílchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Sci. Educ.* 103, 799–822. doi: 10.1002/sce.2152

Master, A., Cheryan, S., and Meltzoff, A. N. (2014). Reducing adolescent girls' concerns about STEM stereotypes: when do female teachers matter? *Rev. Int. Psychol. Soc.* 27, 79–102.

McGarr, O., and Lynch, R. (2017). Monopolising the STEM agenda in second-level schools: exploring power relations and subject subcultures. *Int. J. Technol. Des. Educ.* 27, 51–62. doi: 10.1007/s10798-015-9333-0

McLoughlin, E., Butler, D., Kaya, S., and Costello, E. (2020). *STEM education in schools: what can we learn from the research?* ATS STEM Report #1. Dublin: Dublin City University. doi:10.5281/zenodo.3673728

Means, B., Wang, H., Wei, X., Young, V., and Iwatani, E. (2021). Impacts of attending an inclusive STEM high school: meta-analytic estimates from five studies. *Int. J. STEM Educ* 8, 1–19. doi: 10.1186/s40594-020-00260-1

Mohr-Schroeder, M., Bush, S. B., Maiorca, C., and Nickels, M. (2020). "Moving toward an equity-based approach for STEM literacy," in *Handbook of Research on STEM Education*. eds. C. Johnson, M. J. Mohr-Schroeder, T. Moore and L. English (New York: Routledge), 29–38.

Organization for Economic Co-operation and Development [OECD] (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. Paris: OECD.

RDS (2021). *STEM learning*. Available at: https://www.rds.ie/rds-foundation/science-and-technology/stem-learning (Accessed December 24, 2021).

Riegle-Crumb, C., Moore, C., and Buontempo, J. (2017). Shifting STEM stereotypes? Considering the role of peer and teacher gender. *J. Res. Adolesc.* 27, 492–505. doi: 10.1111/jora.12289

SciFest (2021). *SciFest – A National STEM Fair Programme for Schools*. Available at: https://scifest.ie/ (Accessed December 24, 2021).

Sobieraj, S., and Krämer, N. (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

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

The Scottish Government (2017). *Science Technology Engineering Mathematics: Education and Training Strategy for Scotland*. Edinburgh: The Scottish Government. Available at: https://www.gov.scot/binaries/content/documents/govscot/publications/strategy-plan/2017/10/science-technology-engineering-mathematics-education-training-strategy-scotland/documents/00526536-pdf/00526536-pdf/govscot%3Adocument/00526536.pdf

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

Tyler-Wood, T., Knezek, G., and Christensen, R. (2010). Instruments for assessing interest in STEM content and careers. *J. Technol. Teach. Educ.* 18, 345–368.

Tzu-Ling, H. (2019). Gender differences in high-school learning experiences, motivation, self-efficacy, and career aspirations among Taiwanese STEM college students. *Int. J. Sci. Educ.* 41, 1870–1884. doi: 10.1080/09500693.2019.1645963

United Nations Educational, Scientific and Cultural Organization [UNESCO] (2017). *Cracking the Code: Girls' and Women's Education in Science, Technology, Engineering and Mathematics (STEM)*. Paris: UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000253479

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 Aalderen-Smeets, S., and van der Molen, J. (2018). Modeling the relation between students' implicit beliefs about their abilities and their educational STEM choices. *Int. J. Technol. Des. Educ.* 28, 1–27. doi: 10.1007/s10798-016-9387-7

Van Tuijl, C., and van der Molen, J. H. W. (2016). Study choice and career development in STEM fields: an overview and integration of the research. *Int. J. Technol. Des. Educ.* 26, 159–183. doi: 10.1007/s10798-015-9308-1

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

Wegemer, C. M., and Eccles, J. S. (2019). Gendered STEM career choices: altruistic values, beliefs, and identity. *J. Vocat. Behav.* 110, 28–42. doi: 10.1016/j.jvb.2018.10.020

 10.3389/fpsyg.2022.959972

Appendix

Appendix 1. Kruskal–Wallis H-Test for differences between genders for STEM semantics scale.

Item	Science	Mathematics	Engineering	Technology	STEM career
Fascinating – Dull	$\chi^2(1) = 6.002$, $p = 0.014$	$\chi^2(1) = 2.814$, $p = 0.093$	$\chi^2(1) = 12.133$, $p = 0.000$	$\chi^2(1) = 6.485$, $p = 0.011$	$\chi^2(1) = 0.389$, $p = 0.533$
Appealing – Unappealing	$\chi^2(1) = 14.044$, $p = 0.000$	$\chi^2(1) = 1.709$, $p = 0.191$	$\chi^2(1) = 15.977$, $p = 0.000$	$\chi^2(1) = 7.718$, $p = 0.005$	$\chi^2(1) = 0.301$, $p = 0.584$
Exciting – Unexciting	$\chi^2(1)\ 5.901$, $p = 0.015$	$\chi^2(1) = 1.380$, $p = 0.240$	$\chi^2(1) = 11.003$, $p = 0.001$	$\chi^2(1) = 9.670$, $p = 0.002$	$\chi^2(1) = 0.008$, $p = 0.927$
Means Nothing – Means a Lot	$\chi^2(1) = 5.729$, $p = 0.017$	$\chi^2(1) = 1.263$, $p = 0.261$	$\chi^2(1) = 3.049$, $p = 0.081$	$\chi^2(1) = 0.841$, $p = 0.359$	$\chi^2(1) = 0.134$, $p = 0.715$
Boring – Interesting	$\chi^2(1) = 6.243$, $p = 0.012$	$\chi^2(1) = 4.638$, $p = 0.031$	$\chi^2(1) = 5.255$, $p = 0.022$	$\chi^2(1) = 2.288$, $p = 0.130$	$\chi^2(1) = 0.003$, $p = 0.958$

TYPE Original Research
PUBLISHED 01 December 2022
DOI 10.3389/feduc.2022.957716

OPEN ACCESS

EDITED BY
Karen Blackmore,
University of Worcester,
United Kingdom

REVIEWED BY
Jessica Gladstone,
New York University, United States
Luiz M. G. Gonçalves,
Federal University of Rio Grande do Norte,
Brazil

*CORRESPONDENCE
Antonia Ladewig
ladewig@leibniz-ipn.de

SPECIALTY SECTION
This article was submitted to
STEM Education,
a section of the journal
Frontiers in Education

RECEIVED 31 May 2022
ACCEPTED 07 November 2022
PUBLISHED 01 December 2022

CITATION
Ladewig A, Köller O and Neumann K (2022) Stereotypes in the German Physics Olympiad - Hurdle or no Harm at all? *Front. Educ.* 7:957716.
doi: 10.3389/feduc.2022.957716

Stereotypes in the German Physics Olympiad - Hurdle or no Harm at all?

Antonia Ladewig[1]*, Olaf Köller[2] and Knut Neumann[1]

[1]Department of Physics Education, Leibniz Institute for Science and Mathematics Education, Kiel, Germany, [2]Department of Educational Research and Educational Psychology, Leibniz Institute for Science and Mathematics Education, Kiel, Germany

The German Physics Olympiad is a science competition in which students can compete to measure their Physics knowledge and skills with other students. Female participants are underrepresented and typically drop out of the competition earlier than their male counterparts. As the cause for this underrepresentation, social identity threat theory identifies a threat to women's gender identity in the predominantly male environment. Stereotype threat theory adds negative stereotypes about women's abilities in physics as a heightening factor. In this study, growth mindset and values affirmation interventions, as well as a combination of both methods, were integrated into a weekend seminar of Physics content to protect female participants from the harmful influences of stereotype and social identity threat. As female and male students' sense of belonging and gender identification remained at equal levels, respectively, after the interventions, the results did not show any effects of stereotype threat or social identity threat for the female students. The results suggest that women who are highly interested and talented in physics and have taken first steps to pursue physics and to engage with the physics community beyond mandatory school education are not as susceptible to stereotypes and harmful cues in the environment as might previously have been assumed. Implications for future research and science competitions are discussed.

KEYWORDS

stereotypes, social identity threat, stereotype threat, science competitions, physics

Introduction

Why are females still a minority in physics? One reason for the underrepresentation of women that is discussed in the literature is stereotype threat, that is, the phenomenon that minorities start to unconsciously live up to negatively stereotyped behavior in fields consisting of a predominant majority (e.g., Steele and Aronson, 1995; Steele et al., 2002; Hall et al., 2015; Bedyńska et al., 2018). Women are affected by this as they are underrepresented in many science environments and faced with stereotypes of lacking science talent.

Previous research introduced various interventions to fight stereotype threat (e.g., Cheryan et al., 2011; Lin-Siegler et al., 2016). As the mere presence of negative stereotypes might lead to stereotype threat effects (e.g., Huguet and Régner, 2009; Marchand and Taasobshirazi, 2013) and its various negative consequences (e.g., Schmader and Johns, 2003; Hall et al., 2015, 2018), interventions aim to shield minority groups from stereotype threat instead of eliminating stereotypes from the environment. Especially two short intervention methods showed noticeable success: First, interventions that aim to reduce the impact of stereotypes by teaching students about the malleability of the brain, the use of effort, and struggle to gain success (growth mindset; e.g., Blackwell et al., 2007; Yeager et al., 2016), and, second, interventions that affirm participants in their values (values affirmation; e.g., Cohen et al., 2006, 2009).

However, these interventions have not yet been tested with the important sample of students who chose to engage in extracurricular science activities. These students show interest and intention to engage in the domain by pursuing science outside the mandatory school curriculum. Nonetheless, female students who participate in such extracurricular science activities experience stereotype threat (see Ladewig at al., 2020).

In this study, participants in the German Physics Olympiad, a physics competition, were invited to physics seminars. In these seminars, students participated in either an intervention of growth mindset, an intervention of values affirmation, or a combination of both. We investigated how the different short interventions impacted sense of belonging, social identity threat, stereotype endorsement, perceptions of environmental stereotyping, and gender identification.

Theoretical background

The German Physics Olympiad presents a science environment, which students encounter outside of their mandatory school education. Students decide freely whether they want to participate in the competition. However, female students decide to do so not as often as the male students. Which are the driving influences leading to this gender gap?

Underrepresentation of women in science

Females in science face stereotypes about women having supposedly lower talent or not fitting within the male in-group. In line with the stereotypical association of sciences with the male gender (see Makarova et al., 2019), only about every fifth academic in science, technology, engineering, or mathematics fields in Germany is female (Anger et al., 2019). This underrepresentation of women begins with decreasing science interest in school for girls but not for boys (Sadler et al., 2012) and ends with low numbers of women in science careers (e.g., Kahn and Ginther, 2015; Miller and Wai, 2015; Su and Rounds, 2015). Although the gender gap is less pronounced in some domains, it is very distinct in physics (e.g., Düchs and Mecke, 2019). But why is science such a hindering environment for females?

Science provides cues that drive many women from the field. A person remains in an environment if their perception of their self fit to the person stereotypically expected in the environment (e.g., Setterlund and Niedenthal, 1993; Hannover and Kessels, 2004). Hannover and Kessels (2004) showed that the prototypical student who prefers stereotypically female subjects such as humanities to stereotypically male subjects such as physics is perceived in a more positive manner than a student who prefers science subjects. Additionally, perceived similarity to other members of a field moderates women's interest in continuing in a stereotypically male domain (Cheryan and Plaut, 2010). Therefore, women in science face a high hurdle when intending to stay in this male environment because science appears rather male and unpopular. The stereotypes contradict their fit to the field and encourage women to leave.

Nevertheless, several women master this hurdle and begin to pursue science. Science competitions provide one opportunity for interested students to engage in science domains outside of school. In Germany, 9,065 students participated in in the 2019 Science Olympiads, which include competitions in several science domains.

Yet, gender differences are visible and especially pronounced in the German Physics Olympiad. The German Physics Olympiad is organized into four consecutive rounds. Each round requires students to solve physics tasks or do experiments. The initial registration for the competition is open to any interested student, who is still in school and within the yearly set age limit. In the first round, physics teachers receive a best practice solution for the tasks and judge the participants' solution according to this. Only those participants, who solved the tasks best, continue from there on to the next rounds. Thus, the number of participants continually decreases. Eventually, the best five students form the national team for the international competition. The Physics Olympiad faces two connected problems concerning gender representation. First, fewer females than males choose to participate. In 2018, only 28% of all participants were female. Second, these female participants drop out of the competition in disproportionally higher numbers than the male participants do. This often leads to all-male national teams, as was the case in 2019.

Stereotype threat and social identity threat

What causes this underrepresentation of female participants? An important factor in the context of science competitions is stereotypes (Steegh et al., 2019). Stereotype threat theory explains how stereotypes impact behavior.

Stereotype threat occurs to minority groups that enter a field in which they are underrepresented and faced with negative stereotypes about their groups' characteristics or abilities (e.g., Hall et al., 2015; Bedyńska et al., 2018). It is hypothesized that the negative stereotypes trigger stereotype threat, which inhibits the members of the stereotyped group from performing to their full potential (e.g., Steele and Aronson, 1995; Steele et al., 2002). Implicit and explicit cues in the environment about the stereotypes' eligibility can induce this mechanism (e.g., Spencer et al., 1999; Marchand and Taasobshirazi, 2013).

Stereotype threat effects have been demonstrated for several groups, from women in sciences (e.g., Miller et al., 2015; Smyth and Nosek, 2015), to males in typically female jobs, to older employees in working life (e.g., Hartley and Sutton, 2013; Froehlich et al., 2016; Kalokerinos et al., 2017; Rahn et al., 2020). All of these groups face stereotypes regarding either their abilities or their lacking fit within their chosen environment, e.g., that women lack talent for science and are consequently not able to perform as good as men. Negative stereotypes lead not just to lower performance levels (Steele and Aronson, 1995; Shih et al., 1999; Flore and Wicherts, 2015) but also to various other negative changes such as the minority members' stronger wishes to leave the environment (Kalokerinos et al., 2017).

Although women try to counteract the negative stereotypes in science to prove their falsehood (Jamieson and Harkins, 2011), they still feel less accepted, more mentally exhausted, and less competent (Schmader and Johns, 2003; Hall et al., 2015). Stereotype threat is also connected to heightened anxiety (Ben-Zeev et al., 2005) and burnout (Hall et al., 2018).

Social identity threat intensifies negative effects even further (see van Veelen et al., 2019). Stereotype threat is a specific theory within the theoretical framework of social identity threat (Schmader et al., 2015). The general feeling of being different from the majority group because of one's social identity causes social identity threat (e.g., Steele et al., 2002; Hall et al., 2018). Schmader (2002) showed that gender identification can explain performance differences between males and females in science. Women who strongly identified with their gender performed tasks worse than men if the tasks were linked to gender identity. Women with lower gender identification performed at the same level as men. As women identify with the negatively stereotyped gender identity more easily when in science environments (Marx et al., 2005), gender identification can lead to stereotype threat and, consequently, lower performance (e.g., Shih et al., 1999; Flore and Wicherts, 2015). Endorsing negative stereotypes and believing in their eligibility further heightens these effects (Schmader et al., 2004).

Sense of belonging

Stereotypes about females in physics also go beyond stereotype threat. Members of a stereotyped group that enter a situation in which stereotypes are present then doubt their abilities and are more likely to interpret lower performance as the result of missing fit within the environment (Aronson and Inzlicht, 2004). When doubting their abilities, individuals perceive features in the environment that could justify their doubt. Belonging uncertainty — the feeling of not being sure whether they fit within the group (Walton and Cohen, 2007) — can lead to individuals distancing themselves from the group, environment, and tasks. For example, female students who perceive their environment as negatively stereotyping women in math feel less belonging within the math environment (Good et al., 2012). This can lead to negative performance and, thereby, to the unintentional confirmation of the negative stereotypes (Steele and Aronson, 1995).

Sense of belonging, which is the feeling of connection, membership, trust, participation, positive affect, and acceptance in a group (Good et al., 2012), is closely connected to a wide range of variables relevant for academic success. Among others, value of school (Gillen-O'Neel and Fuligni, 2013), intrinsic motivation (Freeman et al., 2007), and academic adjustment in college (Pittman and Richmond, 2007) are linked to belonging. Consequently, belonging uncertainty is especially disadvantageous in situations that require high performance and achievements.

Women could benefit from feeling more belonging in science. Murphy et al. (2007) showed that women had lower sense of belonging and fewer intentions to participate in science settings that appeared to be predominantly male than in settings with equal numbers of male and female participants. Because men choose science environments more often, they create a predominantly male environment that is associated with male characteristics. Thus, as the sense of belonging influences a woman's decision to leave science or not (e.g., Good et al., 2012; Banchefsky et al., 2019), the association of male characteristics with science is likely to reduce women's aspirations to continue in science fields (e.g., Makarova et al., 2019).

Interventions against stereotype threat

To interrupt the ongoing cycle of stereotype threat and self-selection out of science, a supportive identity and system of values need to be formed; these are predictors of persistence in a domain (Estrada et al., 2011). Especially short psychological interventions that can be implemented easily in various situations (e.g., Yeager et al., 2013; Brady et al., 2020) can be valuable because they can be implemented in school, college and even in physics competitions. Growth mindset and value affirmation interventions both fall into this category and have been shown to be beneficial in protecting participants against stereotype threat and social identity threat.

Growth mindset

One common stereotype is that females do not have science talent. Especially in sciences that require a lot of mathematics such as physics, people endorse the opinion that success is built on talent (Deiglmayr et al., 2019; see also Archer et al., 2020; Johansson, 2020). This, however, results in a vicious cycle: Fewer women in a field are connected to higher beliefs in talent as the basis for success in this domain (Bailey et al., 2019). As stereotypes drive women to leave physics, the predominantly male gender ratio persists and consequently strengthens the stereotypes about female talent, which again drives more women to leave.

How can an intervention break this cycle? Aiming to change implicit theories of intelligence takes away the basis of the assumption that talent is essential for success.

Implicit theories of intelligence can be divided into two groups: Entity theories, which assume that characteristics such as cognitive abilities cannot be changed, and incremental theories, which assume those characteristics change through effort and work (e.g., Blackwell et al., 2007; Burnette et al., 2013). Entity theories of intelligence tend to promote stereotypes about talent in science. Yeager et al. (2016) tested growth mindset interventions that promoted an incremental theory of intelligence. These interventions had three successive steps: First, participants read information on how intelligence can be enhanced. Second, they found examples in their personal experience where the learned information applied and, lastly, they wrote a letter that encouraged other students to handle struggles based on this information. Students' belief changed from an entity theory to an incremental theory of intelligence after the intervention.

Also, stereotypes affected the performance of college students less after such interventions (e.g., Alter et al., 2010), thus showing reduced stereotype threat. Students who held an incremental theory of intelligence also enjoyed and engaged more in academic work (Aronson et al., 2002).

Values affirmation

Female students who enter a physics environment are perceived as a minority whose fit and belonging are threatened and questioned (e.g., Aronson and Inzlicht, 2004). This uncertainty can be reduced with values affirmation interventions that heighten the fit of self and situation (e.g., Cohen et al., 2006). Cohen et al. (2006) asked a stereotypically threatened minority to participate in a values affirmation intervention that aimed to achieve self-integrity and an unthreatening environment. Students received a list of values and were asked to choose the personally most important one and write about why it was important to them. Their results showed that the achievement gap, created by racial stereotypes, went down by 40% after the intervention. The activation of stereotypes in performance situations was also lowered. Similar effects were shown by Cohen et al. (2009) along with the intervention's long-term positive impact on performance. Nevertheless, several studies find no effects of the interventions (de Jong et al., 2016; Bayly and Bumpus, 2019) or even goal disengagement (Vohs et al., 2013). Still, in physics contexts, Miyake et al. (2010) found beneficial effects for women, especially if these endorsed stereotypes about women in science.

The present study

How are females in physics affected by stereotypes and how can we help them pursue their interest in physics? To address these questions, the current study aimed to analyze two different interventions against stereotype threat in the context of the German Physics Olympiad. The results of previous research suggest that brief growth mindset and values affirmation interventions are useful in combating stereotype threat for school (e.g., Cohen et al., 2009; Yeager et al., 2016) and college students (e.g., Aronson et al., 2002; Miyake et al., 2010). However, to the best of our knowledge, these interventions have not been tested in science competitions.

The Physics Olympiad presents a typically predominantly male science environment, while at the same time presenting a selective sample of participants who have shown ongoing interest in science by entering the competition and pursuing physics knowledge beyond school education. Although the competition does not explicitly broadcast stereotypes about women having lacking talent for science, previous research has indicated that stereotype threat has negative effects on female participants of science competitions (e.g., Steegh et al., 2019; Ladewig et al., 2020). It seems advisable to implement growth mindset and values affirmation interventions to protect female participants of the German Physics Olympiad from the potential damage by those stereotypes. Implementing the interventions during the participants' first encounter with other Physics Olympiad participants enables the assessment of whether these interventions prevent stereotype threat because it marks female students' first exposure to the predominantly male environment. The encounter should heighten stereotype threat and social identity threat effects as here, female participants are first personally entering the apparently stereotyping and predominantly male environment of the competition. Researching both participants' personal agreement with stereotypes and their perception of other's beliefs in stereotypes within the Physics Olympiad seems advisable.

Based on previous research on stereotype threat, social identity threat, and sense of belonging, we formed several hypotheses. The first hypothesis focused on the benefits of the interventions on the variables that stereotype threat and social identity threat are theorized to impact. We assumed that both interventions would have equally beneficial effects but that the best results would be achieved by combining the two interventions. Those effects were expected to continue after the seminar. As, however, we do not expect male participants to suffer under any social identity or stereotype threat effects, which the interventions aim to counteract, they should also not experience changes by participating in the interventions.

> *H1a*: We hypothesized that female students would rate sense of belonging higher and gender identification lower directly after the interventions as well as several weeks after the interventions than before the interventions. For male students, no changes were expected.
> *H1b*: We expected that the combination of both interventions would have stronger effects, that is, would lead to a stronger increase in sense of belonging and a stronger decrease in gender identification.

Further, we aimed to compare the female and male participants' assessments of variables crucial for stereotype threat and for perceived social identity threat with the study. Therefore, the second hypothesis can be split into three parts based on the environment the participants were facing. At the first measurement point, students had not yet met the other participants and made their assessments based on previously experienced physics environments. These perceptions were expected to be stereotyped on the basis of participants' experience from school and media. This study used an explicit measure of stereotype endorsement; it needs to be considered that the assessments, especially those of the female participants, might differ from implicit measures (*cf.* Kessels et al., 2006). Women might feel less inclined to explicitly agree with negative stereotypes about women in science than male participants do and might indicate lower beliefs in negative stereotypes about women. An explicit measure seems, nevertheless, appropriate as higher stereotype endorsement was previously shown to predict higher sensitivity to the consequences of stereotype threat (e.g., Schmader et al., 2004; Pennington et al., 2016).

> *H2a*: We hypothesized that, before the interventions, females would endorse negative stereotypes about women in physics less than males. Here, female should perceive more social identity threat than males.

When the participants arrived at the seminar, they encountered a predominantly male group. This again should present a stereotypical environment. The literature suggests that in predominantly male science groups, females identify more with their gender (e.g., Schmader, 2002) while also feeling less belonging and perceiving the environment to be stereotyping (e.g., Good et al., 2012).

> *H2b*: We hypothesized that, after the first meeting with other participants of the German Physics Olympiad, females would perceive stronger environmental stereotyping than males. We expected females to endorse negative stereotypes about women in physics less than males.

The interventions took place in the seminars. The interventions aimed to reduce perceptions of stereotyping for the female participants but not for the male participants and we thus did not expect women to perceive stereotypical cues in the environment differently than the male participants.

> *H2c*: We hypothesized that, directly after the interventions, females would have equal perceptions of environmental stereotyping and social identity threat to males in the physics environment. We expected females to endorse negative stereotypes about women in physics less than males.

Materials and methods

To study the hypotheses, we conducted a study within two successive years of the German Physics Olympiad.

Project "Identiphy – Identity development in physics!"

The study presented in this paper was part of the larger project "Identiphy – Identity development in physics!" (see also Ladewig et al., 2022). The project included a longitudinal study with two cohorts of German Physics Olympiad participants. Participation was voluntary. The study included four measurement points that took place before and after a weekend-long seminar, which was advertised to teach additional physics knowledge and give participants the chance to meet other participants of the German Physics Olympiad. The study took place directly after the first competition round, which consisted of solving physics tasks at home, so students encountered other participants for the first time at the seminars. Thereby, we aimed to study the regular competition conditions: Students, most often, encounter other highly interested and talented students for the first time when entering the higher competition's rounds. This was replicated with our placing of the study after filling in the first tasks at home. Also, we did not explicitly trigger stereotypes. In a normal competition, this would also not happen. Students just perceive the regular broadcasting of the stereotypes in the competition and environment.

Participants

All participants in the German Physics Olympiad received an invitation to participate in voluntary weekend seminars. Invitations went out *via* e-mail or letter, and participants were informed that declining participation would not lead to any disadvantages in the competition and that questionnaires would remain anonymous. All students or, if they were underaged, their legal guardians provided informed consent before participation.

Overall, 298 students participated. Of these, 82 were female (age: $M = 15.87$, $SD = 1.22$) and 216 male (age: $M = 15.93$, $SD = 1.36$). 167 students participated in the first (age: $M = 15.87$, $SD = 1.26$; 42 female, 125 male) and 131 in the second cohort (age: $M = 15.95$, $SD = 1.40$; 40 female, 91 male participants). Students could choose one out of six weekend seminars in the first year and one out of four weekend seminars in the second year. At the weekend seminars, participants formed 11 groups in the first and 10 groups in the second year. Each group was randomly assigned to an intervention approach. In each case, the whole group was assigned to one approach because each intervention demanded different amounts of time and could only be easily included in the program if all students within one group participated in the same intervention. Table 1 shows the numbers of participants in each intervention method for both years and the overall sample, who were included in the analyses based on the intervention groups they could be clearly associated with.

We deferred from separating the cohorts in the analysis because they did not vary significantly in gender

$$\chi^2[1] = 1.13,\ p = 0.287$$

or age

$$t[264] = -0.51,\ p = 0.609.$$

Using G*Power 3.1 (Faul et al., 2007, 2009) to calculate a sensitivity analysis, the sample size of 278 participants, whose assessments were clearly allotted to the three intervention groups, has a critical population effect size of 0.08 for comparing two groups, in this case male and female participants, for an analysis of variance. The sample size of 278 participants has a critical population effect size of 0.10 for comparing three groups with a mixed analysis of variance.

Procedure and intervention methods

Figure 1 gives an overview of the study proceedings and the scales. Students were able to choose one of the four cities in which the seminar took place according to their own preferences. Students registered for the seminar and received further information on the study as well as informed consent forms. Following the submission of those forms, students filled in a questionnaire online and received preparatory materials covering the physics content of the seminar.

Students arrived at the seminars without having knowledge of the planned intervention. They were first assigned to groups before filling in the second questionnaire. Next, the interventions took place. At the seminar, these questionnaires and interventions took place on paper under supervision of the seminar staff.

The growth mindset intervention was adapted from Yeager et al. (2016). Participants read a text that explained how learning changes the brain, how to improve performance, and how to handle struggles. Participants repeated the taught information by answering two questions about the text's content before writing about a personal experience where this information could be or had been used. The writing task was not limited in length.

The values affirmation intervention was adapted from Cohen et al. (2006). Participants chose one value from a list of 13 (e.g., "my family and friends," "being intelligent") and explained its importance and meaning for their life and possibly for their interest in physics. The writing task was not limited in length.

In both interventions, we did not study the students' results of the writing task. We wanted the students to feel free to write anything personal without feeling hindered by knowing someone else would read it.

Every student participated in an intervention. We thus assumed stereotype threat in the sample as previous research has shown that female participants in science competitions suffer under negative stereotypes (e.g., Steegh et al., 2020). No untreated group was implemented, which is why we can only compare the differences between the intervention methods. Thereby, we also wanted to ensure that every student had the chance to benefit from an intervention as stereotype threat, based on previous literature, is to be expected within the whole sample.

At the end of the second seminar day, students again filled in a questionnaire. Approximately 6 and 12 weeks after the seminar students were given the chance to do further physics tasks, which recaptured the seminar contents, and to receive feedback on their solutions. The feedback was focused on the way the students were able to apply the seminars' new physics content on other physics task. Thereby, the feedback was written as a positive support to further use the new knowledge, while still pointing out where mistakes were made and how to further improve in applying the knowledge. Subsequently, students were asked to fill in the last questionnaire online.

TABLE 1 Demographic data of the participants, split by intervention groups.

			Cohort 1	Cohort 2	Overall sample
Growth mindset	Age	*M* (*SD*)	15.96 (1.29)	16.35 (1.23)	16.13 (1.22)
	Gender	*N* male/female	50 (37/13)	37 (25/12)	87 (63/25)
Values affirmation	Age	*M* (*SD*)	15.92 (1.32)	15.73 (1.53)	15.84 (1.41)
	Gender	*N* male/female	38 (31/7)	30 (22/8)	68 (53/15)
Combination	Age	*M* (*SD*)	15.76 (1.28)	15.85 (1.38)	15.80 (1.33)
	Gender	*N* male/female	70 (51/19)	53 (35/17)	123 (86/36)

Measures

The questionnaires included five scales, which were used for the analyses.

Sense of belonging

Sense of belonging was measured on a scale adapted from Good et al.'s (2012) Math Sense of Belonging scale. All 30 items were ranked from 1 (*strongly disagree*) to 5 (*strongly agree*). In the first and last questionnaires, all items referenced the group of the participants' school physics class (e.g., "When I am in my physics lessons, I feel that I belong to the group."; first questionnaire: Cronbach's $\alpha = 0.63$; last questionnaire: Cronbach's $\alpha = 0.94$),

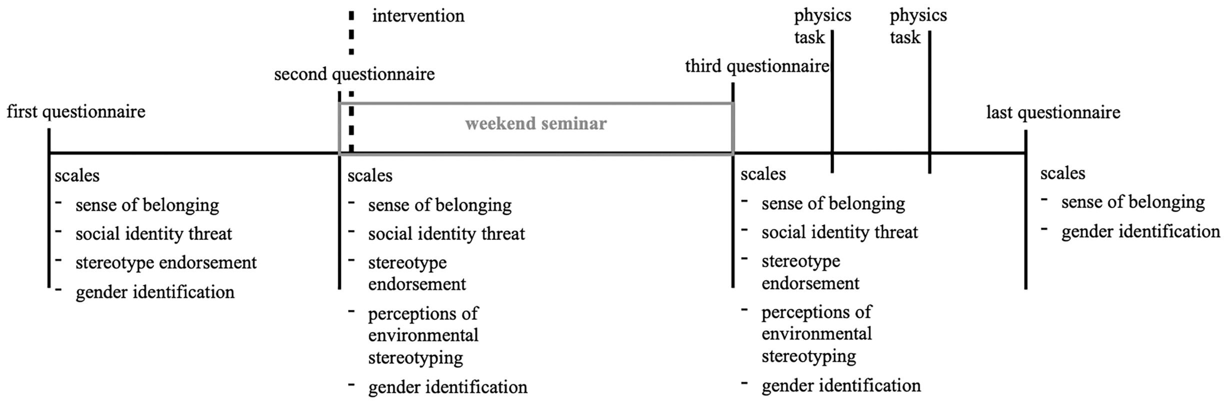

FIGURE 1
Graphical depiction of the study proceedings including the scales, which were used for the analyses.

whereas, upon arrival at the weekend seminar, the seminar group was the reference group for the first impression (e.g., "At the moment, I feel that I belong to the seminar group."; Cronbach's $\alpha = 0.91$) and at the end of the seminar (e.g., "During the weekend seminar, I feel that I belong to the group."; Cronbach's $\alpha = 0.91$).

The questionnaire consisted of five subscales measuring trust, acceptance, negative affect (reverse coded), desire to fade (reverse coded), and membership. As suggested by the authors of the scale (Good et al., 2012), we used the scale without splitting it further into subscales.

Social identity threat

Social identity threat was measured with a scale adapted from Rattan et al. (2018), which is itself an adapted version of a scale from Steele and Aronson (1995). The scale consists of four items (e.g., "My gender influences the perception that others have of my physics abilities"). The items were assessed on a scale ranging from 1 (*Not at all*) to 5 (*Extremely*). High values on this scale indicate high social identity threat. Cronbach's alpha was 0.82 in the first questionnaire, 0.80 at the beginning of the seminar, and 0.76 at the end of the seminar.

Stereotype endorsement

Stereotype endorsement was measured with a scale by Schmader et al. (2004) consisting of three items, which were adapted to physics (e.g., "It is possible that men have greater physics ability than women do."). Items were ranked from 1 (*strongly disagree*) to 5 (*strongly agree*). High values indicate high stereotype endorsement. Cronbach's alpha was 0.71 in the first questionnaire, 0.69 at the beginning of the seminar, and 0.69 at the end of the seminar.

Perceptions of environmental stereotyping

Perceptions of environmental stereotyping were measured with a shortened 4-item version of a scale by Good et al. (2012), which is an adapted version of a scale by Fennema and Sherman (1976). Items (e.g., "The other students in my seminar group believe that men are naturally better in physics than women.") were ranked from 1 (*I do not agree*) to 5 (*I agree*). Cronbach's alpha was 0.81 at the beginning and 0.82 at the end of the seminar.

Gender identification

Gender identification was measured with four items from a scale by Schmader (2002), which is an adapted version of a scale by Luhtanen and Crocker (1992). Items (e.g., "Being a male/female is important for the perception I have of myself.") were ranked from 1 (*strongly disagree*) to 5 (*strongly agree*). Cronbach's alpha ranged between 0.80 (first questionnaire), 0.83 (beginning of seminar), 0.83 (end of seminar), and 0.87 (last questionnaire).

Growth mindset

A scale to assess students' growth mindset was included in the study but not further for the here mentioned analyses. The scale was based on a scale by Dweck (2000). The data for the intervention group, which was supposed to achieve a growth mindset, showed high values previous to the intervention in the first questionnaire ($M_{male} = 3.68$, $SD = 0.48$; $M_{female} = 3.67$, $SD = 0.47$) and very high values in the last questionnaire after the interventions ($M_{male} = 4.16$, $SD = 0.83$; $M_{female} = 4.15$, SD = 0.76).

Results

The analyses were, first, focusing on the effects of the interventions, and, second, on the gender differences in stereotype threat and social identity threat.

Effects of the interventions

With regard to Hypotheses 1a and 1b, a mixed analysis of variance was calculated in SPSS (version 26, IBM Corp., 2015) for sense of belonging and gender identity across all four measurement points. The means and standard deviations can be found in Table 2 for sense of belonging and in Table 3 for gender identity.

First, a significant main effect of measurement point was found on belonging, $F(3, 471) = 379.34$, $p < 0.001$, but not on gender identification, $F(3, 471) = 0.46$, $p = 0.713$. There was no significant main effect of gender on either belonging, $F(1, 157) = 1.34$, $p = 0.249$, or gender identification, $F(1, 157) = 0.16$, $p = 0.695$. There was also no significant main effect of the intervention group on belonging, $F(2, 157) = 1.36$, $p = 0.259$, or on gender identification, $F(2, 157) = 0.39$, $p = 0.681$. These results only partly support Hypothesis 1a by showing changes in sense of belonging but no gender differential effect on either variables or changes over time for gender identification.

Hypothesis 1b is also not supported by these results as the intervention groups did not differ. Looking more closely, the interaction term gender x intervention group did not have a significant effect on belonging, $F(2, 157) = 0.55$, $p = 0.578$, or on gender identification, $F(2, 157) = 0.036$, $p = 0.965$. The further interaction term measurement point × gender also did not show a significant effect on belonging, $F(3, 471) = 1.89$, $p = 0.131$, or on gender identification, $F(3, 471) = 0.50$, $p = 0.685$, indicating no significant differences between the changes in the groups along the measurement points. The third interaction term measurement point x intervention group showed a significant effect on belonging, $F(6, 471) = 2.97$, $p = 0.008$, but not on gender identification, $F(6, 471) = 0.91$, $p = 0.486$. The interaction term measurement point x gender x intervention group did not show a significant effect on belonging, $F(6, 471) = 0.467$, $p = 0.828$, or on gender identification, $F(6, 471) = 1.21$, $p = 0.301$, indicating that all groups experienced the study similarly.

Gender differences in stereotype threat and social identity threat

To test Hypotheses 2a, b, and c, analyses of variance were calculated for each measurement point to compare the assessments of male and female participants and the assessments between the intervention groups. The means and standard deviations, split by gender and intervention group, can be found in Table 4.

Hypothesis 2a focused on the first measurement point. Our results show a significant main effect of gender on social identity threat, $F(1, 259) = 15.53$, $p < 0.001$, but not on stereotype endorsement, $F(1, 259) = 3.56$, $p = 0.060$. The intervention group did not have a significant main effect on either social identity threat, $F(2, 259) = 1.49$, $p = 0.227$, or stereotype endorsement, $F(2, 259) = 0.17$, $p = 0.842$, nor did the interaction term gender × intervention group prove significant for social identity threat, $F(2, 259) = 0.09$, $p = 0.916$, or stereotype endorsement, $F(2, 259) = 1.01$, $p = 0.367$. The results indicate a confirmation of the hypothesis regarding social identity threat, which was rated significantly higher by females than males.

Next, Hypothesis 2b targeted the measurement point at the beginning of the first seminar day. Our results show a significant main effect of gender on social identity threat, $F(1, 268) = 9.39$, $p = 0.002$, but not on stereotype endorsement, $F(1, 268) = 1.56$, $p = 0.213$, or on perceptions of environmental stereotyping, $F(1, 268) = 2.74$, $p = 0.099$. The intervention group did not have a significant main effect on social identity threat, $F(2, 268) = 0.76$, $p = 0.467$, stereotype endorsement, $F(2, 268) = 0.46$, $p = 0.629$, or perceptions of environmental stereotyping, $F(2, 268) = 0.12$, $p = 0.988$. Lastly, the interaction term gender x intervention group also did not prove to be significant for social identity threat, $F(2, 268) = 0.18$, $p = 0.836$, stereotype endorsement, $F(2, 268) = 0.16$, $p = 0.852$, or perceptions of environmental stereotyping, $F(2, 268) = 0.43$, $p = 0.653$. Again, these results mostly contradict the hypothesis, while the results for social identity threat — with a higher mean for females — confirm it.

Finally, Hypothesis 2c assumed changes would occur in the assessment of the variables due to the interventions, which is at the third measurement point. Again, a significant main effect of gender was found on social identity threat, $F(1, 267) = 6.63$, $p = 0.011$, but not stereotype endorsement, $F(1, 267) = 3.28$, $p = 0.071$, or perceptions of environmental stereotyping, $F(1, 267) = 0.15$, $p = 0.703$. The intervention group did not have a significant main effect on social identity threat, $F(2, 267) = 0.53$,

TABLE 2 Means and standard deviations of sense of belonging for all measurement points, split by gender and intervention method.

	Growth mindset				Values affirmation				Combination			
	Boys		Girls		Boys		Girls		Boys		Girls	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
First assessment point	3.14	0.19	3.07	0.18	3.07	0.14	3.08	0.21	3.08	0.19	3.07	0.21
Second assessment point	4.33	0.42	4.16	0.54	4.13	0.44	4.13	0.42	4.36	0.4	4.43	0.29
Third assessment point	4.52	0.38	4.41	0.41	4.26	0.59	4.36	0.39	4.41	0.39	4.45	0.45
Fourth assessment point	4.3	0.4	4.14	0.54	4.18	0.6	4.02	0.42	4.13	0.53	3.92	0.47

TABLE 3 Means and standard deviations of gender identification for all measurement points, split by gender and intervention method.

	Growth mindset				Values affirmation					Combination		
	Boys		Girls		Boys		Girls			Boys	Girls	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
First assessment point	2.88	1.23	2.79	1.03	2.56	0.77	2.85	1.11	2.64	0.91	2.4	0.85
Second assessment point	2.8	1.14	2.92	1	2.54	1.07	2.58	0.9	2.5	1.16	2.83	1.19
Third assessment point	2.59	1.28	2.67	1.02	2.55	1.02	2.6	1.3	2.47	1.07	2.76	1.31
Fourth assessment point	2.76	1.19	2.78	1.17	2.47	1.34	2.63	1.35	2.69	1.08	2.59	1.07

TABLE 4 Means and standard deviations of perceived social identity threat, stereotype endorsement, and perceptions of environmental stereotyping for the first three measurement points, split by gender and intervention group.

		Growth mindset				Values affirmation				Combination			
		Boys		Girls		Boys		Girls		Boys		Girls	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Social identity threat	First assessment point	1.7	0.82	2.23	0.85	1.5	0.71	1.92	1.02	1.59	0.74	2.03	1.01
	Second assessment point	1.75	0.84	2	0.69	1.58	0.67	1.98	1.01	1.77	0.76	2.11	0.87
	Third assessment point	1.77	0.87	2.02	0.61	1.8	0.86	2.14	1.03	1.84	0.81	2.18	0.92
Stereotype endorse-ment	First assessment point	2.19	1.08	1.79	0.81	2.07	0.94	2.11	0.83	2.26	1.02	1.84	0.8
	Second assessment point	2.12	1.05	1.85	0.98	2.15	0.99	2.09	0.86	2.22	1	2.03	0.94
	Third assessment point	2.18	1.07	1.88	0.96	2.31	1.12	2.12	1.05	2.4	1.05	2.04	0.98
Perceptions of environ-mental stereo-typing	Second assessment point	1.67	0.72	2	0.91	1.75	0.74	1.88	1	1.77	0.78	1.88	0.89
	Third assessment point	1.68	0.8	2.13	0.88	1.95	0.95	1.7	0.87	1.74	0.81	1.74	0.89

$p = 0.589$, on stereotype endorsement, $F(2, 267) = 0.67$, $p = 0.514$, or on perceptions of environmental stereotyping, $F(2, 267) = 0.55$, $p = 0.580$. The same result was found for the interaction term gender x intervention group, with no significant effect found on either social identity threat, $F(2, 267) = 0.14$, $p = 0.873$, stereotype endorsement, $F(2, 267) = 0.09$, $p = 0.917$, or perceptions of environmental stereotyping, $F(2, 267) = 1.91$, $p = 0.150$. Whereas the results regarding environmental stereotyping and stereotype endorsement confirm the hypothesis, the results regarding social identity threat contradict it.

Discussion

What can be done to reduce the gender gap in participation in the German Physics Olympiad? This question was addressed in the present study. We tested a growth mindset and a values affirmation intervention as well as a combination of both interventions regarding their impact on the assumed stereotype threat and social identity threat for females in the competition. We assumed that female participants suffered from social identity threat, which was hypothesized to be expressed in higher perceived social identity threat and stereotype endorsement, as well as from stereotype threat, which was expected to be seen in higher stereotype endorsement. We, nevertheless, expected females and males to rate sense of belonging and gender identification to a similarly high degree after the interventions.

This expectation was partially fulfilled. Females did not appear to be negatively impacted due to stereotype endorsement or perceived social identity threat after the interventions: female contestants rated their perceived social identity threat higher than male contestants did at the beginning of the study, and this was still the case after the interventions, even though it had been expected that the interventions would reduce social identity threat. Therefore, we did not find any changes in assessments due to the interventions. Likewise, none of the groups showed higher impact. Females' assessment of sense of belonging, gender identification, stereotype endorsement, and perceptions of environmental stereotyping did not vary before or directly after the interventions; several weeks after the interventions, female contestants' sense of belonging was lower than that of males. This contradicts the assumption of both a social identity threat and a stereotype threat, suggesting that the interventions hindered these — all in similarly effective ways, with no one type of intervention being more advantageous than the others.

Nevertheless, it needs to be mentioned that prior to the interventions, the highly selective sample of female participants apparently did not differ from the male participants on the decisive variables: Females and males who participated in the study were similar in their assessments of the used scales Noticeably, female participants did not appear to suffer under either stereotype threat or social identity threat effects.

No stereotype threat or social identity threat?

The results of this study seem to indicate that this highly selective sample is at least not hindered or harmed by stereotype or social identity threat effects either before or after the interventions; this contradicts previous findings on stereotypes. However, we cannot say if these effects would be comparable without interventions, as we did not include a control group.

First, although we expected females to suffer under threats while participating in the predominantly male physics seminars (see, e.g., Marx et al., 2005; Good et al., 2012; Schmader et al., 2015), our results suggest that the interventions prevented this. Even though we did not find any reduction in the effects of stereotype threat after the interventions, the interventions seemed to help the females to not fall behind in sense of belonging, which could have been expected without interventions based on social identity and stereotype threat theory (see, e.g., Aronson and Inzlicht, 2004; Murphy et al., 2007). Participating in the interventions apparently prevented a split between the genders. Even though we cannot draw the conclusion that the interventions reduced any of the negative effects, we can assume that the interventions prevented the effects from appearing in the first place.

Second, there could be several different reasons for the results we found. Female participants in the German Physics Olympiad are a specific sample. They have mastered several hurdles to compete in the competition: They have resisted rejecting cues in a stereotypically male domain and have not shied away from a competition rarely won by females. Entering the competition can already be seen as an indicator of high resistance (e.g., Gonsalves et al., 2021) to negative stereotypes, male predominance, and to associations of physics with the male gender, which indicate less fit and reduce sense of belonging and cause women to leave physics (see, e.g., Cheryan and Plaut, 2010; Makarova et al., 2019). Nevertheless, several studies showed that females in science competitions suffer under stereotypes (see, e.g., Steegh et al., 2019).

Why was the sample of this study even more resilient?

This study was tied to participation in a weekend seminar. Participants were willing to spend a whole weekend in a group of interested and talented students (see Höffler et al., 2019), solving physics tasks and conducting experiments. The group thus most likely consisted of female students who had even more interest and wanted to deepen their knowledge in physics even more than other participants of the Physics Olympiad. As all 539 female participants of the two cohorts of the German Physics Olympiad could have participated in this study but only 82 did, our sample possibly was more engaged in the field and, therefore, more resistant to cues that hinder females from pursuing science.

Further, our sample might vary in motivational profiles and success expectations for proceeding to the competition's next round to the overall group of participants (see Steegh et al., 2021). The weekend seminars took place after the end of the first round but before the announcement of the participants who would continue to the next round. It could be assumed that the participants of our study chose to be in the seminars either in expected preparation for the continuing competition or just to engage further in physics (see Höffler et al., 2019). This would imply higher interest, higher learning goal orientation, and a higher self-concept (see Höffler et al., 2017) than the average participant — factors that could interact with an individual's susceptibility to stereotype or social identity threat.

Stereotype threat theory apparently does not apply to this group of participants. Thus, other factors that might account for the gender ratio and the achievement differences in this field need to be considered. Previous studies have suggested further starting points, which can be separated into two main groups with regard to their implications for science competitions.

First, possible causes might be found in the differences between the best contestants' characteristics and characteristics of females, who are highly interested but not as successful (see also the competition proceedings; e.g., Petersen and Wulff, 2017): Two questions are interesting here: First, in which characteristics or abilities are these contestants especially advanced? Differences between male and female participants that may cause gender differences and be relevant for success in the competition might regard self-concept (see, e.g., Saß and Kampa, 2019; Vinni-Laakso et al., 2019), competence (see, e.g., Schorr, 2019), or parental support (see, e.g., Hoferichter and Raufelder, 2019; Schorr, 2019). These variables have been found to closely align with gender differences and the achievement gap. Second, do personal characteristics vary between the best contestants and the female participants who drop out of the competition earlier? It appears useful to look closely at empathy (see Ghazy et al., 2019), motivation (see, e.g., Watt et al., 2019; Luttenberger et al., 2019a; Dietrich and Lazarides, 2019), and interest (see, e.g., Ertl and Hartmann, 2019; Song et al., 2019); these variables are closely related to persistence in science and to gender differences. Analyzing differences in these variables between male and female participants might give insights into the personal characteristics that hinder talented young women from succeeding to the finale.

Further research should also analyze the competition regarding success factors that are independent of the contestants' personal characteristics: Do tasks and experiments especially favor male participants (see, e.g., Sanchis-Segura et al., 2018)? Is the content of tasks (see, e.g., Wille et al., 2018; Wheeler and Blanchard, 2019) or the context of the examination process more disadvantageous for females (e.g., Sobieraj and Krämer, 2019)? Studying those factors could help make the competition more equitable for both genders.

However, it might also be a possibility that the general effect of stereotypes on females in science — in our case physics — is changing. Younger generations could perceive the world as being more equal thus suffering less under stereotypes. The fact that the rates of girls dropping out of the competition is higher than the boys', could, for example, just be out of a lack of personal importance of investing into succeeding in the competition. Which are the personal advantages that girls gain out of competing — aside from, e.g., knowledge and getting to know other interested students? Are girls and boys perceiving investing time and studying for the competition as justifiable for their personal outcome? Future research is thus not just asked to focus on the impact of changing perceptions of gender and societal issues, which are likely to impact feelings and perceptions of equality in our societies, but also on gender difference in the significance attributed to science.

Limitations

First, all participants in the German Physics Olympiad received invitations to the study. However, participation was voluntary and we cannot assume that this sample is representative of the overall sample of contestants. Rather, we assume the participants of our study to be more interested, engaged, and more likely to continue in science than the other contestants. For this highly selective sample, which is most likely to pursue a career in physics, the theory of stereotype or social identity threat apparently does not suffice as an explanation for the gender gap. Future research should apply new theories to find more useful explanatory approaches.

Second, stereotype threat was measured explicitly in this study by measuring stereotype endorsement. Participants were asked how much they agree with common derogatory stereotypes about females' physics abilities. Previous research showed higher stereotype endorsement as a predictor of higher susceptibility to stereotype threat (e.g., Schmader et al., 2004; Pennington et al., 2016). We considered our explicit method as preferential to activating stereotypes or implicitly measuring participants' agreement with stereotypes, as the weekend seminar depicted the regular environment of the competition and the measurement was combined within the questionnaires without drawing any special attention to its purpose. Nevertheless, explicitly measuring stereotype endorsement might lead to divergent results, although previous research is not congruous (see, e.g., Kessels et al., 2006). As the assessments of stereotype endorsement were equally high throughout the study, we do not believe that our results are biased. Regardless, future research should add an implicit measure to control for social desirability effects.

Lastly, our study did not use a control group. We can thus only assume that finding no differences between male and female participants' assessments after the interventions indicates beneficial effects of the interventions on stereotype and social identity threat. This seems the appropriate conclusion as previous literature rather consistently showed the existence of stereotype threat for females in science competitions. Nevertheless, future research should include a control group to measure the extent of the interventions' effects.

Conclusion

This study addressed factors that are potentially responsible for male predominance in the German Physics Olympiad. Previous research showed that stereotype and social identity threat are useful models to explain the underachievement and consequent underrepresentation of women and females in science. The results of this study, however, suggest that stereotype threat and social identity threat are possibly not applicable to the highly interested and engaged female participants of the German Physics Olympiad. From the beginning, females who chose to participate in the competition and in the study's weekend seminars were not affected more negatively by stereotypes or social identity threat than their male counterparts. It thus seems that the commonly expected harm done by stereotypes did not occur to the extent expected. Nevertheless, further pursuing to include interventions against stereotype threat in environments, which are highly likely to induce the mechanism, seems important. Shielding more young women from harmful impacts could reduce the gender gap even further.

Why then are these females still not as successful in the continuing competition as their male counterparts? The results of this study suggest that other approaches need to be tested to examine this question. We suggest looking more closely at new approaches and concepts that do not focus on stereotypes and social identity as the reasons for deciding against a career in science and, instead, focus on examining a combination of the internal and external factors behind this decision (see, e.g., Luttenberger et al., 2019b).

Overall, the results of this study provide a ray of hope for physics: If the females who are most likely to continue in science are immune to or not as affected by stereotypes as the average female student, stereotypes might not be such a big problem for the domain anymore. Although continuing to fight stereotypes might encourage more females to proceed to this level of pursuing physics, how to support those who have already reached this stage should also receive more attention. Nevertheless, this study shows that promising starting points for supporting interested females in science could be interventions that promote resilience and support the development of abilities and useful characteristics.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

AL wrote the first draft and collected the data. AL and OK contributed to the data analysis. All authors contributed to the manuscript, revised the manuscript critically, and approved it for submission.

Funding

The study was part of the project "Identiphy – Identity development in physics!," which was funded by the Bundesministerium für Bildung und Forschung (Grant number: 01 FP1717).

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alter, A. L., Aronson, J., Darley, J. M., Rodriguez, C., and Ruble, D. N. (2010). Rising to the threat: reducing stereotype threat by reframing the threat as a challenge. *J. Exp. Soc. Psychol.* 46, 166–171. doi: 10.1016/j.jesp.2009.09.014

Anger, C., Koppel, O., Plünnecke, A., Röben, E., and Schüler, R. M. (2019). *MINT - Grundlage für Forschung und Digitalisierung. Gutachten für BDA, BDI, MINT Zukunft schaffen und Gesamtmetall [STEM – Basis for research and digitalization].*

Archer, L., Moote, J., and MacLeod, E. (2020). Learning that physics is 'not for me': pedagogic work and the cultivation of habitus among advanced level physics students. *J. Learn. Sci.* 29, 347–384. doi: 10.1080/10508406.2019.1707679

Aronson, J., Fried, C. B., and Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *J. Exp. Soc. Psychol.* 38, 113–125. doi: 10.1006/jesp.2001.1491

Aronson, J., and Inzlicht, M. (2004). The ups and downs of attributional ambiguity: stereotype vulnerability and the academic self-knowledge of African American college students. *Psychol. Sci.* 15, 829–836. doi: 10.1111/j.0956-7976.2004.00763.x

Bailey, K. A., Horacek, D., Worthington, S., Nanthakumar, A., Preston, S., and Ilie, C. C. (2019). STEM/non-STEM divide structures undergraduate beliefs about gender and talent in academia. *Front. Sociol.* 4:26. doi: 10.3389/fsoc.2019.00026

Banchefsky, S., Lewis, K. L., and Ito, T. A. (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

Bayly, B. L., and Bumpus, M. F. (2019). An exploration of engagement and effectiveness of an online values affirmation. *Educ. Res. Eval.* 25, 248–269. doi: 10.1080/13803611.2020.1717542

Bedyńska, S., Krejtz, I., and Sedek, G. (2018). Chronic stereotype threat is associated with mathematical achievement on representative sample of secondary schoolgirls: the role of gender identification, working memory, and intellectual helplessness. *Front. Psychol.* 9:428. doi: 10.3389/fpsyg.2018.00428

Ben-Zeev, T., Fein, S., and Inzlicht, M. (2005). Arousal and stereotype threat. *J. Exp. Soc. Psychol.* 41, 174–181. doi: 10.1016/j.jesp.2003.11.007

Blackwell, L. S., Trzesniewski, K. H., and Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child Dev.* 78, 246–263. doi: 10.1111/j.1467-8624.2007.00995.x

Brady, S. T., Cohen, G. L., Jarvis, S. N., and Walton, G. M. (2020). A brief social-belonging intervention in college improves adult outcomes for black Americans. *Sci. Adv.* 6:eaay3689. doi: 10.1126/sciadv.aay3689

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., and Finkel, E. J. (2013). Mind-sets matter: a meta-analytic review of implicit theories and self-regulation. *Psychol. Bull.* 139, 655–701. doi: 10.1037/a0029531

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., Siy, J. O., Vichayapai, M., Drury, B. J., and Kim, S. (2011). Do female and male role models who embody STEM stereotyped hinder women's anticipated success in STEM? *Soc. Psychol. Personal. Sci.* 2, 656–664. doi: 10.1177/1948550611405218

Cohen, G. L., Garcia, J., Apfel, N., and Master, A. (2006). Reducing the racial achievement gap: a social-psychological intervention. *Science* 313, 1307–1310. doi: 10.1126/science.1128317

Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., and Brzustoski, P. (2009). Recursive processes in self-affirmation: intervening to close the minority achievement gap. *Science* 324, 400–403. doi: 10.1126/science.1170769

de Jong, E. M., Jellesma, F. C., Koomen, H. M. Y., and de Jong, P. F. (2016). A values- affirmation intervention does not benefit negatively stereotyped immigrant students in the Netherlands. *Front. Psychol.* 7:691. doi: 10.3389/fpsyg.2016.00691

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

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

Düchs, G., and Mecke, K. (2019). Vielfalt statt Einfalt. Statistiken zum Physikstudium an den Universitäten in Deutschland 2019 [Statistics to physics courses at German Universities in 2019]. *Phys. J.* 18, 26–31.

Dweck, C.S. (2000). *Self-Theories: Their Role in Motivation, Personality and Development.* Taylor & Francis Philadelphia, PA.

Ertl, B., and Hartmann, F. G. (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

Estrada, M., Woodcock, A., Hernandez, P. R., and Schultz, P. W. (2011). Toward a model of social influence that explains minority student integration into the scientific community. *J. Educ. Psychol.* 103, 206–222. doi: 10.1037/a0020743

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fennema, E., and Sherman, J. A. (1976). Fennema-Sherman Mathematics Attitude Scales: instruments designed to measure attitudes toward the learning of mathematics by females and males. *J. Res. Math. Educ.* 7, 324–326. doi: 10.2307/748467

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

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

Froehlich, L., Martiny, S. E., Deaux, K., Goetz, T., and Mok, S. Y. (2016). Being smart or getting smarter: implicit theory of intelligence moderates stereotype threat and stereotype lift effects. *Br. J. Educat. Psychol.* 55, 564–587. doi: 10.1111/bjso.12144

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

Gillen-O'Neel, C., and Fuligni, A. (2013). A longitudinal study of school belonging and academic motivation across high school. *Child Dev.* 84, 678–692. doi: 10.1111/j.1467-8624.2012.01862.x

Gonsalves, A. J., Cavalcante, A. S., Sprowls, E. D., and Iacono, H. (2021). "Anybody can do science if they're brave enough": understanding the role of science capital in science majors' identity trajectories into and through postsecondary science. *J. Res. Sci. Teach.* 58, 1117–1151. doi: 10.1002/tea.21695

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

Hall, W. M., Schmader, T., Aday, A., and Croft, E. (2018). Decoding the dynamics of social identity threat in the workplace: a within-person analysis of women's and men's interactions in STEM. *Soc. Psychol. Personal. Sci.* 10, 542–552. doi: 10.1177/1948550618772582

Hall, W. M., Schmader, T., and Croft, E. (2015). Engineering exchanges: daily social identity threat predicts burnout among female engineers. *Soc. Psychol. Personal. Sci.* 6, 528–534. doi: 10.1177/1948550615572637

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

Hartley, B. L., and Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Dev.* 84, 1716–1733. doi: 10.1111/cdev.12079

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. Education* 4:14. doi: 10.3389/feduc.2019.00014

Höffler, T., Bonin, V., and Parchmann, I. (2017). Science vs. sports: motivation and self- concepts of participants in different school competitions. *Int. J. Sci. Math. Educ.* 15, 817–836. doi: 10.1007/s10763-016-9717-y

Höffler, T., Köhler, C., and Parchmann, I. (2019). Scientists of the future: an analysis of talented students' interests. *Int. J. STEM Educat.* 6, 1–8. doi: 10.1186/s40594-019-0184-1

Huguet, P., and Régner, I. (2009). Counter-stereotypic beliefs in math do not protect school girls from stereotype threat. *J. Exp. Soc. Psychol.* 45, 1024–1027. doi: 10.1016/j.jesp.2009.04.029

IBM Corp. (2015). *IBM SPSS Statistics for Windows, Version 23.0*. Armonk, NY: IBM Corp.

Jamieson, J. P., and Harkins, S. G. (2011). The intervening task method: implications for measuring mediation. *Personal. Soc. Psychol. Bull.* 37, 652–661. doi: 10.1177/0146167211399776

Johansson, A. (2020). Negotiation intelligence, nerdiness, and status in physics master's studies. *Res. Sci. Educ.* 50, 2419–2440. doi: 10.1007/s11165-018-9786-8

Kahn, S., and Ginther, D. K. (2015). Are recent cohorts of women with engineering bachelors less likely to stay in engineering? *Front. Psychol.* 6:1144. doi: 10.3389/fpsyg.2015.01144

Kalokerinos, E. K., Kjelsaas, K., Bennetts, S., and von Hippel, C. (2017). Men in pink collars: stereotype threat and disengagement among male teachers and child protection workers. *Eur. J. Soc. Psychol.* 47, 553–565. doi: 10.1002/ejsp.2246

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

Ladewig, A., Keller, M., and Klusmann, U. (2020). Sense of Belonging as an important factor in the pursuit of physics: Does it also matter for female participants of the German Physics Olympiad? *Front. Psychol.* 11:548781. doi: 10.3389/fpsyg.2020.548781

Ladewig, A., Köller, O., and Neumann, K. (2022). Persisting in physics and the physics olympiad — impact of gender identification and sense of belonging on expectancy-value outcomes *Europ. J. Psychol. Educ.* doi: 10.1007/s10212-022-00600-5

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

Luhtanen, R., and Crocker, J. (1992). A collective self-esteem scale: self-evaluation of one's social identity. *Personal. Soc. Psychol. Bull.* 18, 302–318. doi: 10.1177/0146167292183006

Luttenberger, S., Paechter, M., and Ertl, B. (2019a). 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

Luttenberger, S., Steinlechner, P., Ertl, B., and Paechter, M. (2019b). 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

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. Frontiers. *Education* 4:60. doi: 10.3389/feduc.2019.00060

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

Marx, D. M., Stapel, D. A., and Muller, D. (2005). We can do it: the interplay of construal orientation and social comparisons under threat. *J. Pers. Soc. Psychol.* 88, 432–446. doi: 10.1037/0022-3514.88.3.432

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., and Wai, J. (2015). The bachelor's to Ph.D. STEM pipeline no longer leaks more women than men: a 30-year analysis. *Front. Psychol.* 6:37. doi: 10.3389/fpsyg.2015.00037

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., and Ito, T. A. (2010). Reducing the gender achievement gap in college science: a classroom study of values affirmation. *Science* 330, 1234–1237. doi: 10.1126/science.1195996

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

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

Petersen, S., and Wulff, P. (2017). The German physics Olympiad – identifying and inspiring talents. *Eur. J. Phys.* 38, 1–16. doi: 10.1088/1361-6404/aa538f

Pittman, L. D., and Richmond, A. (2007). Academic and psychological functioning in late adolescence: the importance of school belonging. *J. Exp. Educ.* 75, 270–290. doi: 10.3200/JEXE.75.4.270-292

Rahn, G., Martiny, S. E., and Nikitin, J. (2020). Feeling out of place: internalized age stereotypes are associated with older employees' sense of belonging and social motivation. *Work Aging Retire.* 7, 61–77. doi: 10.1093/workar/waaa005

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

Sadler, P., 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

Sanchis-Segura, C., Aguirre, N., Cruz-Gómez, Á. J., 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

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

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., Hall, W., and Croft, A. (2015). "Stereotype threat in intergroup relations" in *APA Handbooks in Psychology®. APA Handbook of Personality and*

Social Psychology, vol. 2. Group Processes. eds. M. Mikulincer, P. R. Shaver, J. F. Dovidio and J. A. Simpson (American Psychological Association), 447–471.

Schmader, T., and Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *J. Pers. Soc. Psychol.* 85, 440–452. doi: 10.1037/0022-3514.85.3.440

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

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

Setterlund, M. B., and Niedenthal, P. M. (1993). "Who am I? Why am I here?" Self-esteem, self-clarity, and prototype matching. *J. Pers. Soc. Psychol.* 65, 769–780. doi: 10.1037/0022-3514.65.4.769

Shih, M., Pittinsky, T. L., and Ambady, N. (1999). Stereotype susceptibility: identity salience and shifts in quantitative performance. *Psychol. Sci.* 10, 80–83. doi: 10.1111/1467-9280.00111

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

Sobieraj, S., and Krämer, N. C. (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

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

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

Steegh, A. M., Höffler, T. N., Höft, L., and Parchmann, I. (2020). First steps toward gender equity in the chemistry Olympiad: understanding the role of implicit gender-science stereotypes. *J. Res. Sci. Teach.* 58, 40–68. doi: 10.1002/tea.21645

Steegh, A. M., Höffler, T. N., Höft, L., and Parchmann, I. (2021). Exploring science competition participants' expectancy-value perceptions and identification: a latent profile analysis. *Learn. Instr.* 74:101455. doi: 10.1016/j.learninstruc.2021.101455

Steegh, A. M., Höffler, T. N., Keller, M. M., and Parchmann, I. (2019). Gender differences in mathematics and science competitions: a systematic review. *J. Res. Sci. Teach.* 56, 1431–1460. doi: 10.1002/tea.21580

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

Steele, C. M., Spencer, S. J., and Aronson, J. (2002). Contending with group image: the psychology of stereotype and social identity threat. *Adv. Exp. Soc. Psychol.* 34, 379–440. doi: 10.1016/s0065-2601(02)80009-0

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

van Veelen, R., Derks, B., and Endedijk, M. D. (2019). Double trouble: how being outnumbered and negatively stereotyped threatens career outcomes of women in STEM. *Front. Psychol.* 10:150. doi: 10.3389/fpsyg.2019.00150

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

Vohs, K. D., Park, J. K., and Schmeichel, B. J. (2013). Self-affirmation can enable goal disengagement. *J. Pers. Soc. Psychol.* 104, 14–27. doi: 10.1037/a0030478

Walton, G. M., and Cohen, G. L. (2007). A question of belonging: race, social fit, and achievement. *J. Personal. Soc. Psychol. Res.* 92, 82–96. doi: 10.1037/0022-3514.92.1.82

Watt, H. M. G., 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

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. doi: 10.3389/fpsyg.2019.00594

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

Yeager, D. S., Walton, G. M., Brady, S. T., Akcinar, E. N., Paunesku, D., Keane, L., et al. (2016). Teaching a lay theory before college narrows achievement gaps at scale. *Proc. Natl. Acad. Sci. U. S. A.* 113, E3341–E3348. doi: 10.1073/pnas.1524360113

Yeager, D. S., Walton, G. M., and Cohen, G. L. (2013). Addressing achievement gaps with psychological interventions. *Phi Delta Kappan* 94, 62–65. doi: 10.1177%2F003172171309400514

TYPE Original Research
PUBLISHED 13 March 2023
DOI 10.3389/feduc.2023.1018241

OPEN ACCESS

EDITED BY
Lisbet Rønningsbakk,
UiT The Arctic University of Norway, Norway

REVIEWED BY
Silvina Ponce Dawson,
University of Buenos Aires, Argentina
Susan Howitt,
Australian National University, Australia

*CORRESPONDENCE
Matthew C. Fleenor
✉ mfleenor@umw.edu

SPECIALTY SECTION
This article was submitted to STEM Education, a section of the journal Frontiers in Education

RECEIVED 12 August 2022
ACCEPTED 23 February 2023
PUBLISHED 13 March 2023

CITATION
Balasubramanian R, Findley-Van Nostrand D and Fleenor MC (2023) Programmatic innovations that accord with the retention of women in STEM careers.
Front. Educ. 8:1018241.
doi: 10.3389/feduc.2023.1018241

Programmatic innovations that accord with the retention of women in STEM careers

Rama Balasubramanian[1,2], Danielle Findley-Van Nostrand[3] and Matthew C. Fleenor[1,4]*

[1]Department of Mathematics, Computer Science, and Physics, Roanoke College, Salem, VA, United States, [2]Chance to Change Lives (CCL-US), Pittsburgh, PA, United States, [3]Department of Psychology, Roanoke College, Salem, VA, United States, [4]Department of Chemistry and Physics, University of Mary Washington, Fredericksburg, VA, United States

Gender representation in the physical sciences remains inequitable and continues to lag behind other fields. Even though there exists adequate documentation regarding programmatic postures, difficulties persist within the physics discipline. In this paper, we present innovative, programmatic elements over an 8-year period at an undergraduate, liberal arts, physics program. These elements were added in response to the following two questions: "What practices cultivate an increase of physics major numbers in an undergraduate, liberal arts setting?" and "What practices facilitate a depth of experience for individual physics graduates?" Some of these innovations aligned with published, "best practices" for undergraduate physics programs, while others were novel to the program's context. Within this 8-year period, alterations were separated into curricular and co-curricular elements. Innovations are described, and data are presented in 3-year timeframes before, during, and after their implementation. The number of total majors and graduates increased, including a 200% increase of women degree recipients compared to the previous 10 years. This boosted average graduation rates for women above the national average (30% > 20%). Moreover, women were retained within the undergraduate physics major at a higher percentage during this time period when compared to men in the program. Lastly, these women physics majors maintained careers in science advancement fields at a rate of 80±% after ≤ 5 years post-graduation. While this paper presents a singular case study, the purpose is two-fold: (a) to validate quantitatively the work of national physics organizations within the context of a liberal arts institution, and (b) to suggest that a multi-level approach is most efficacious when considering programmatic innovations.

KEYWORDS
physics undergraduate education, experiential learning, women retention, STEM identity, STEM belonging

1. Introduction

1.1. Global landscape

One of 17 United Nations Sustainable Development Goals reads, "achieve gender equality and empower all women and girls" (United Nations, 2023). Within purportedly more rational and logical disciplines in Science, Technology, Engineering, and Mathematics (STEM), one might expect a better fulfillment of this goal. Yet, STEM disciplines reveal a glaring "gender gap" at the professional level, despite the increase of women undertaking higher-education STEM studies. Gender gaps express themselves as differences "between women and men in terms of their levels of participation, access, rights, remuneration, or benefits" (World Economic Forum, 2019). Education is one of the four key areas of identification and measurement. Within the sciences, fewer than 30% of the world's researchers are women, which reflects a clear gender gap (UNESCO Institute for Statistics, 2019). Through a conglomeration of global scientific organizations, individual survey results from global scientists, and publication pattern information, a deep synthesis by Roy et al. (2020) reveals a persistent and significant gender gap within the sciences globally. Sampling equally over 30,000 men and women from 159 countries, this gap remains regardless of STEM discipline, geographic location, or economic development. While what follows is a focused examination of one physics program at an undergraduate institution in the United States (US), the local environment of the study and the national context of the institution are representative of the global landscape.

1.2. National landscape

Over the past 30 years in the US, gender representation in the physical sciences continues to be disproportionately skewed against women (Ivie and Tesfaye, 2012), despite a three-fold increase in the percentage of undergraduate physics degrees awarded to women (6% in 1970 vs. 22% in 2018, Porter and Ivie, 2019). According to a 2021 report by the American Physical Society (APS) on building America's workforce, the importance of developing an inclusive and diverse workforce is crucial to boost the innovation and productivity of science and technology and to maintain America as a global leader in these areas (Johnson, 2021). Problems persist in gender equity issues within most STEM fields (e.g., Hill et al., 2010), though physics remains one of the most inequitable. At approximately 20% women recipients of undergraduate degrees, physics continues to lag behind mathematics (40%), chemistry (~50%), and biology (60%). This considerable difference is displayed clearly in Porter and Ivie (2019), where the percentage of women physics-degree recipients remains approximately constant since 2,000. Within this same climate, the number of undergraduate-degree recipients is more balanced. For example, the National Center for Education Statistics reports that women receive 58% of all undergraduate degrees, while only 36% of those are awarded to STEM majors (US Department of Education, National Center for Education Statistics, 2019).

There are various potential reasons for this continued gender disparity in the physical sciences, including stereotypes relating the practice of physical science to hyper-masculinity (Good et al., 2008; Smyth and Nosek, 2015; Francis et al., 2017), or a lack of perceived representation of women or women role models in the field, in turn leading to issues of belonging and fit (Nelson and Brammer, 2010; Good et al., 2012; Gisler et al., 2018). There are also individual-level factors to consider, including women's lower self-efficacy for physics (not coupled with lower objective skills, Kalender et al., 2020), their experienced science identity (Trujillo and Tanner, 2014; Eren, 2021), or a perceived lack of opportunity to fulfill communal career goals, which tend to be valued more by women (Diekman et al., 2017).

Simultaneous to persistent gender inequity for physics undergraduate recipients, US organizations like the Statistical Research Center of the American Institute of Physics (AIP) address and publicize these known inequities through detailed study, survey, and site visitation (AIP, 2021). The AIP continues to produce valuable statistics on historically underrepresented populations in physics specifically, and STEM fields as a whole. This important work presents the results of labor within the field, as well as the persistent gaps and needed future focus. In addition, commissions of specialists within the discipline have produced and published manuals containing guidelines on undergraduate program efficacy. This is particularly true for undergraduate programs as a whole (SPIN-UP, Hilborn et al., 2003), career preparation (J-TUPP, Heron and McNeil, 2016), and African-American undergraduates within physics (TEAM-UP, James and Bertschinger, 2020). Resources related to programmatic guidance, change, and growth are available through the interactive AAPT-EP3 website, supported by the American Association of Physics Teachers (AAPT/APS, 2021). It is essential that physics programs consult and digest the results of these resources and apply the general guidelines to their specific institutional contexts.

Because the current study focuses on an undergraduate physics program in the US, some differences between traditional American and non-US undergraduate universities are made clear. Most popularly, it is well-known that an undergraduate degree in the US takes approximately 1-year longer than systems outside of the US, particularly European-based systems. Part of this prolonging by the US system accords with an emphasis on general education coursework, which usually delays the declaration of an intended undergraduate degree program, usually referred to as a "major." Moreover, student choice of an undergraduate major is a novelty within a US-based, post-secondary education, and it is a source of intense study. For example, how students choose an undergraduate major is a recent, scholarly focus for economics (Wiswall, 2021), race and gender studies (Rainey et al., 2018), and social psychology (Denice, 2021). A major declaration is further complicated in a liberal arts environment, which is discussed below in Section 1.3. In contrast, choosing a major is not a feature of many/most non-US systems. Whether in favor of an Dual System, like Germany (e.g., Nash, 2012), or a traditional 3-year framework, undergraduate degree pursuits outside the US are more streamlined.

Within the context of total undergraduate degrees awarded in the US, physics programs maintain a meager percentage of STEM bachelor's recipients. Approximately 20% of all undergraduate degrees are awarded to STEM fields, while approximately 2.2% of those degrees were awarded to physics over the past 20 years (APS, 2020). Even though Ph.D. granting institutions comprise only one-quarter of all institutions offering a physics bachelor's degree, they award approximately one-half of all physics bachelor's degrees. When these national percentages are applied to the reduced numbers of undergraduate populations at liberal arts colleges, where a STEM culture does not often exist, the corresponding enrollments for STEM courses and physics majors are drastically reduced.

While demand for STEM-degree recipients increases, it may appear to the public that the numbers of physics graduates are insignificant compared to the whole. Even though a small percentage of bachelor's degree recipients will pursue graduate studies in physics, the importance of a physics degree persists through the application of the discipline's transferable skills in other fields (Hunt, 2013). "Hidden physicists" (that is, "those who are trained in physics but actually work in a job more widely," Heron and McNeil, 2016) populate not only other STEM fields and education landscapes, but also peripheral fields like law and management. Physics degree recipients historically score among the highest on standardized, pre-professional entrance exams like the Medical College and Law School Admission Tests (MCAT and LSAT, Tesfaye and Mulvey, 2013; Tyler and Mulvey, 2022). Furthermore, physics degree recipients maintain a privileged position in terms of earning opportunities and employment satisfaction (cf., Figure 3 in Heron and McNeil, 2016). For these reasons, along with the continued need for STEM innovation across various career paths, it is vital that physics programs continue to produce degree recipients for careers as "hidden physicists." This is especially true at institutions where typically less than half of the bachelor's recipients matriculate to graduate physics programs.

1.3. Local landscape

The current investigation is situated within a context of small number statistics, both nationally within the discipline of physics and locally within an undergraduate, liberal arts college. The physics program discussed here (referred to as the Physics Group) is part of a shared department with mathematics and computer science. Bachelor's degree-granting, physics programs within shared departments comprise less than 10% of all US programs (Mulvey, 2021). Moreover, while offering approximately one-half of the bachelor's degrees, non-Ph.D. granting programs comprise three-quarters of the undergraduate physics landscape. Therefore, a fertile opportunity exists for these programs to affect STEM students with physics content and potential future trajectories.

Beyond the broad differences of matriculation length and focus of study outlined in Section 1.2, a liberal arts setting adds another layer of nuance to the US undergraduate system. Liberal arts colleges in the US typically provide an undergraduate-focused education emphasizing a general curriculum balanced with humanities and social sciences content, which is required for all students. Both public, state-funded and selective, private-funded institutions are labeled "liberal arts colleges." In addition, most liberal arts institutions have enrollments of less than 5,000 students. Regardless of the differences, the selection of a major is not rushed for undergraduates at these institutions. It is not uncommon to declare a major in the second year of undergraduate studies at a liberal arts institution, where becoming a (physics) major is a celebrated event.

Many non-Ph.D. granting physics programs contain less than five full-time faculty and produce less than 10 graduates per year. For instance, Tyler et al. (2020) present statistics demonstrating that a majority of undergraduate-only programs have five or less graduates per year *and* five or less full-time equivalent faculty. The Physics Group also finds itself in a similar situation, where it has three tenure-track (TT) members, a non-permanent visiting position (VAP), and a permanent, non-tenure track (NTT) position. This was also the case 10–15 years ago when the Physics Group contained four TT and one NTT faculty, while it awarded 3.6 degrees per year. The number of women Physics Group faculty during the time-period of the study was either two (1 TT, 1 NTT), or three (1 TT, 1 NTT, 1 VAP) of five. While the NTT and visiting faculty members primarily taught physics courses for non-majors, they actively contributed to additional learning experiences for women students such as training teaching assistants, troubleshooting lab equipment, facilitating travel to CUWiP conferences. It is not the focus of the current study to evaluate the effects of faculty gender distribution. The average percentage of degrees awarded to women during the same time period (2003–2012) was approximately 20%, with the national average. When considering published national statistics, the program in this study could be classified as a typical, undergraduate physics program in the US, prior to the study focus (2013–2021).

The current study maintains the following trajectory: innovative, programmatic elements are introduced and described in the next section. Next, we provide representation and retention of physics degree recipients before, during, and after the implementation of the programmatic elements. While causality is not the purpose of the presentation, it will be clear that the increases for the program accord with the implementation period. A discussion of the impact on the Physics Group in light of these increases follows with some suggestions for other similar programs. While each institutional context is unique, and the synergistic effect of innovative elements cannot be deduced *a priori*, the experience of the Physics Group substantiates two demonstrative realities: published guidelines and statistics support programmatic growth goals, and multi-compartment implementation provides an effective programmatic impact.

2. Materials and methods

2.1. Data acquisition

The Office of Institutional Research and the Office of the Registrar were queried for the numbers of majors and graduates for the previous 20 years so that a baseline of physics participation was determined. For the 8 years of interest within the study (graduates from 2013 to 2021), more detailed information was acquired and collated, including course rosters, extra-curricular participation, and post-graduation employment. This information was accessed through interpersonal communication, social media, and programmatic assessment. The figures within the study were compiled from the results of these data.

The timetable for programmatic alteration was carried out with the following provisos: only one element was implemented in a given year, innovative elements were connected directly to items within the annual programmatic assessment plan, and these elements were carried out for a minimum of 3 years before conducting evaluation. This 3-year baseline for introduction and implementation of programmatic elements forms a justification for the presentation of data in subsequent sections. Innovative elements were implemented locally based on consultation of the Physics Education Research (PER) literature as well as familiarity with the program, as detailed in the following sections. The addition of innovative elements for programmatic growth was driven equally by the following two questions: "What practices cultivate an increase of physics major numbers in an undergraduate, liberal arts setting?"; and, "What practices facilitate a depth of experience for individual physics graduates?"

Responsibilities for implementing innovative elements were shared among the Physics Group faculty. Within courses, the instructor of record maintained primary leadership over the specific implementation and data collection. For intra-programmatic or extra-curricular elements, a particular physics faculty member oversaw implementation and collection. While literature-based elements were implemented as accurately as possible, concessions for context and student populations were judiciously applied. In what follows, physics program alterations are categorized and discussed within the contexts of undergraduate education literature and the particular institutional environment.

2.2. Introduction of innovative elements

Programmatic transformation within STEM undergraduate programs is not clear-cut and often dependent on localized factors. Despite public usage of the STEM acronym, it is clear that even the "S" (science) is not monolithic in its practice (Marder, 2013). Moreover, when considering programmatic change initiatives, points of emphasis vary based on discipline (Reinholz et al., 2019). These differences stand outside the unique departmental and institutional cultures on each campus (Henderson et al., 2015). Detailed case studies of undergraduate programmatic growth exist in the literature for large, US institutions (Stewart et al., 2013) and diverse "thriving" programs (Hilborn et al., 2003). The "best practices" documentation contains some, but not all, of the innovations that are introduced here.

Consideration of significant programmatic change began in 2011 while program faculty were preparing for an external review. This physics program review occurred in 2012, during which some of the proposed elements were discussed with evaluators. Accordingly, a focus on growth pivoted from recruitment to retention, because program faculty overlooked that retention could double the number of physics graduates. While some changes were contextual and on-going, major elements were implemented over a 4-year period, 2013–2017, along two distinct compartments of the program (curricular and co-curricular). With a 4-year undergraduate matriculation period, eight academic years within the Physics Group provide the basis for the present study (i.e., 2013–2021).

2.2.1. Curricular

While faculty may have significant autonomy within the classroom or individual course offerings, instituting change within a curriculum is not an individual escapade. The process requires input and collaboration from all the physics faculty, where consensus is the aspirational goal. Moreover, the students themselves must also display a receptivity to any modifications and justifications that are offered. Here, we present details regarding two major curricular changes.

(A) First-year colloquium. Belonging and science identity are interrelated, drive retention in STEM majors, and may be especially important for women and minority students (Good et al., 2012; Rainey et al., 2018). Outside of a traditional curriculum, these qualities can be facilitated by affording students early and ample opportunities to connect with one another and with faculty members. The first-year, Physics and Engineering Colloquium meets weekly as an exploratory course emphasizing overarching themes in the physical sciences. While maintaining a high relational component for cohort-building, the half-credit course is graded on a "pass-fail" basis and is based on participation, completion of assignments, and written reflection quality. Throughout the semester, the first-year students are also introduced to several different cohorts within the physics major (upper-level students and faculty), while also engaging with a breadth of generalized content (order-of-magnitude estimates, physical modeling, and "how things work"). Hands-on investigations supplement classroom sessions in order to emphasize the experimental aspects of the discipline, as well as increase self-efficacy for tasks needed in subsequent physics courses. In summary, a successful colloquium experience cultivates the following: social capital (cohort-building and inclusion, Abbott and Sapsford, 2005), content engagement in the discipline (identity), and active learning in the discipline (self-efficacy). By instituting a first-semester course where students of similar interests gather, a like-minded cohort of learners is formed within an inclusive environment, which is supported as a means of establishing a STEM identity and sense of belonging (Lewis et al., 2017).

(B) Upper-level laboratories. The Physics Group also made alterations to increase active and applied learning opportunities, as well as facilitate essential experimental skill development. Due to students' interests in applied physics and engineering, two laboratory augmentations were made, while adhering to the college-wide constraints for number of major-only credits. First, the previous, junior-level, one-semester "advanced laboratory" course, which consisted of verifying physical constants, was converted into an intermediate laboratory (The Advanced Laboratory Physics Association, ALPhA, 2021). This laboratory accompanied a third-semester Modern Physics course that introduced the following novel facets: more developed experimentation and report writing, deepened uncertainty quantification, and emphasized historical and philosophical aspects of science.

With space created at the junior-level, a second laboratory course was added to effectively and efficiently address interdisciplinary topics within the physics major. Following guidance from the literature on Course-based Undergraduate Research Experiences (CUREs, Auchincloss et al., 2014; Wooten et al., 2018), authentic perspective was provided on the specific techniques, while also instructing the students on content that they would not otherwise receive (Mordacq et al., 2017). Students collaboratively completed four mini-research projects in the following areas: Astrophysics, Biophysics, Materials, and Optics/Spectroscopy. These four areas coincided with four, rotating, upper-level electives. As a result, students were introduced to all interdisciplinary electives through the CURE at a cursory level, even though they will not take all the courses in their entirety.

2.2.2. Co-curricular

Guidance from US-based organizations on undergraduate physics education emphasizes several co-curricular strategies to facilitate student and program success (Hilborn et al., 2003; AAPT/APS, 2021). All published studies mention a vibrant Society of Physics Students (SPS) chapter, or similar institution-based student cohort. With very little previous involvement, the Physics Group became more active with SPS in 2013, though its first year-end Chapter Report was not submitted until 2015. Here, five co-curricular, programmatic innovations are introduced that extend beyond a thriving SPS chapter.

(A) Public science outreach (informal programming). Cohort-building is a significant component of the program already mentioned (e.g., the first-year physics colloquium). Cohort-building implies an individual belonging and inclusion that traditional usage of "community" does not (Gowar, 2013). Another way to build interest-based inclusion is through student groups and science outreach to the public (Hinko et al., 2016). Science outreach opportunities (or "informal programs") not only serve public scientific literacy by raising awareness at an early age, but informal programs also empower undergraduates (Rethman et al., 2021). When the Physics Group began a concerted informal program effort 10 years ago, most of the events were faculty-organized and led. Within responsible and eager undergraduate leaders, informal programs transformed into a student-led effort. One such example was the total solar eclipse of 2017, where students served as Eclipse Ambassadors, which resulted in a nationally recognized award (Blake Lilly Prize). After conducting a well-organized event that served over 500 citizens, physics majors presented their experiences to their peers after their return. This further resulted in local news articles about their ambassadorship. Experiences such as these provide demonstrable opportunities for increased inclusion, efficacy, and identity within physics.

(B) Junior review. A second, related co-curricular addition to the program is Junior Review, an informal interview involving at least two faculty members and the individual physics major. This addition to the program is beneficial for multiple reasons. First, having multiple faculty members in attendance allows students to participate in collegiality and camaraderie first-hand. This approach also fosters belonging within an inclusive learning environment, which seems particularly meaningful for women (Lewis et al., 2017). Secondly, informal questions encourage each student to verbalize the ways and directions in which their interests may have changed (e.g., "In what ways has your interest in physics increased and/or decreased?"). Such self-reflection contributes positively to learning and achievement, and may help develop students' sense of meaning or purpose within physics (Fleenor, 2018). Instances of "hidden physicist" trajectories often arose within Junior Review conversations, which encourage new avenues of exploration are not hindered by presumptive assumptions (Alon, 2009). Third, the review is also an opportunity to facilitate participation in "high-impact practices" tied to deep learning, including research mentored by faculty, supportive minors and/or concentrations, and off-campus internships (Heron and McNeil, 2016).

(C) Conference attendance. Prior to 2012, student conference attendance within the Physics group was primarily synchronized with the mentoring faculty researcher. This was sparse, totaling less than five instances in 10 years. There were many contributing factors, including a lack of faculty attendance, lack of results, and lack of funding. Beginning in 2012, students attended conferences where undergraduate participation was encouraged regardless of faculty presence (e.g., regional opportunities). The institution developed on-campus poster sessions where students could present their work in less-threatening environments. With the establishment of a campus-wide Director of Undergraduate Research, monetary funding opportunities for students increased. These college-wide initiatives led to increased numbers of physics majors attending conferences.

(D) Definitions of physics excellence. Another co-curricular initiative related to the number and definition of year-end physics recognition, for (not-yet) majors. Traditionally, the institution sponsored one "Senior Scholar" award for the highest academic grades. To incorporate a holistic picture of excellence reflecting more than academic achievement, and to facilitate identity and belonging within the discipline, several new awards were added. For example, year-end recognition for majors was given for research within and service to the Physics Group. First-year awards were given for early achievement in the discipline to those considering a major in physics. These emphases properly reminded students that grades (marks) do not solely determine their undergraduate success, their inclusion within the discipline, nor their future trajectory as a physicist.

(E) Experiential learning. Traditional extracurricular research and internship opportunities are widely recognized as best-practices for cultivating STEM identity and belonging, particularly within traditionally underrepresented STEM populations (Estrada et al., 2018). Due to the limited number of research projects within the Physics Group, faculty pursued creative avenues for physics-related extracurricular experiential learning (EEL). Beyond more common, widely-publicized Research Experiences for Undergraduates (REUs), EEL opportunities for majors within the Physics Group were initiated with regional industry corporations, regional and on-campus collaborators, and Physics Group alumni. An introductory independent study course was also created to better prepare students for their (predominantly) summer EEL participation. The pre-emptive courses gauged student interest, facilitated research prowess, and built resilience. By not making a research requirement within the curriculum, the Physics Group invites a student to discover for themselves how best to uniquely experience physics. These opportunities not only increase the total number of students who participate in EEL, but they also broaden and diversify how a physicist is defined in society.

3. Results

Since specific changes and their effects are not isolated, program data are presented before, during, and after implementation. This precedent follows best-practices by recognizing departmental culture and multiple change agents (Dunne and Zandstra, 2011; Reinholz et al., 2019). Where results pertain to primarily one element, it is recognized that other elements are also "running in the background." Entanglement between innovative elements is discussed after the presentation of the increases in programmatic markers.

Two features of the results deserve clarification. There is a preference to average (or sum) over 3-year increments in the data presented. There are three justifications for this approach. One, the innovations were staggered and repeated for a 3-year timescale before evaluating their effectiveness. Curricular implementations did not initiate in the same academic year, so the 3-year average provides an opportunity to see the partial development of one innovation and its integration within the program more holistically. Second, a 3-year timescale defines the active trajectory for an undergraduate physics major. Upper-level core courses are taught on every-other-year basis, which also serves to compress the third and fourth year of the program. Moreover, by a student's fourth year, much of what they do serves as preparation for post-graduate decisions. While a physics major may undergo some significant intellectual transformation within their fourth year, this is an exception not the rule. Third, a 3-year timescale provides an opportunity to discuss numbers of physics majors (and graduates) in more meaningful quantities. Since the issue of small-number statistics forms the basis for a second clarification, we now transition to that discussion.

Historically, the number of people pursuing physics degrees has always been small in comparison to the whole of undergraduate degrees awarded. Compounded with that fact is the total enrollment at the liberal arts college where the study was conducted ($\sim 2,000$). Since Poisson distribution statistics characterize samples that are collected at random but with a definite average rate, we believe that they accurately describe the occasion of a woman choosing and matriculating through an undergraduate physics program. The definite average rate is set by the typical 4-year matriculation period through the US undergraduate system, where the randomness is based on the gender of a student and their choice of a major. The uncertainties for Poisson samples are provided by $\sqrt{N}$, where N is the number of events. Such uncertainties are valid whether cumulative or average populations are examined, since the standard deviation for a Poisson distribution is also given as $\sqrt{\mu}$, where μ is the mean count (Taylor, 1997). Based on the programmatic justification for 3-year increments given above, Poisson uncertainties measure the benefit of innovation implementation above a typical statistical noise.

The use of 3-year increments and Poissonian uncertainties are utilized throughout the results of the study. Therefore, a description is provided for the general presentation of the results section. In all histograms outlining the numbers of students for time-periods before and after the study, the results are binned according to the same years. A marker denotes the initiation of the study. When error bars are provided, they are calculated by the square-root of the number, whether cumulative or average numbers are presented. When other presentations of data are utilized, they are explained in context.

3.1. Increased number of physics majors and graduates

During the time period of implementation (2012–2021), the Physics Group did not have any external changes regarding number of faculty, departmental situation, or physical location. Over the 20-year period presented here, there was some faculty turnover though the number of positions remained constant. With respect to national standards, the Physics Group is considered "normal" regarding the number of physics faculty (Tyler et al., 2020). However, the number of women faculty could be considered important since it was higher than the national average (40% compared to 20%). The Physics Group's building did not change during the period of implementation and remained the oldest academic building without renovation.

Figure 1 displays the 3-year averages for the number of physics majors officially declared and bachelor's degree recipients. The data covers a 20-year time period, which extends significantly before the implementations were added. As a reminder, the data column "2013–2015" covers the initiation period for several innovative elements, including the first-year colloquium. The following averages are more revealing when the innovative elements permeate a physics major's full matriculation. Specifically, from 2016 to 2021, the average number of graduates was 9.8, while previously it was 3.9 (2001–2015). Not only did the number of graduates increase by more than two-fold, but the yearly fluctuation stabilized considerably.

One immediate result of adopting an inclusive, cohort-building mindset was the admittance of Sophomores (second-year students) as declared majors. This practice began in 2013 and helped explain the significant increase in "declared" column between 2012 and 2013. Therefore, declared majors include second-, third-, and fourth-year students intentionally pursuing a physics degree. Prior to 2013, a "gate-keeping" mindset was more prevalent in the Physics Group, which required majors to show proficiency in upper-level coursework. This change in mindset and practice afforded students earlier access as members of the cohort, including a greater sense of connection to the discipline. This decision does not fully account for all the increases observed, since the number of graduates also increased significantly after 2016.

Lastly, the noticeable increases of physics majors and graduates cannot be due to a weakening of the program or a loosening of accountability for its majors. During the time period covered in Figure 1, the number of credit units for the major remained roughly the same. In fact, with the addition of the first-year colloquium and lab restructuring (c.f., Section 2.2.1), one could make the argument that the amount of coursework increased by at least one unit during the time period. Student expectations and engagement heightened due to programmatic augmentation, which was evidenced by recognition for both individuals (Goldwater Scholar) and the program (SPS Chapter Award).

3.2. Increased number of women degree recipients

While the numbers of physics degree recipients have increased considerably over the last 40 years, the percentages for women recipients remain approximately constant at 20±5% since 2,000 (Porter and Ivie, 2019). Figure 2 displays the number of women bachelor's recipients in the Physics Group over a 20-year period. Specifically, the yearly average of women graduating in physics from 2001 to 2012 is 0.75, while the same average from 2013 to 2021 is 2.4. As a 200% increase is weighed, a few considerations are addressed.

The undergraduate institution where the Physics Group is located enrolls a higher percentage of women undergraduate students, usually around 60%, which is also the case for the period of the study. However, prior to the study (2001–2012), the average percentage of women was higher (closer to 65%). No particular alterations were made that would easily explain the significant rise in women physics graduates. For the years of focused implementation (2013–2021), the percentage of women graduates held steady around 30%, which is about 10% above the US average. The retention of women in physics during the implementation period is now examined as a function of two innovative elements.

3.3. Increased participation in experiential learning

Over the last 10 years, there has been a steady increase in the number of EEL opportunities from which our students have benefited. Figure 3 displays data to support evidence of growth in the number of physics majors participating in one or more EELs. This includes all research performed by students on campus supervised by physics faculty, collaborative and/or inter-disciplinary research projects with other departments (chemistry, computer-science, mathematics), off campus research experiences such as NSF sponsored REUs and collaborative projects supervised by off campus mentors, and internships. Several of our physics graduates often have more than one such EEL. Between 2013 and 2018, the 5-year average number of on-campus EEL for our students increased to 90% compared to 69% for the previous 5 years 2007–2011. More starkly, the numbers of off-campus EEL for our majors has increased by a factor of six (7–45%) over a similar period. Figure 3 reflects the growth in each area over the time-period considered.

3.4. Retention of women physics majors

Prior to 2013, there was no consistent means of accounting for matriculation in the physics major as a function of original interest. For example, there was no direct measure of incoming interest compared with the number of students who enrolled in the first semester of calculus-based physics (in Spring). Therefore, the retention rates presented since the implementation period (2013) have no prior comparative data. That said, it is still clear that the number of women retained within the physics major accord with the innovative elements introduced in the program (cf., Section 2.2).

3.4.1. With respect to first-year colloquium enrollment

Since 2013, a first-year colloquium was required for the physics degree, and it was strongly recommended by pre-admission advising to all incoming students who were interested in physics and/or engineering. The enrollment for the course increased steadily for every pre-pandemic year, from 19 in 2013 to 36 in 2019. Figure 4 presents the retention percentages by gender for original enrollers who persisted until graduation. Due to the frequency of course offering and higher enrollments, retention percentages are provided for each year during the study. The asterisk for 2019 indicates those who had not officially graduated before the completion of the study, so their retention was measured prematurely (at the end of the third year). Since it is in the third year that upper-level coursework begins, it is held with more certainty that a student would be retained. With the exception of a single year (2015), women were retained within the physics major at a higher percentage than men. Therefore, during the time period of the study, women were retained in the physics major at a higher average percentage than men.

Even though the first-year colloquium continues, more recent numbers are not given here for two reasons. In Figure 4, 2019 is the last year that majors can confidently be reported as matriculating through the major, since a student's first- and second-year remain less certain. Second, since the course maintains a high relational component, pandemic effects are unclear for both institutional and colloquium enrollments.

3.4.2. With respect to conference participation

In particular, the Conference for Undergraduate Women in Physics (CUWiP) provided a timely and specific opportunity for women physics majors in the Physics Group. These regional conferences are supported by the APS through funding from the National Science Foundation (NSF) and the Department of Energy (DOE). Typically, associated costs for conference participants are subsidized. Figure 5 shows the increase in numbers of women participants at CUWiP conferences, where each participant is counted only once. These increases are independent and irrespective of the increases in total numbers of majors.

3.4.3. With respect to EEL participation

As a reminder, the following types of opportunities are included as EEL participation: REU experiences, on-campus research with

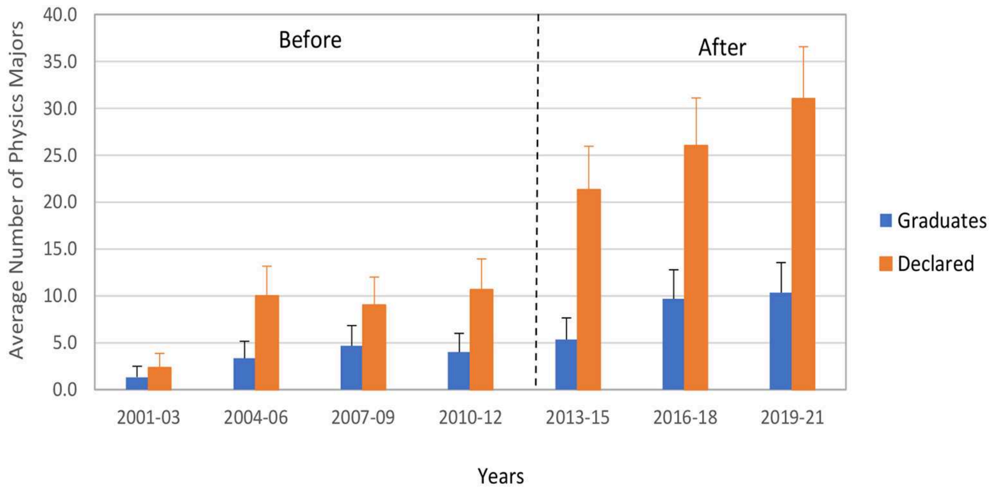

FIGURE 1
Three-year averages for the numbers of declared physics majors ("declared") and the number of undergraduate degree recipients ("graduates"). Differences between the number of declared and the number of graduates are based on the definition of physics majors as discussed in the text (Section 1.2). Vertical dashed line demarcates the beginning of the innovation initiation.

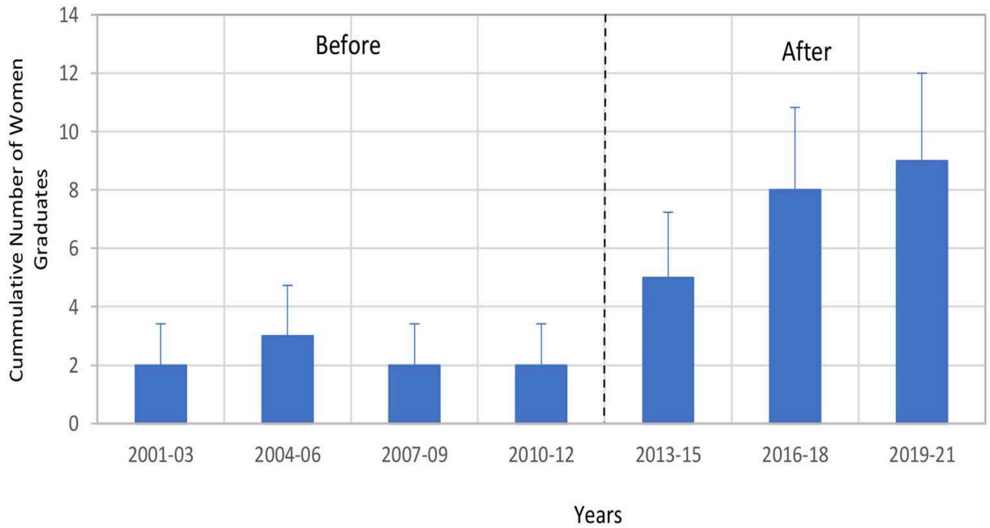

FIGURE 2
Three-year, cumulative numbers of women who received an undergraduate physics degree in the Physics Group over a 20-year period. Vertical dashed line demarcates the beginning of the innovation initiation.

STEM faculty, STEM-related educational experiences, and physics-related internships. Students who participate in multiple EEL, particularly within historically underrepresented groups, are shown to positively correlate with self-evaluated STEM identity (Estrada et al., 2018). Figure 5 shows the relationship between women physics graduates and their participation in EEL opportunities. The number of EEL accords with the increases in women physics degree recipients. For Figure 5, multiple EEL for the same woman, physics major are included in the total, so only for the years "2016–2018" are there an average of > 2 EEL per woman graduate. During the years of innovative element implementation (2013–2021), there was greater than one EEL opportunity per woman graduate. For years prior to 2007, there was no EEL information cataloged.

3.5. Retention of women within STEM careers

As displayed in Figure 2, noticeable and consistent increases in women degree recipients begin during the "2013–2015" segment, which coincides with the initiation of most innovative elements. This same 3-year segment also marks the first time where the average number of physics majors eclipses 20.0, and total graduate

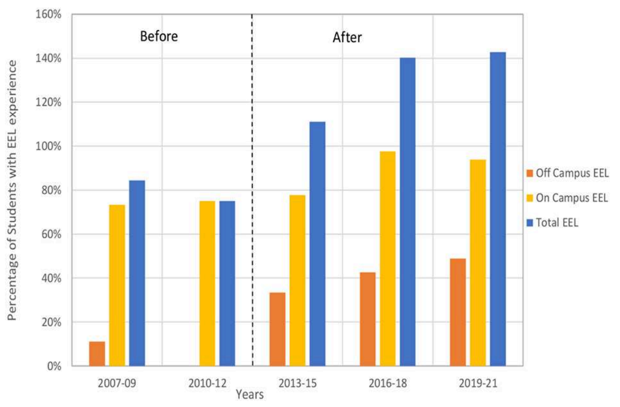

FIGURE 3
Percent participation for all Physics Group majors in one (or more) EEL opportunities averaged over 3-year periods. Multiple experiences by the same student were tallied, therefore 100% participation should not be interpreted as every major had a single experience. "On Campus" EEL include inter-disciplinary opportunities with programs like biology, chemistry, computer science, or mathematics. "Off Campus" EEL include internships, REUs, and other collaborative partnerships.

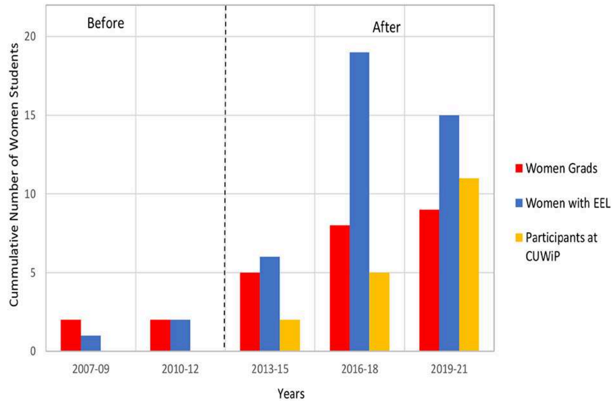

FIGURE 5
Retention-to-Graduation for women in the Physics Group in comparison with EEL and CUWiP participation. Cumulative number of participants at the APS Conference for Undergraduate Women in Physics (CUWiP) by 3-year grouping. Prior to 2013, there was no CUWiP attendance in the Physics Group. If a participant attended more than once, it is only tallied at the first instance and suppressed for subsequent years. Multiple EEL for the same student were included in the tallies.

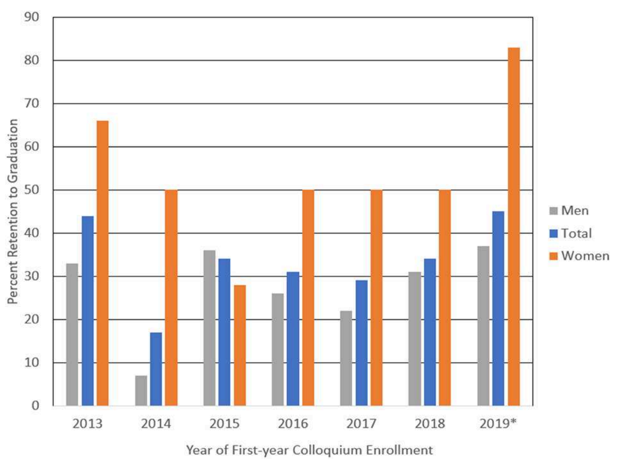

FIGURE 4
Retention-to-graduation within the Physics Group for participants in the first-year colloquium by gender. The percentages are based on final enrollments in colloquium and final numbers of degree recipients. The asterisk for 2019 indicates those who had not officially graduated before the completion of the study. Their retention was calculated after the third-year of completion.

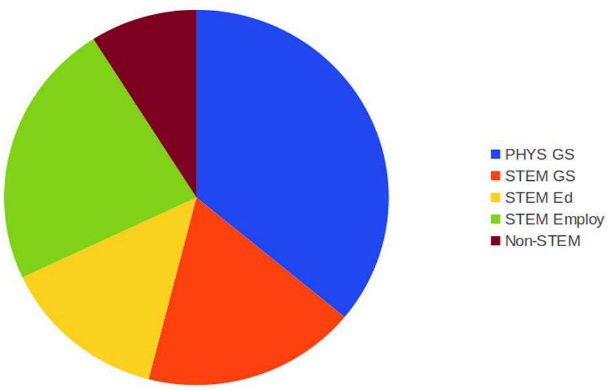

FIGURE 6
Post-graduate career trajectory decisions for 22 women degree recipients from the Physics Group in the years 2013–2021. Decisions were categorized 1-year after graduation. All but two of the graduates were retained within STEM advancement positions. Legend categories are discussed in the text (Section 3.5).

average above 5.0 (c.f. Figure 1). Since immediately after graduation presents another pressure point for attrition from STEM fields, we maintain an interest in what these retained physics graduates do after they receive their degree.

Figure 6 presents the decisions for 22 women degree recipients from 2013 to 2021, 1-year after graduation. Decisions are separated loosely along the following categories: Physics-related graduate school (PHYS GS), non-Physics, STEM-related graduate school (STEM GS), STEM-related employment professions (STEM Employ), STEM-related education professions (STEM Ed), and non-STEM employment (non-STEM). Of the 22 women, <10% (2/22) persist in a career not directly related to STEM advancement. Here, advancement is defined as involvement in STEM research and support (STEM Employ), learning (GS), or education (STEM Ed). Therefore, women graduates from the Physics Group matriculate into STEM-related trajectories at ∼90% during the 2013–2021 time period.

Graduate training, whether physics-related or STEM broadly, comprised over half of the post-graduate decisions (12/22) for these women. Physics-related training (PHYS GS) included physics subdisciplines, engineering, and materials science. More broadly, STEM-related, graduate training (STEM GS) included fields like computer science, applied math, and veterinary medicine. A few of the women graduates (3) completed their graduate training since leaving the Physics Group, and all three matriculated into STEM-related employment. Moreover, at the time of publishing, all of these women graduate students remained in their programs, ranging from 1 to 6 years later, or graduated from them.

When considering the entire group of women bachelor recipients from 2013 to 2021, a more broad statement emerges about their longevity within STEM advancement careers. Through continuing to follow the post-graduate trajectories of the 20 women already involved in STEM advancement, we find that all of them continue to find meaningful, STEM-related employment, even up to 8 years after graduation, beginning in 2013. Even though some of them have transitioned into parenthood, parallel fields, and/or promotion, these women continue to thrive within STEM-advancement careers.

4. Discussion

To summarize, several innovative elements augmented the undergraduate program in the Physics Group over a period of 8 years. With a staggered initiation, these additions were categorized as curricular, co-curricular, or experiential (Section 2.2). Some of these elements were introduced as a result of national organization documentation, some from STEM literature, and some were novel within the Physics Group. Changes and growth in the program were noted over the same 8-year period, from 2013 to 2021. There were three primary results that accorded with the addition of these innovative elements. First, the total number of majors and graduates increased by approximately 200% compared with the previous 13 years (Figure 1). Second, the number of women in the major were retained at a higher rate than men (Figure 4). Third, these women graduates were employed in STEM-advancement positions at ∼90%, from 1-year after graduation and extending out to 8 years post-graduation (Figure 6). Since the implementation and evaluation period of the elements overlapped, it was impossible to determine which elements contributed specifically to each result. That said, we provide a guiding analogy to frame an interpretative discussion of these results within the context of programmatic change.

4.1. Leaky pipelines and delta distributaries

Historically, "leaky pipelines" refer to losses in STEM representation, ranging from aggregate (National Institute of Food and Agriculture, 2014) to minoritized populations (Liu et al., 2019), and even both (Metcalf, 2010). More specifically, the analogy pertains to a decrease in STEM participation for women, initially beginning in transitions from secondary to undergraduate (Archer et al., 2016), then extending to Ph.D. representation (Miller and Wai, 2015) and careers in academia (Sheltzer and Smith, 2014). Systemic problems are prevalent with historically underrepresented groups that are more clearly recognized at formal transitions, though other alternatives are suggested (Rainey et al., 2018; Witteveen and Attewell, 2020). Logically, since the numbers of men and women in K-12 are roughly equal, including the numbers of students taking physical science coursework in high school, then there should be roughly equal numbers of women and men

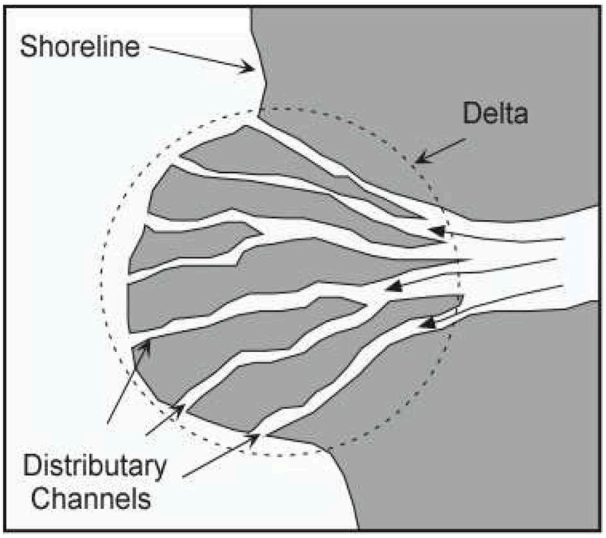

FIGURE 7
A typical schematic of a river delta. The primary tributary (right side) branches into several distributaries as it nears the much larger body of water. The Physics Group image of a typical undergraduate's matriculation through the program loosely follows this image (https://www2.tulane.edu/~sanelson/Natural_Disasters/riversystems.htm).

in STEM careers. For women in physics and engineering, there are greatly reduced numbers of bachelor's recipients, which is interpreted as a "leak." Sometimes the cause of a leak is focused on gender discrimination (Grogan, 2018) and sometimes more broadly on the nature of science (Blickenstaff, 2005). However, questions about the analogy have been raised, either as to its efficacy toward improvements (Cannady et al., 2014) or its over-simplicity (Hinton et al., 2020).

The innovative changes implemented in the Physics Group offer a more inclusive and branched approach than a pipeline analogy. By cultivating a "hidden physicist" model, the program encourages students to view physics as a pathway to diverse STEM careers. Perhaps, a more flexible image is more helpful. Specifically, the Physics Group considers a river delta analogy where distributaries (post-graduation decisions) are initially kept broad within a primary tributary (major program) but allowed to spread and disseminate at the delta by the third- and fourth-years. Figure 7 shows a schematic of a typical river delta. While not completely accurate in every detail, we note that river delta plains become fertile areas for new growth. By leaving unanswered the question, "What can a physics major do?," any number of post-graduation options are encouraged (Fleenor, 2018). We examine a river delta analogy within two important concepts, STEM identity and programmatic structure.

4.2. Programmatic structure, building the banks of the tributary

Within a river delta analogy, the number and strength of the distributaries depends on the tributary. In our analogy, the augmented physics program serves as the tributary and includes student majors with faculty. The innovative elements discussed in Section 2.2 provide fortified banks for the increased number of majors. To be clear, "fortifying" does not mean constraining individual choices, whereby students may choose to leave the program and/or switch majors. In the establishment of STEM identity, freedom and encouragement must be given as students persist through difficult circumstances (Fleenor, 2021). By drawing on best-practices from PER literature, in combination with unique innovations specific to the Physics Group (Henderson et al., 2015), a well-fortified tributary is kept. Two implemented examples from the physics program confirm the imagery of building strong banks for women in undergraduate trajectories.

The first-year colloquium serves as the tributary entry point. Keeping the course in an exploratory state, while also introducing a broad diversity of those who are in the Physics Group, seems to keep undecided participants in an interested state. The connection seems clear between the increase of majors and graduates overall (Figure 1) and the introduction of the first-year colloquium (Figure 4). An open award structure for physics excellence rewards the unique contributions and accomplishments of many who persevere in one of the most rigorous disciplines (Section 2.2), while early participation within many informal programs sponsored by the Physics Group helps to establish identity (Rethman et al., 2021). By continuing to keep the tributary entry broad, a student comes to the Junior Review event with a greater opportunity of successful STEM experiences.

Synergizing CURE results from the literature with the curriculum structure in the Physics Group was an innovation beyond documentation (Corwin et al., 2015; Wooten et al., 2018; Reichart, 2019). With the implementation of a CURE-based laboratory, students were able to receive some partial instruction in all upper-level electives, but it also facilitated further EEL persistence. By implementing a half-credit research/independent study experience that usually precedes a larger, full-credit summer experience, the Physics Group better connects research experiences between the coursework and beyond. These pre-experiences parlay well and create inertia for a student's formative summer EEL opportunities (particularly after their third-year). Such opportunities allow scaffolding of skills and strengthening a knowledge-base through a research experience, which promote a heightened persistence in undergraduates (Estrada et al., 2011). All of these EEL experiences serve to fortify the banks of the program tributary facilitating students to successful STEM trajectories post-graduation.

4.3. STEM identity, flooding the delta

Many of the innovative elements reinforce the development of interpersonal and intrapersonal factors known to affect retention in the STEM disciplines, such as science (STEM) identity and sense of belonging. As science identity is predictive of longer-term persistence in STEM-related fields (Eren, 2021), and sense of non-belonging is cited as a reason for "leaving" the sciences (especially for women, Lewis et al., 2017), these aims are essential to the presented study and well-supported in the literature. Specifically, Estrada et al. (2018) note that completing multiple semesters of research and/or internship within STEM correlates strongly with establishing identity within STEM. Related, Findley-Van Nostrand and Pollenz (2017) demonstrate that participation in an intensive week-long co-curricular program just as students enter college, including engaging with peers with similar interests in addition to several elements related to reflecting on goals and career aims and getting connected to faculty, is related to increased science identity and sense of belonging. In addition, increases in science identity seem to be driven by increases in belonging (to STEM and institutional community, Kuchynka et al., 2019), suggesting that efforts such as providing greater connection to faculty and peers from early on in the program facilitate such development. As research suggests, formation of identity is iterative as students grow in confidence through repeated interaction with content, problem-solving, and experimental techniques (Keagan, 2018). It would be unsurprising if women were better established as shareholders and valued members, where the numbers of women graduates increased.

To substantiate a distributaries analogy within the Physics Group, measures in Figure 4 seem to reduce attrition at a high school to undergraduate transition, since women are retained in the major at higher rates than men. Specifically, if a high school student shows high aptitude and interest in physics (so-called, "Exceptional Physics Girls," Archer et al., 2016), then she is retained at rates above national averages in the Physics Group. The result in Figure 4 accord with an attempt to keep the primary tributary as broad as possible, early within the undergraduate experience.

Similarly, if a woman receives her bachelor's degree in the Physics Group, then her STEM identity seems more solidified as she moves into a STEM advancement position (Figure 6). When examining the approximate percentages shown in Figure 6, it is clear that a greater number of women in STEM advancement occupy the non-PHYS GS categories. That is, fewer women are retained to become (traditional) physicists than not. However, the Physics Group still considers this successful, since the goal is to flood the delta. A "flooded delta" represents scenarios where there is a greater likelihood that young women (K-12 students) will see someone like them in STEM-advancement positions. Without becoming elitist, the Physics Group believes that "hidden physicists" critically participate in other non-physics, STEM fields, because physics provides a unique way of knowing (Marder, 2013). These unique pathways serve to diversify the face of science by increasing the number of distributaries for physics bachelor recipients.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: Roanoke College Institutional Research Office.

Author contributions

RB and MF contributed to conception, implementation of programmatic innovation, constructed the histograms, organized, and collected the pertinent data. MF wrote most of the manuscript's first draft, where DF-VN provided significant contribution. DF-VN contributed to the construction of STEM identity and belonging framework. All authors contributed to the article and approved the submitted version.

Funding

MF thanks the Roanoke College sabbatical program for its generosity in reassign-time. MF also thanks the Margaret Duke Endowment at the University of Mary Washington from which reassign-time was provided.

Acknowledgments

All authors appreciate the Office of Institutional Research and the Registrar at Roanoke College for their gracious cooperation.

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.

Publisher's note

References

AAPT/APS. (2021). *Effective Practices for Physics Programs (EP3).* Available onliine at: https://ep3guide.org/ (accessed November 17, 2021).

Abbott, P., and Sapsford, R. (2005). Living on the margins: older people, place and social exclusion. *Policy Stud.* 26, 29–46. doi: 10.1080/01442870500041660

AIP. (2021). *Statistical Research Center (SRC).* Available online at: https://www.aip.org/statistics (accessed November 17, 2021).

Alon, U. (2009). How to choose a good scientific problem. *Mol. Cell* 35, 726–728. doi: 10.1016/j.molcel.2009.09.013

APS. (2020). *Bachelor's Degrees in Physics and STEM.* Available online at: https://www.aps.org/programs/education/statistics/bachelors.cfm(accessed November 18, 2021).

Archer, L., Moote, J., Francis, B., DeWitt, J., and Yeomans, L. (2016). The "exceptional" physics girl: a sociological analysis of multimethod data from young women aged 10–16 to explore gendered patterns of post-16 participation. *Am. Educ. Res. J.* 54, 88–126. doi: 10.3102/0002831216678379

Auchincloss, L., Laursen, S., Branchaw, J., Eagan, K., Graham, M., Hanauer, D., et al. (2014). Assessment of course-based undergraduate research experi- ences: a meeting report. *CBE Life Sci. Educ.* 13, 29–40. doi: 10.1187/cbe.14-01-0004

Blickenstaff, J. C. (2005). Women and science careers: leaky pipeline or gender filter? *Gend. Educ.* 17, 369–386. doi: 10.1080/09540250500145072

Cannady, M. A., Greenwald, E., and Harris, K. N. (2014). Problematizing the STEM pipelinemetaphor: is the STEM pipeline metaphor serving our studentsand the STEM workforce? *Sci. Educ.* 98, 443–460. doi: 10.1002/sce.21108

Corwin, L. A., Graham, M. J., and Dolan, E. L. (2015). Modeling course-based undergraduate research experiences: an agenda for future research and evaluation. *CBE Life Sci. Educ.* 14, es1. doi: 10.1187/cbe.14-10-0167

Denice, P. (2021). Choosing and changing course: postsecondary students and the process of selecting a major field of study. *Sociol. Perspect.* 64, 82–108. doi: 10.1177/073112142091903

Diekman, A., Steinberg, M., Brown, E., Belanger, A., and Clark, E. (2017). 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

Dunne, E., and Zandstra, R. (2011). *Students as Change Agents – New Ways of Engaging with Learning and Teaching in Higher Education, 1 Edn.* Bristol, UK: University of Exeter/ESCalate/HE Academy.

Eren, E. (2021). Exploring science identity development of women in physics and physical sciences in higher education. *Sci. Educ.* 30, 1131–1158. doi: 10.1007/s11191-021-00220-3

Estrada, M., Hernandez, P. R., and Schultz, P. W. (2018). A longitudinal study of how quality mentorship and research experience integrate underrepresented minorities into stem careers. *CBE Life Sci. Educ.* 17, ar9. doi: 10.1187/cbe.17-04-0066

Estrada, M., Woodcock, A., Hernandez, P. R., and Schultz, P. W. (2011). Toward a model of social influence that explains minority student integration into the scientific community. *J. Educ. Psychol.* 103, 206–222. doi: 10.1037/a0020743

Findley-Van Nostrand, D., and Pollenz, R. (2017). Evaluating psychosocial mechanisms underlying stem persistence in undergraduates: evidence of impact from a six-day pre–college engagement stem academy program. *CBE Life Sci. Educ.* 16, ar36. doi: 10.1187/cbe.16-10-0294

Fleenor, M. C. (2018). Cultivating experimental innovation within undergraduate physics majors. *Glob. Educ. Rev.* 5, 73–87. Available online at: https://ger.mercy.edu/index.php/ger/article/view/419

Fleenor, M. C. (2021). Grounded mentoring as a pathway for program-building. *Chron. Mentor. Coach.* 5, 421–428. Available online at: https://www.mentor-cmc.com/cmc/cmc2021/MobilePagedReplica.action?pm=2&folio=420#pg420

Francis, B., Archer, L., Moote, J., DeWitt, J., MacLeod, E., and Yeomans, L. (2017). The construction of physics as a quintessentially masculine subject: young people's perceptions of gender issues in access to physics. *Sex Roles* 76, 156–174. doi: 10.1007/s11199-016-0669-z

Gisler, S., Kato, A. E., Lee, S., and Leung, D. (2018). One size does not fit all: gender inequity in STEM varies between subfields. *Indust. Organ. Psychol.* 11, 314–318. doi: 10.1017/iop.2018.21

Good, C., Aronson, J., and Harder, J. (2008). Problems in the pipeline: stereotype threat and women's achievement in high-level math courses. *J. Appl. Dev. Psychol.* 29, 17–28. doi: 10.1016/J.APPDEV.2007.10.004

Good, C., Rattan, A., and Dweck, C. (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

Gowar, C. (2013). *Inclusive Communities: A Research Report.* Technical report, Disability Rights UK. Available online at: https://www.disabilityrightsuk.org/sites/default/files/pdf/3.%20InclusiveCommunitiesResearch.pdf (accessed November 22, 2021).

Grogan, K. E. (2018). How the entire scientific community can confront gender bias in the workplace. *Nat. Ecol. Evol.* 3, 3-6. doi: 10.1038/s41559-018-0747-4

Henderson, C., Cole, R., Froyd, J., Friedrichsen, D., Khatri, R., and Stanford, C. (2015). *Designing Educational Innovations for Sustained Adoption: A How-to-Guide for Education Developers Who Ant to Increase the Impact of Their Work, 1st Edn.* Kalamazoo, MI: Increase the Impact.

Heron, P., and McNeil, L. (2016). *Phys21: Preparing Physics Students for 21st-Century Careers (J-TUPP), 1st Edn.* College Park, MD: The American Physical Society (APS).

Hilborn, R. C., Howes, R. H., and Krane, K. S. (2003). *Strategic Programs for Innovations in Undergraduate Physics: Project Report (SPIN-UP), 1st Edn.* College Park, MD: The American Association of Physics Teachers (AAPT).

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

Hinko, K. A., Madigan, P., Miller, E., and Finkelstein, N. D. (2016). Characterizing pedagogical practices of university physics students in informal learning environments. *Phys. Rev. Phys. Educ. Res.* 12, 010111. doi: 10.1103/PhysRevPhysEducRes.12.010111

Hinton, A. O. Jr., Termini, C. M., Spencer, E. C., Rutanganira, F. U. N., Cherry, D., Roby, R., et al. (2020). Patching the leaks: revitalizing and reimagining the STEM pipeline. *Cell* 183, 568–575. doi: 10.1016/j.cell.2020.09.029

Hunt, J. (2013). An alternate view: is industry really a "non-traditional" career? *APS Forum on Industrial and Applied Physics Newsletter*.

Ivie, R., and Tesfaye, C. L. (2012). Women in physics: a tale of limits. *Phys. Today* 65, 47–50. doi: 10.1063/PT.3.1439

James, M., and Bertschinger, E. (2020). *The Time Is Now: Systemic Changes to Increase African Americans wiht Bachelor's Degrees in Physics and Astronomy (TEAM-UP), 1st Edn*. College Park, MD: The American Institute of Physics (AIP).

Johnson, T. A. (2021). APS releases new report: building America's STEM workforce. *APS News* 30, 1–4. Available online at: https://www.aps.org/publications/apsnews/202111/congressional.cfm

Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., and Singh, C. (2020). Damage caused by women's lower self-efficacy on physics learning. *Phys. Rev. Phys. Educ. Res.* 16, 010118. doi: 10.1103/PhysRevPhysEducRes.16.010118

Keagan, R. (2018). "What "form" transforms? a constructive-developmental approach to transformative learning," in *Contemporary Theories of Learning: Learning Theorists . . . In Their Own Words, 2nd Edn.*, ed K. Illeris (New York, NY: Routledge), 35–52.

Kuchynka, S., Findley-Van Nostrand, D., and Pollenz, R. (2019). Evaluating psychosocial mechanisms underlying stem persistence in undergraduates: scalability and longitudinal analysis of three cohorts from a six-day pre–college engagement stem academy program. *CBE Life Sci. Educ.* 18, 1–13. doi: 10.1187/cbe.19-01-0028

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: belonging, gender, and persistence in the physical sciences, technology, engineering, and mathematics (pSTEM). *Psychol. Women Quart.* 41, 420–436. doi: 10.1177/0361684317720186

Liu, S.-N. C., Brown, S. E. V., and Sabat, I. E. (2019). Patching the "leaky pipeline": interventions for women of color faculty in STEM academia. *Arch. Sci. Psychol.* 7, 32–39. doi: 10.1037/arc0000062

Marder, M. (2013). A problem with STEM. *CBE Life Sci. Educ.* 12, 148–150. doi: 10.1187/cbe.12-12-0209

Metcalf, H. (2010). Stuck in the pipeline: a critical review of STEM workforce literature. *Interact. UCLA J. Educ. Inform. Stud.* 6, 1–21. doi: 10.5070/D462000681

Miller, D. I. and Wai, J. (2015). The bachelor's to Ph.D. stem pipeline no longer leaks more women than men: a 30-year analysis. *Front. Psychol.* 6:37. doi: 10.3389/fpsyg.2015.00037

Mordacq, J. C., Drane, D. L., Swarat, S. L., and Lo, S. M. (2017). Development of course-based undergraduate research experiences using a design-based approach. *J. Coll. Sci. Teach.* 46, 64–75. Available online at: http://www.jstor.org/stable/44579916

Mulvey, P. (2021). Electronic communication.

Nash, B. J. (2012). Journey to work: European model combines education with vocation. *Fourth Quart.* 16, 17–19, Available online at: https://fraser.stlouisfed.org/title/3941/item/476957

National Institute of Food and Agriculture. (2014). *Addressing Challenges in Education.* Technical report, National Institute of Food and Agriculture. Available online at: https://nifa.usda.gov/sites/default/files/resource/ed_facts_8_8_fnl.pdf

Nelson, D. J., and Brammer, C. N. (2010). *A National Analysis of Minorities in Science and Engineering Faculties at Research Universities.* Technical report. Available online at: http://drdonnajnelson.oucreate.com//diversity/Faculty_Tables_FY07/07Report.pdf (November 25, 2021).

Porter, A. M., and Ivie, R. (2019). *Women in Physics and Astronomy, 2019.* AIP Reports.

Rainey, K., Dancy, M., Mickelson, R., Stearns, E., and Moller, S. (2018). Race and gender differences in how sense of belonging influences decisions to major in STEM. *Int. J. STEM Educ.* 5, 14. doi: 10.1186/s40594-018-0115-6

Reichart, D. (2019). *Exploring the Impact of Robotic Telescope-Based Observing Experiences on Students' Learning and Engagement in STEM.* Technical Report, National Science Foundation, USA. https://www.nsf.gov/awardsearch/showAward?AWD_ID=2013300&HistoricalAwards=false

Reinholz, D. L., Matz, R. L., Cole, R., and Apkarian, N. (2019). STEM is not a monolith: a preliminary analysis of variations in STEM disciplinary cultures and implications for change. *CBE Life Sci. Educ.* 18, mr4. doi: 10.1187/cbe.19-02-0038

Rethman, C., Perry, J., Donaldson, J. P., Choi, D., and Erukhimova, T. (2021). Impact of informal physics programs on university student development: creating a physicist. *Phys. Rev. Phys. Educ. Res.* 17, 020110. doi: 10.1103/PhysRevPhysEducRes.17.020110

Roy, M.-F., Guillope, C., Cesa, M., Ivie, R., White, S., Mihaljevic, H., et al. (2020). *A Global Approach to the Gender Gap in Mathematical, Computing, and Natural Sciences: How to Measure It, How to Reduce It?, 1st Edn.* Geneva: Zenodo.

Sheltzer, J. M., and Smith, J. C. (2014). Elite male faculty in the life sciences employfewer women. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10107–10112. doi: 10.1073/pnas.1403334111

Smyth, F., and Nosek, B. (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

Stewart, J., Oliver III, W., and Stewart, G. (2013). Revitalizing an undergraduate physics program: a case study of the University of Arkansas. *Am. J. Phys.* 81, 943–950. doi: 10.1119/1.4825039

Taylor, J. (1997). *An Introduction to Error Analysis, 2nd Edn.* Sausalito, CA: University Science Books.

Tesfaye, C. L., and Mulvey, P. (2013). *MCAT, LSAT and Physics Bachelor's.* Technical report, AIP. Available online at: https://www.aip.org/statistics/reports/mcat-lsat-and-physics-bachelors

The Advanced Laboratory Physics Association, ALPhA. (2021). *Alpha Homepage.* Available online at: https://advlab.org/ (November 22, 2021).

Trujillo, G., and Tanner, K. (2014). Considering the role of affect in learning: monitoring students' self-efficacy, sense of belonging, and science identity. *CBE Life Sci. Educ.* 13, 6–15. doi: 10.1187/cbe.13-12-0241

Tyler, J., and Mulvey, P. (2022). *How Well do Physics Bachelor's Degree Recipients Perform on the MCAT and LSAT Exams?* Technical Report, AIP. Available online at: https://www.aip.org/statistics/reports/how-well-do-physics-bachelor%E2%80%99s-degree-recipients-perform-mcat-and-lsat-2022

Tyler, J., Mulvey, P., and Nicholson, S. (2020). *Size of Undergraduate Physics and Astronomy Programs: Results from the Enrollments and Degrees and Academic Workforce Surveys.* Technical report, AIP. Available online at: https://files.eric.ed.gov/fulltext/ED607301.pdf (November 15, 2021).

UNESCO Institute for Statistics. (2019). *Women in Science.* Available online at: http://uis.unesco.org/en/topic/women-science (accessed January 26, 2023).

United Nations. (2023). *United Nations Sustainable Development Goals.* Available online at: http://www.un.org/sustainabledevelopment (accessed January 26, 2023).

US Department of Education. National Center for Education Statistics. (2019). *Indicator 26: STEM Degrees.* Available online at: https://nces.ed.gov/programs/raceindicators/indicator/reg.asp (accessed November 16 2021).

Wiswall, M.J., and Zafar, B. (2021). New approaches to understanding choice of major. *NBER Report.* 2, 18–21. Available online at: https://www.nber.org/reporter/2021number2/new-approaches-understanding-choice-major

Witteveen, D. and Attewell, P. (2020). The STEM grading penalty: an alternative to the "leaky pipeline" hypothesis. *Sci. Educ.* 104, 714–735. doi: 10.1002/sce.21580

Wooten, M., Coble, K., Puckett, A., and Rector, T. (2018). Investigating introductory astronomy students' perceived impacts from participation in course-based undergraduate research experiences. *PRPER* 14, 010151. doi: 10.1103/PhysRevPhysEducRes.14.010151

World Economic Forum. (2019). *The Global Gender Gap Index 2020.* Technical report, World Economic Forum. Available online at: http://reports.weforum.org/global-gender-gap-report-2020/the-global-gender-gap-index-2020/

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 17 April 2023
DOI 10.3389/feduc.2023.1138285

Check for updates

OPEN ACCESS

EDITED BY
Karen Blackmore,
University of Worcester, United Kingdom

REVIEWED BY
Cucuk W. Budiyanto,
Sebelas Maret University, Indonesia
Aris Budianto,
Sebelas Maret University, Indonesia
Rosihan Ari Yuana,
Sebelas Maret University, Indonesia

*CORRESPONDENCE
Greta Heim
✉ greta.heim@uit.no

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to STEM Education, a section of the journal Frontiers in Education

RECEIVED 05 January 2023
ACCEPTED 28 March 2023
PUBLISHED 17 April 2023

CITATION
Heim G and Wang OJ (2023) Block and unplugged programming can be mutually beneficial: A study of learning activities in a 6th grade class in Norway.
Front. Educ. 8:1138285.
doi: 10.3389/feduc.2023.1138285

Block and unplugged programming can be mutually beneficial: A study of learning activities in a 6th grade class in Norway

Greta Heim*† and Oskar Jensen Wang†

Department of Education, Faculty of Humanities, Social Sciences and Education, UiT The Arctic University of Norway, Tromsø, Norway

In the new Norwegian national curricula, programming and interdisciplinary work have been introduced as two central elements. Furthermore, computational thinking is part of the core elements of the mathematics curriculum. In this paper, we present the findings from a small-scale study within the subjects of mathematics and food and health. The aim was to see if these two subjects could be used as an arena for working with parts of computational thinking, in this case algorithmic thinking and collaboration, and expanding the students' understanding of what programming can be. We think there is a gap in the research regarding this topic, and therefore we wanted to look into this. The students who took part in the study carried out two lessons, one in each subject. In mathematics, the focus was on block programming, while food and health can be seen as unplugged programming. Our findings are based on feedback from 36 students and observations from the lessons. The main finding is that most of the students could not see a link between programming and food and health. Only seven students expressed something that indicated that they did see a link. Thus, it seems like most of the students could not see similarities between the algorithmic thinking in block programming and recipes in food and health.

KEYWORDS

computational thinking, unplugged programming, block programming, mathematics, food and health

1. Introduction

In Norway, programming has entered as a central element in the Norwegian national curricula from 2020 (Sevik, 2016). In this article, we present a small study we have carried out together with pre-service teachers. The purpose was to investigate whether students in the 6th grade could recognize parts of computational thinking in two lessons involving block programming and unplugged programming in the subjects mathematics and food and health. Elements of computer science, including programming and computational thinking, have been included in the school curricula in an increasing number of countries in the later years (Nouri et al., 2020). To our knowledge there is done little research on the topic of using the subject food and health as an arena for programming. Therefore, we believe this research could not only be of interest to Norwegian schools, but also a wider audience.

Computational thinking is closely linked to programming and coding (Gjøvik and Torkildsen, 2019), and has been included as a part of the core elements of the mathematics curriculum (Ministry of Education and Research, 2019). Computational thinking can be seen as a thought process that involves creating solutions that can be executed by computers or humans (Bocconi et al., 2018), or involves problem-solving (Ardito et al., 2020). Algorithmic thinking is one of several key concepts within computational thinking (Bocconi et al., 2018), and collaboration is a key component (Ardito et al., 2020). Many of the concepts and approaches within computational thinking can be practiced through unplugged programming (Bell and Vahrenhold, 2018).

By giving different kinds of problems to students, programming can be used to practice collaboration and discussions/reasoning (Sevik, 2016). Furthermore, it is pointed out that students can use their creativity and imagination in the work from idea to action. Interaction, communication and exploring and creating have been highlighted as important competences to be emphasized in the school of the future with the intention of educating future professionals (NOU, 2015: 8). A central element in the new Norwegian core curriculum, that decides the values and principles for primary and secondary education, is interdisciplinary, and one of the goals is that students achieve understanding and see connections across subjects (Ministry of Education and Research, 2017). We wanted to try to combine mathematics and a practical-aesthetic subject, such as food and health, and see if this combination could give some benefits. Food and health was chosen because it is an arena where following instructions is often used, which requires algorithmic thinking. Thus, perhaps one can use block programming in mathematics and unplugged programming in food and health to work on students' algorithmic thinking and their understanding of what programming can be.

Therefore, our research question was:

> In what way can two lessons within the subjects of mathematics and food and health make possibilities in programming visible to the students?

2. Materials and methods

In this project we used two types of programming: block programming and unplugged programming. Block programming can be described as a visual representation of programming code, for instance graphic icons or blocks. These blocks can be put together to form a code or an algorithm (Humble et al., 2019). An algorithm is step-by-step instructions that describe how to do something. Humble et al. (2019) describes unplugged programming as programming without using a computer. Board games and controlling each other with commands or written instructions are some examples of unplugged programming.

Another example of unplugged programming is recipes in food and health. A recipe is an example of an algorithm (Berry, 2014), which tells you step by step what you should do to make the desired dish. As mentioned, algorithmic thinking is a part of computational thinking and it consists of following and explaining step-by-step instructions (Gjøvik and Torkildsen, 2019). Therefore, as recipes can be seen as an algorithm, we think it can be used to practice algorithmic thinking.

Ardito et al. (2020) also includes collaboration as a skill within computational thinking, which can also be an element when several people cook together. Wang et al. (2021) emphasize the importance of collaboration when learning computational thinking. These are some of the similarities that can be found between computational thinking and cooking, which made us think that working in food and health can provide an opportunity to increase students understanding of algorithmic thinking and programming and broaden their view of what programming can entail.

Furthermore, Eidslott (2021) argues that the motivation of students who initially have a greater interest in other subjects than science can be increased by combining learning objectives from several subjects and making programming something practical. Thus, perhaps food and health can be used as an introduction to programming for students who are initially not interested in it.

2.1. Study design

Action research can be seen as a small-scale intervention that combines action and reflection on what has been done (Cohen et al., 2018). Furthermore, in action research, the researcher can take an active role in the studied field (Tiller, 2006). In our project, we wanted the students to experience several aspects of programming. As programming is relatively new in the curriculum, we assumed that the students had a limited image of what programming is or could be. Thus, we wanted the students to make use of computational thinking in other areas than digital programming, in the form of unplugged programming. An action research process consists of planning, implementation and evaluation of a scheme, preferably in several rounds (Carr and Kemmis, 1986). In our project, we only carried out the lessons in one round. The reason for this is presented under descriptions of the lessons.

2.2. Participants

Two sixth grade school classes from a regular city school participated in the project. Both classes consisted of 22 students, therefore the total amount of students which participated were 44. Of these students, 36 gave feedback. A small sample size fits with action research. At the same time, it will not be possible for us to generalize the results from this project beyond the project's participants, but hopefully we can draw some inferences from it.

2.3. Lessons

The project consists of two lessons, one within food and health and one in mathematics, where the students participate in both. Both lessons were planned by pre-service teachers in collaboration with university lecturers and were carried out by the pre-service teachers.

To prevent the size of the student groups to be too large, we chose to carry out the project over 2 days where one class visited us at the university each day. Furthermore, each class was divided into two groups. One group started with mathematics, while the other one started with food and health. Halfway through the day the two groups switched places. As this was on consecutive days, we were not able to evaluate and plan and make any major changes from 1 day to the next, which is one of the main ideas of action research. In addition, we did not make any changes on the overall structure of the implementation of the lessons because we wanted the students to have the same experiences so that they had a common starting point for any conversations that took place at a later stage. Furthermore, as all the students did more or less the same, we got a larger number of responses that can give us an indication on the study rather than doing two separate run-troughs and getting half the responses on each.

Below follows a description of the two lessons.

2.3.1. Mathematics

A competence aim after year five in mathematics is student "is expected to be able to create and programme algorithms with the use of variables, conditions and loops" (Ministry of Education and Research, 2019, p. 9). Before the study the students have followed the curriculum for a year and should be familiar with this competence aim and it was plausible that they needed some of the same competence in this lesson. We chose to use block programming, including micro:bit and Bit:Bot, as the students already had some familiarity with this. The students were divided into groups of two or three students in each group, where each group had a micro:bit and Bit:Bot. The groups were given two to three tasks to solve, depending on how quickly they solved the first two tasks. The first task was to get the Bit:Bot to drive one meter forward, turn around 180°and drive back to start. This task gives the students information about how the Bit:Bot moves. For example, how many milliseconds it takes to drive one meter or turn around at the selected percentage of maximum speed. The students could use this information in the next task, which consisted of making the Bit:Bot to drive around a track. The track contained five straight stretches with two 90° turns and two 45° turns. The groups that finished driving the track, could try to make a traffic light. The traffic light consisted of a red, a yellow and a green LED. The traffic light had to be programmed to change to a new colour every 2 s. The tasks were taken from: https://www.vitensenter.no/superbit/elev/superbit-ogsätning/.

2.3.2. Food and health

The students were divided into four groups of three or four students and in the first session the students were organized into groups with the same recipe, while in the second session each group had four different recipes. For the first group, the recipes consisted of counting, brushing, washing and boiling potatoes, preparing trout and butter sauce, making raw carrot salad with dressing and dessert Greek yogurt with strawberries and toasted oatmeal. The other group had to make fish gratin from scratch, with raw carrot salad and potatoes. For dessert the students made fruit salad. In addition, the groups had pre- and post-work in connection the making of the food, such as preparation and cleaning afterward. The unplugged programming consisted of understanding and following the given recipes and working together in groups. Throughout the process, pre-service teachers were available to give guidance to the students on "decoding" the recipes.

2.4. Data collection

To answer the research question, we collected data through observation and a questionnaire. A questionnaire consists of written questions requiring written responses (Kemmis et al., 2014).

The day after the students had participated in the lessons, they answered the questionnaire together with their teacher at their own school. It was voluntary for the students to answer the questionnaire and it was anonymous. We had no way of finding out which students had answered what because it was anonymous and we were not present when they answered the questionnaire. The questionnaire consisted of three open ended questions that they had to answer. Open-ended questions is useful for research on a smaller scale (Cohen et al., 2018). The questions we asked were:

- Describe the programming you did at UiT.
- Describe how you experienced programming in mathematics.
- Describe how you experienced the programming in food and health.

With the first question, we wanted to get the students' overall picture of the day (both lessons). Next, we wanted the students to describe the programming within the two lessons. Open-ended questions could provide answers which might not otherwise have been possible in a questionnaire, and allow the respondents to answer in their own words (Cohen et al., 2018). Therefore, we chose to have open-ended questions where the students had to describe what they did because we wanted to get their own thoughts. At the same time, we did not know what the students would answer. If we had closed questions, possibly with answer categories, we could perhaps get answers to exactly what we wanted. On the other hand, we did not want to steer the students toward any answers. We chose to use the term programming which the students have heard about before, and we did not know whether the students were familiar with what computational thinking was or some of the concepts or components of computational thinking.

As the pre-service teachers had the main responsibility for the implementation of the lessons, we made our observations as non-participants. As there was only one student teacher at the mathematics lesson, the one of us who were there occasionally had to help some groups of students to progress in their work. We chose to make use of unstructured observation, as we wanted to have the opportunity to write down interesting events we captured. During the lessons we observed and subsequently wrote what we observed after the lessons.

2.5. Data analysis

In the planning phase, we had expectations that the pupils would most likely not see the connection between block and analog programming on the basis that LK20 had only been in use for two and these were pupils in the 6th grade. Before the analysis process started, all the student answers were imported into an excel-document, and each question was placed on a separate sheet in this document. This allowed us to analyse one question at a time, while at the same time we had an overview of what each student had answered on every question. The analysis process of qualitative data is often inductive (Cohen et al., 2018). Thus, we started the analysis without predefined categories.

We used thematic analysis following the phases of Braun and Clarke (2006). To familiarize ourselves with the data we read through the answers from the students several times and noting down ideas, patterns and interesting answers. In the second phase we identified initial codes in the responses. Third, we gathered similar codes into larger categories. For questions two and three, answers could often contain more than one code and could fit in more than one category. Fourth we reviewed the categories that had emerged to see if anything had been left out or if it was natural to combine some of them. To get a better overview of the data, all categories that we experienced as consistent was noted in the excel-document which made it possible for us to count and categorize the qualitative answers. Fifth, we settled on names for the categories.

After we had finished the thematic analysis of the questionnaire we discussed the categories against our observations. Through the observation, we had more control over what we saw, but we did not have that in the students' feedback and in that way we experienced that they gave information about each other. This repeated and multi-sided systematization can have an impact on reliability. The results of the analysis were in line with our expectations.

3. Results

In this section, we will present the parts of the data material that can shed light on the problem. In other words, not everything the students have answered will be presented here. Because the students answered the questions in Norwegian, the responses presented here is our translation from Norwegian to English of what the students have answered.

3.1. Students did not recognize programming in food and health

It seems that there is a clear difference to what extent the students think programming was involved in the two lessons. After the completion of these two lessons, it may appear that the students did not link programming to the lesson that was carried out in food and health. On question one where the students were asked to describe the programming, they did on the day they were at the university campus, food and health were only mentioned in four out of 36 answers. In contrast, 34 out of 36 of the answers can be directly linked to the lesson that was carried out in mathematics. In these answers, students wrote that they programmed a Bit:Bot, a car or a robot.

Below are the four answers that mentioned the lesson in food and health.

- *"I programmed and cooked. In the programming we programmed Bit:Bots to follow a line and drive one meter forward, turn and go back one meter."*
- *"We programmed a robot to drive one meter back and forth, we also made it drive on a track and we also made a mini traffic light. We also made fish gratin and fruit salad in food and health."*
- *"It was fun, but a bit challenging sometimes, for example with the traffic lights. But everything was a lot of fun, both food and programming."*
- *"First, I had food and health in the kitchen, and after that I had programming. We programmed the car to go one meter back and forth."*

If you look at these responses, the lesson in food and health are not linked to programming. The students just answered that they cooked or made food. From these four responses it seems like the students think they programmed in one lesson and made food in the other. Furthermore, food and health were never mentioned alone in these answers, but programming in mathematics was also mentioned.

3.2. Some programming in food and health nonetheless

It seems like seven of the students were, nonetheless, able to see the connection between programming and the recipe they used in food and health on the question "Describe how you experienced the programming in food and health." In two of these answers, the students clearly state a connection between the two lessons that were carried out. If seven students could see a connection, this means that 29 students gave answers that did not give any indication that they saw a connection between programming and the lesson in food and health.

In **Table 1**, there is an overview of all the categories on this question where we believe that we can see answers that can be related to programming in one way or another. It is worth adding that none of the students who mentioned food and health in the first question answered anything that could be linked to programming to this question.

The most obvious link to programming can be found in the category *Description of roles*. In the answers in this category, the students have linked the lesson in mathematics together with the lesson in food and health by seeing the similarity between themselves in the kitchen and the car/Bit:Bot. Below are the two answers that ended up in this category.

- *"I'm kind of the micro:bit and the recipe is the MAKER."*
- *"The recipe was what we had to follow and it's a bit like the car as well because in a way it gets a recipe."*

As you can see from the first response above, the student draws a direct link between itself in food and health and the micro:bit (which is in the Bit:Bot), and the recipe and the person who programs the Bit:Bot. We can see the same in the second response, where the student wrote that in the same way that student followed the recipe in food and health, the Bit:Bot also follows a recipe that determines what it will do.

The category *Recipe* contains the answers that, in one way or another, mentions recipe or part of a recipe. In five of the six answers in this category, it seems like the students have realized

TABLE 1 Categories that can be connected to programming on the question: "Describe how you experienced the programming in food and health".

Category	Number of responses
Recipe	6
Description of roles	2
Mentions the word programming	3

that it is the recipe in food and health that can be the link to programming. Below are all the answers in this category.

- *"I think that it is that we are reading a recipe."*
- *"The recipe."*
- *"The recipe! It was actually easy since we didn't use the recipe, since we were told what to do."*
- *"I experienced programming in the recipe."*
- *"I followed a line of code most of the time. I let the potatoes cook for 40 and occasionally look at them."*
- *"I don't know, I just did what was written on the recipe."*

It varies how certain the students seem to be about whether the recipe can be programming or not. For example, the first answer shows that the student is somewhat unsure about this, while others seem more certain. In the fifth answer, the word recipe is not mentioned like the other answers. Here it can seem like the student connects lines of code to the description of how to cook the potatoes. As mentioned earlier, a recipe is step-by-step instructions, and this can be compared to lines of code in an algorithm. This can thus be interpreted as the student linking the recipe in food and health to lines of code and algorithms in programming. Although the recipe is mentioned in the sixth answer, the student expressed that they did not know and just followed the recipe. Thus, it can be interpreted that this student does not see the programming in the recipe.

All answers where the word programming has been mentioned has ended up in the category *Mentions the word programming*. In these responses, the students answered that the programming in food and health was fun or good. Thus, it is not easy to say whether they have seen what the programming can be in food and health. It is worth mentioning that in addition to the categories shown in Table 1, there were also six students who answered: "I don't know" or "I do not understand the question."

3.3. Other interesting categories

Table 2 shows two of the categories that cannot be directly linked to programming, but which nevertheless can be interesting. These categories are *Fun* and *Collaboration*. In the fun category, 25 out of 36 students answered that they thought the mathematics lesson was fun. Correspondingly, 16 of the students wrote that the lesson in food and health was fun.

A total of seven of the responses about the experience of programming in food and health have ended up in the collaboration category, six of these refer to a functioning collaboration and one case where the collaboration has not worked. These answers describes whether the group members worked well as a team and whether everyone did their part of the work. On the question of the experience of programming in mathematics, two answers can be categorized as the collaboration worked well and three where the collaboration has not worked. Here, too, the answers focus on whether each group member had done their part of the job or not.

TABLE 2 Number of responses in two categories that cannot be directly linked to programming to the questions that asked how students experienced the programming in mathematics and food and health.

Question	Fun	Collaboration	
		Good	Bad
Describe how you experienced programming in mathematics	25	2	3
Describe how you experienced the programming in food and health	16	6	1

3.4. Observations

Here we will present some of the things which we observed that can be related to our research question.

In food and health, we saw that the pre-service teachers were active tutors in reading and following the recipes. In the event of a lack of description or ambiguities in the recipe, the pre-service teachers supported the students in their process. An example of this is in the procedure for white sauce, one instruction is: "Melt butter in a large saucepan." The student opened the kitchen cupboard and wondered which pot was big enough and took out the largest pot, whereupon a pre-service teachers guided the student to take a smaller one and which the pre-service teacher considered suitable for the amount of sauce that was going to be made. In food and health, it was also observed that the pre-service teachers did not tell the students they were doing unplugged programming.

The observations of very happy students who worked as if they were highly motivated were very prominent both in mathematics and in food and health. It felt as if the students found the assignments meaningful. Despite much joy, two individual students were observed in food and health who did not participate in the work and one student who finished early with his part of the work and did not help the rest of the group. In mathematics, certain groups or group members were also observed who occasionally opted out and did not participate in the lesson. In one of the cases, it seemed like the group just needed some guidance on how to think in order to figure out how to get the Bit:Bot to follow the track and how to work together.

In mathematics, it was observed that the majority of the groups largely used trial and error as a way of working to solve the tasks. The students made some changes to their code and then tested the Bit:Bot on the track.

4. Discussion

The purpose of this study was to investigate whether two lessons in the subjects mathematics and food and health could highlight possibilities in programming for the students. In particular, whether the students recognized the programming in both subjects.

4.1. Transfer of learning

Our main finding was that none of the students answers to the first question could be linked directly to programming in food and health. This is further supported by the fact that the majority could not clarify what the programming was in question three either, and six students answered that they did not know or understand this question. This may indicate that the majority of the students did not see the connection between programming and food and health.

Zhuang et al. (2020) defines transfer of learning as the result of generalization of experiences. If a person generalizes his experiences, it is possible that he can transfer knowledge from one situation to another. An advantage of transfer of learning between two areas is that it can strengthen learning in the new area you are studying. A prerequisite for transfer of learning is that there must be a link between the two learning activities (Zhuang et al., 2020). The fact that most of the students were unable to see the connection between the lessons may indicate that they were unable to generalize their experiences. One can speculate whether it could be because the students may think of programming as a separate area and not linked to other subjects, and thus they did not see the link between block programming and unplugged programming in food and health. In addition, maybe the students did not consider reading recipes in food and health to be the same as reading algorithms in programming. In food and health, the pre-service teachers filled in gaps in the recipe so that the students succeeded in the process, or as one student wrote that they did not need the recipe as they were told what to do. In contrast, a Bit:Bot will follow the algorithm and steps exactly as the students have set it up, and if something is wrong or missing, compared to the students' intention, the Bit:Bot will not do what the students want.

In contrast, 34 out of 36 answers could be connected to the lesson that was carried out in mathematics. Thus, it can appear that students largely associate programming with block programming, and not unplugged programming. In light of the intention of the school of the future to facilitate students to develop problem-solving skills and critical assessment skills (NOU, 2015: 8), it may appear that students did not automatically express competence in computational thinking. However, we cannot rule out that students are aware of how they think (Ways of Thinking) and work (Ways of Working), which tools they use (Tools for Working) in relation to the world they live in (Living in the World) (Binkley et al., 2012) in this case in mathematics or food and health.

One of the reasons why the majority of students may not connect programming and lesson in food and health may be that the term *unplugged programming* was not used in food and health, at least not that we could observe. We chose not to use the term unplugged programming to avoid influencing the results and we wanted to see if the students themselves would make the connection, this in line with the criticism of action research that the researcher can influence the action. If the term was more actively used, it is conceivable that the result could change, but at the same time we would not have given the students opportunity to discover the connection themselves.

When the students were asked directly about programming in food and health, there were seven who expressed in writing what programming could be. Here, connections were made between algorithm and recipe, and some students saw the similarity between themselves and Bit:Bot. Thus, it may appear that some students nevertheless managed to generalize their experiences so that they saw the connections that we hoped to find in the project, and maybe would be able to transfer some of their knowledge from one situation to the other. The fact that there are only seven students who saw the connection does not necessarily mean that the other students cannot make use of the knowledge they acquired in both block and unplugged programming.

4.2. Programming as an approach to learning

A prominent research result was that 25 out of 36 students experienced programming in mathematics as fun. Our interpretation of fun can show that this arrangement in programming made the students more active participants. According to Jordet (2020), school today is still characterized by students sitting on their chairs doing theoretical work. A more practical school is in line with Dewey's (1915) "Learning by doing," and the importance of stimulating the senses and use the body while learning. In our project the students did not sit still but were in motion. In the kitchen, they used their bodies and senses to prepare the food, while they had to walk between the table they were working on and the track the Bit:Bot was driving on. This can be linked to the intention in "embodied cognition" where the body is in interaction with the brain and the world around (Shapiro, 2019) and to bodily learning which in the OECD report is highlighted as important in pedagogy (Paniagua and Istance, 2018). Embodied cognition can take place in both the digital and the analog space (Østern and Knudsen, 2021) and can be summarized through Vygotskij et al. (1978) socio-cultural perspective where all intellectual development is based on social activity. Based on this, programming can be a learning arena, where the students work together with other students and the pre-service teachers where knowledge is exchanged and contributes to the tasks being solved. Vygotskij et al. (1978) calls it the zone of proximal development where a student can solve problems under guidance or in collaborations with more capable peers. The pedagogical challenge lies in making use of the zone by providing help and support, so that next time the students can manage to do the task on their own.

In both lessons, we hoped to challenge the students to try both unplugged and digital tasks that they had not done before. We tried to arrange for the students to experience two sides of programming and which could hopefully clarify programming and algorithmic thinking. The students explored together with the guidance and encouragement of fellow students and pre-service teachers. We experienced that this created a positive feeling of mastery among most of the students, which may be due to the fact that it was an informal learning situation where students and pre-service teachers developed good relationships with each other. Bruner et al. (1997) emphasizes that human learning is an interactive process where people learn from each other. Jordet (2020) summarizes that "Recognition is, in other words, the most central psychosocial prerequisite for children's academic and social learning at school, for the development of good self-esteem and for good mental health," and highlights three forms of mutual recognition; love, justice and social values where the sum is included in and helps to shape the self-worth or identity of a person. Through the programming lessons where everyone was active in one way or another, an opportunity was created for the pre-service teachers to meet the students with friendly eyes, interest and support and in that way could help to form the student's self-esteem. The students could contribute through their rights and duties and in that way strengthen their self-respect (Jordet, 2020). However, it was observed that some students did not make use of their rights or that the task was too small in relation to the student's capacity to perform his duty. The reason for this may be poor planning of work tasks or supervision of the pre-service teachers. This could prevent the students from showing their skills and competence in a social community and thus may not get recognition (Jordet, 2020). The social interaction may be related to that.

Eidslott (2021) writes that programming is more than just writing codes. He thinks that it is more about that the students, through their creativity and ability to collaborate, can be able to solve problems by getting an overview of the problem and are able to divide it up to smaller problems or tasks, and arrive at a solution through trial and error. Shute et al. (2017) also highlight that collaboration and problem-solving skills such as trial and error as elements of computational thinking. In conjunction with Eidslott's (2021) point that students can use their ability to cooperate to solve challenges, there were six students who described a functioning collaboration in food and health and only one who experienced a malfunctioning collaboration. In mathematics, only two mentioned that the collaboration was good, while there were three who chose to point out that it was bad. As the answers largely focused on the distribution of work in the groups, it may be that the students fulfilled their duty in the collaboration to varying degrees. Of course, we cannot rule out that, in the groups that did not work, there may be other reasons why the collaboration did not work, such as a lack of skills.

In the lesson in mathematics, the most used method of working was trial and error. The students constantly made small adjustments to their code, followed by testing if the Bit:Bot did as they wanted. In contrast, trial and error can result in an undesirable result in food and health. It is always possible to use trial and error in food and health too, but it is not sustainable to throw away ingredients or food due to experimentation in terms of the UN's Sustainable Development Goals (United Nations, 2021), especially goal number 12 about ensuring sustainable consumption and production patterns.

4.3. The connection between unplugged and block programming

There were 34 students that related programming to the lesson in mathematics and only a few to the lesson in food and health. There can be many reasons for this. Firstly, maybe the students did not know how to answer. secondly, it may be due to the students' prior knowledge, they could possibly have encountered digital programming both inside and outside of school. It is not necessarily that they relate programming to mathematics either, but as an independent activity. Based on our understanding that the students acquired different skills through programming, it seems like they did not see the connection between block and unplugged programming. Dewey (1938) and Vygotskij et al. (1978) are known for their philosophy that learning takes place through words. Dewey (1938) sees it in the context of doing, experiencing and reflecting. The students' exploration takes place based on what they already know. They have previously had a visit from *The Science Centre of Northern Norway* which focused on programming, and perhaps some of the students have done some programming in their spare time. They could share experiences with and understanding of programming with the others in the group. This is in line with the intention of the school of the future to educate workers for the future, and that students should acquire competence in learning, communicating, interacting and participating in addition to exploring and creating (NOU, 2015: 8).

One of the goals in the national core curriculum is to find a solution to issues by using approaches from various subjects through interdisciplinary work (Ministry of Education and Research, 2017). Initially, our intention was to create an interdisciplinary collaboration between mathematics and food and health, in order to fulfill this part of the core curriculum where interdisciplinary is emphasized. On the other hand, this project probably cannot be called interdisciplinary, as the two lessons are carried out in parallel and are not closely connected. We planned to have a common theme of programming, but the lessons could have been carried out individually since they do not build on each other. On the other hand, the project may perhaps fall under the concept of multidisciplinary as it's called by Drake and Reid (2020). In both lessons, students work with programming and computational thinking, where mathematics and food and health illuminate this from their respective viewpoints, and the subjects are coordinated, but are carried out separately.

5. Conclusion

In our project, we carried out two lessons with unplugged and block programming in the subjects of mathematics and food and health, where we looked at whether these can contribute to developing computational thinking in students in the 6th grade and whether the students could see the connection between the lessons. As we have a limited sample, we cannot say anything that applies in general, but for the students we have, we can summarize that it may appear that most of the students did not see the link between programming and the lesson in food and health/unplugged programming. But we cannot know for sure whether the students can use computational thinking in areas other than digital programming. The same applies to other possibilities in programming. The students may have acquired skills in, for example, collaboration, communication and problem solving.

6. Implications

This is a small study, almost a pilot, we think the article is relevant for a larger audience because, based on our knowledge, there has not been much research into programming where there is a combination of practical-aesthetic subjects, such as food and health, and mathematics. There are a few reasons why we think this study may be relevant to others. This article may help to make accessible and clarify what programming can be, and perhaps motivate others to work with programming in several arenas and angles. One of the approaches could be to work with in-depth learning through multidisciplinary and different activities. This is an example of a multidisciplinary scheme that can perhaps be developed into an interdisciplinary one. It can show how to strengthen the natural bond between food and health as a practical aesthetic subject and mathematics as a science subject. Perhaps an opportunity can be created for practical aesthetic subjects and science subjects to develop familiarity, understanding and respect for the special nature of the subjects–content and working methods.

If we look a little ahead, we have made a couple of thoughts about what we think might be interesting for us to work on later. One possibility is to further develop our lessons so that it becomes more interdisciplinary, or at least so that the lessons in the two subjects are more closely linked. Hopefully, it can make the programming in food and health more apparent to the students. One way to do this could be for students to a greater extent "code" more themselves in food and health. It could, for example, be that they do some research in order to make their own recipes with precise instructions, such as in an algorithm.

Another possibility is to do something similar again in a few years. As the new mathematics curriculum was introduced in the autumn of 2020 and the students have had a lot of home schooling during the corona pandemic, the students in this study have had limited programming lessons. In a few years, students will have had more programming in the mathematics education and perhaps in several subjects. Thus, it can be interesting to compare the results to, among other things, see if there are changes in the students' approach to programming.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements.

Author contributions

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

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ardito, G., Czerkawski, B., and Scollins, L. (2020). Learning computational thinking together: effects of gender differences in collaborative middle school robotics program. *TechTrends* 64, 373–387. doi: 10.1007/s11528-019-00461-8

Bell, T., and Vahrenhold, J. (2018). "CS unplugged—how is it used, and does it work?," in *Adventures Between Lower Bounds and Higher Altitudes: Essays Dedicated to Juraj Hromkoviè on the Occasion of His 60th Birthday*, eds H.-J. Böckenhauer, D. Komm, and W. Unger (Cham: Springer International Publishing).

Berry, M. (2014). *Does not Compute. Teach Primary*. Colchester: Maze Media.

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., et al. (2012). "Defining twenty-first century skills," in *Assessment and Teaching of 21st Century Skills*, eds P. Griffin, B. Mcgaw, and E. Care (Dordrecht: Springer Netherlands). doi: 10.1007/978-94-007-2324-5_2

Bocconi, S., Chioccariello, A., and Earp, J. (2018). *The Nordic Approach to Introducing Computational Thinking and Programming in Compulsory Education*. Report prepared for the Nordic@BETT2018 Steering Group. doi: 10.17471/54007

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Bruner, J. S., Grøver, V., and Christensen, B. (1997). *Utdanningskultur og laering*. Oslo: Ad notam Gyldendal.

Carr, W., and Kemmis, S. (1986). *Becoming Critical: Education, Knowledge, and Action Research*. London: Falmer Press.

Cohen, L., Manion, L., and Morrison, K. (2018). *Research Methods in Education*. London: Routledge. doi: 10.4324/9781315456539

Dewey, J. (1915). *The School and Society, by John Dewey*. Illinois: The University of Chicago press.

Dewey, J. (1938). *Experience and Education*. New York: Kappa Delta Pi/Touchstone.

Drake, S. M., and Reid, J. L. (2020). 21st century competencies in light of the history of integrated curriculum. *Front. Educ.* 5:122. doi: 10.3389/feduc.2020.00122

Eidslott, H. (2021). *5 Råd for å Komme i Gang med Programmering i Skolen [Online]. Utdanningsnytt.no*. Available Online at: https://www.utdanningsnytt.no/fagartikkel-pedagogikk-programmering/5-rad-for-a-komme-i-gang-med-programmering-i-skolen/282639 (accessed December 15, 2022).

Gjøvik, Ø, and Torkildsen, H. A. (2019). Algoritmisk tenkning. *Tangenten Tidsskrift Matematikkundervisning* 30, 31–37.

Humble, N., Mozelius, P., and Sällvin, L. (2019). "On the role of unplugged programming in K-12 education," in *Proceedings of the European Conference on e-Learning 2019 (ECEL2019)* (Denmark).

Jordet, A. N. (2020). *Anerkjennelse i Skolen: en Forutsetning for Laering*. Oslo: Cappelen Damm akademisk.

Kemmis, S., Mctaggart, R., and Nixon, R. (2014). *The Action Research Planner: Doing Critical Participatory Action Research*. Singapore: Springer. doi: 10.1007/978-981-4560-67-2

Ministry of Education and Research (2017). *Core Curriculum – Values and Principles for Primary and Secondary Education. Established as a Regulation by Royal Decree. Laereplanverket for Kunnskapsløftet 2020*. Oslo: Ministry of Education and Research.

Ministry of Education and Research (2019). *Curriculum for Mathematics Year 1–10 (MAT01-05). Established as Regulations by the Ministry of Education and Research. Laereplanverket for Kunnskapsløftet 2020*. Oslo: Ministry of Education and Research.

NOU (2015). *8. Fremtidens Skole — Fornyelse av Fag og Kompetanser*. Oslo: Ministry of Education and Research.

Nouri, J., Zhang, L., Mannila, L., and Norén, E. (2020). Development of computational thinking, digital competence and 21st century skills when learning programming in K-9. *Educ. Inquiry* 11, 1–17. doi: 10.1080/20004508.2019.1627844

Østern, A.-L., and Knudsen, K. N. (2021). "Kroppslig laering i digitale/analoge rom," in *Kroppslig Laering: Perspektiver og Praksiser*, eds Ø Bjerke, G. Engelsrud, A. G. Sørum, and T. Østern (Oslo: Universitetsforlaget).

Paniagua, A., and Istance, D. (2018). *Teachers as Designers of Learning Environments*. Berlin: OECD Publishing. doi: 10.1787/9789264085374-en

Sevik, K. (2016). *Programmering i Skolen*. Oslo: Senter for IKT i utdanningen.

Shapiro, L. A. (2019). *Embodied Cognition*, 2nd Edn. New York, NY: Taylor & Francis Group. doi: 10.4324/9781315180380

Shute, V. J., Sun, C., and Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educ. Res. Rev.* 22, 142–158. doi: 10.1016/j.edurev.2017.09.003

Tiller, T. (2006). *Aksjonslaering - Forskende Partnerskap i Skolen: Motoren i det nye laeringsløftet*. Kristiansand: Høyskoleforlaget.

United Nations (2021). *The Sustainable Development Goals Report 2021*. New York, NY: United Nations.

Vygotskij, L. S., Cole, M., John-Steiner, V., Scribner, S., and Souberman, E. (1978). *Mind in Society: the Development of Higher Psychological Processes*. Cambridge: Harvard University Press.

Wang, X. C., Choi, Y., Benson, K., Eggleston, C., and Weber, D. (2021). Teacher's role in fostering Preschoolers' computational thinking: an exploratory case study. *Early Educ. Dev.* 32, 26–48. doi: 10.1080/10409289.2020.1759012

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., et al. (2020). A comprehensive survey on transfer learning. *Proc. IEEE* 109, 43–76. doi: 10.1109/JPROC.2020.3004555

frontiers | Frontiers in Education

TYPE Review
PUBLISHED 17 April 2023
DOI 10.3389/feduc.2023.1138607

Check for updates

OPEN ACCESS

EDITED BY
Lisbet Rønningsbakk,
UiT The Arctic University of Norway,
Norway

REVIEWED BY
Danielle Harris,
University of Canberra,
Australia
Clodagh Reid,
Technological University of the
Shannon: Midlands Midwest,
Ireland

*CORRESPONDENCE
Caiwei Zhu
✉ c.zhu-1@tudelft.nl

SPECIALTY SECTION
This article was submitted to
STEM Education,
a section of the journal
Frontiers in Education

RECEIVED 05 January 2023
ACCEPTED 14 March 2023
PUBLISHED 17 April 2023

CITATION
Zhu C, Leung C.O-Y., Lagoudaki E, Velho M,
Segura-Caballero N, Jolles D, Duffy G,
Maresch G, Pagkratidou M and
Klapwijk R (2023) Fostering spatial ability
development in and for authentic STEM
learning.
Front. Educ. 8:1138607.
doi: 10.3389/feduc.2023.1138607

Fostering spatial ability development in and for authentic STEM learning

Caiwei Zhu[1]*, Chloe Oi-Ying Leung[2], Eleni Lagoudaki[3], Mariana Velho[4], Natalia Segura-Caballero[3], Dietsje Jolles[2], Gavin Duffy[4], Günter Maresch[3], Marianna Pagkratidou[4] and Remke Klapwijk[1]

[1]Department of Science Education and Communication, Delft University of Technology, Delft, Netherlands, [2]Institute of Education and Child Studies, Leiden University, Leiden, Netherlands, [3]Department of Mathematics, Paris Lodron University of Salzburg, Salzburg, Austria, [4]School of Electrical and Electronic Engineering, Technological University Dublin, Dublin, Ireland

Empirical interdisciplinary research has explored the role of spatial ability in STEM learning and achievement. While most of this research indicates that fostering spatial thinking in educational contexts has the potential to positively impact students' enrollment and performance in STEM subjects, there is less agreement on the best approach to do so. This article provides an overview of various types of effective spatial interventions and practices in formal or informal educational contexts, including targeted training of STEM-relevant spatial skills, spatialized curricula embedded in schools, integrated STEM practices addressing students' use of spatial skills, and spatial activities in informal STEM education. Gender and socio-economic status of students – two variables that have been found to moderate the relationship between students' spatial ability and their STEM performance – are also discussed in this article. Drawing on a wide spectrum of perspectives on situating spatial ability research in STEM education contexts, this article underscores the need for further inquiry into opportunities for developing K-12 students' spatial ability through integrated and informal STEM practices. This article proposes a conjecture that the relationship between developing students' spatial ability and enhancing their abilities to solve spatially complex STEM problems is bidirectional. Recommendations for future research are made on lingering questions about the effect of interventions, untapped resources for spatial ability training in formal and informal STEM education, and educational strategies for developing students' spatial ability in authentic learning environments.

KEYWORDS

spatial ability, spatial intervention, mathematics education, integrated STEM education, informal STEM education, gender difference, socio-economic status

1. Introduction

Much of society values scientific knowledge and technological development as vital contributors to economic growth, innovation, and welfare (Davies and Horst, 2016; Freeman et al., 2019). Attracting more students to study Science, Technology, Engineering, and Mathematics (STEM) from primary, secondary, and university-levels of education has been on national agendas across the world (Rocard et al., 2007; Joyce and Dzoga, 2011; National Research

Council, 2011; Gough, 2015; Fatourou et al., 2019). Variables such as interest in STEM (Caprile et al., 2015), a sense of belonging in STEM fields (Dortch and Patel, 2017; Murphy et al., 2020), and self-efficacy in learning STEM (Lent et al., 2010; Tracey, 2010) all influence students' decisions to follow STEM studies and careers, and have been the focus of efforts to encourage students to enter the STEM field (Kearney, 2011; Caprile et al., 2015). Surprisingly, initiatives promoting STEM education have paid relatively little attention to cognitive factors such as the spatial ability levels of students, which appears to be a key predictor of future academic and professional involvement in the STEM field (Shea et al., 2001; Wai et al., 2009; Kell et al., 2013).

Spatial ability refers to the competence in representing and processing the location of objects, their shape, their relation to each other, and the orbits they take as they move (Newcombe, 2010); and can be developed through training and education (Uttal et al., 2013a). In a meta-analysis of over 200 studies, Uttal et al. (2013a) found that training students to think spatially and apply spatial skills led to significant increases in their spatial ability, with an average effect size of 0.47. Furthermore, spatial training interventions can potentially lead to transfer of gains to other domains, such as mathematics (Gilligan et al., 2020). Although most research on spatial ability and STEM performance has been situated in college education (e.g., Lord, 1987; Sorby, 2009; Hegarty, 2014), it is critical to make use of the formal and informal resources from early childhood on (Newcombe, 2010; Newcombe and Frick, 2010; Hawes et al., 2017). In fact, an increasing amount of research resources is being devoted to developing spatial interventions or spatialized curricula for K-12 education (e.g., Casey et al., 2008; Burte et al., 2017; Hawes et al., 2017; Lowrie et al., 2017; Sorby and Veurink, 2019).

Through this paper, we aim to characterize a general state of knowledge on the various types of practices that can foster spatial ability development in K-12 educational contexts, drawing insights from different disciplines, including cognitive psychology (e.g., Uttal et al., 2013a; Hawes et al., 2017; Mix, 2019; Gilligan et al., 2020), educational science (e.g., Sorby, 2009; Newcombe and Frick, 2010; Casey et al., 2011; Lowrie et al., 2019), and STEM education (e.g., Burte et al., 2017; Ramey and Uttal, 2017; Julià and Antolì, 2018; Atit et al., 2020). We specifically examine how educational spatial interventions and practices have been situated in authentic, formal and informal STEM environments. By authentic learning environments we mean "situating learning tasks in the context of future use" and in "real-world situations" (Herrington et al., 2014, p. 401). We begin with a section outlining spatial ability and its relation to STEM learning from a cognitive psychology perspective. In this section, we draw heavily on the literature about space-math associations, as math is the first STEM subject officially taught in school and fundamental to most other STEM subjects, and is also the subject that has been researched most extensively in the field of spatial cognition. We then inspect the different approaches used to develop students' spatial ability, covering a wide range of interventions and educational practices, including targeted training of STEM-relevant spatial skills, spatialized curricula, integrated STEM practices that address students' use of spatial skills, and spatial activities in informal STEM education. We focus on if and how integrated STEM practices and informal STEM learning can be harnessed and further developed to stimulate spatial skills development, as they present opportunities for spatial training in authentic learning environments. Lastly, as the goals of spatial interventions or spatialized educational practices always point to developing individual students' spatial ability, we find it important to also discuss how individual variables, such as gender and socio-economic status, may moderate the relationship between students' spatial ability development and STEM learning. We discuss how future interventions and educational practices aiming to develop students' spatial abilities in STEM education contexts may benefit from taking these two variables into account.

2. The role of spatial ability in STEM learning

2.1. Understanding spatial ability: Typology of spatial skills

A clear classification and precise description of spatial skills are important for the development and evaluation of interventions (Buckley et al., 2018), as it provides a basis for systematic exploration of whether, how, and why some interventions may or may not be effective for training specific spatial skills. To create classifications, several researchers have followed a psychometric approach. Linn and Petersen (1985) distinguished between three categories of spatial skills, i.e., mental rotation, spatial perception, and spatial visualization. Mental rotation refers to the ability to mentally manipulate and rotate two- or three-dimensional objects; spatial perception requires individuals to ignore distracting information and determine the spatial relationship regarding their own orientation, e.g., perceiving the water level when the container is tilted; spatial visualization refers to the ability to carry out multistep manipulations of spatial information, e.g., visualizing a piece of paper being folded. Others have argued that this typology is too ambiguous and too broad to categorize tasks that may not be conceptually related (Voyer et al., 1995). The extended Cattell-Horn-Carroll framework (McGrew, 2009) suggested 11 spatial factors based on empirical support, including visualization, spatial relation, closure speed, flexibility of closure visual memory, spatial scanning, serial perceptual integration, length estimation, perceptual illusions, perceptual alternation, and imagery. Despite the fact that more spatial factors have been identified, there are discrepancies between their conceptualizations. A lack of theory-driven classification may be one of the reasons for these discrepancies (Uttal et al., 2013a).

In response to the lack of consensus regarding the definition and classification of spatial ability, Uttal et al. (2013a) adopted a classification system that grew out of linguistics, cognitive science, and neuroscience to distinguish different spatial abilities along two dimensions: intrinsic-extrinsic and static-dynamic (Palmer, 1978; Talmy, 2000; Chatterjee, 2008). The intrinsic-extrinsic dimension is classified according to whether the spatial information is within a single object or between multiple objects, while the static-dynamic dimension is classified according to whether transformation or movement is involved (Uttal et al., 2013a; Newcombe and Shipley, 2014). A two-by-two classification of spatial skills by combining these two fundamental dimensions renders four distinct sub-domains: intrinsic-static, intrinsic-dynamic, extrinsic-static, and extrinsic-dynamic, as shown in Figure 1. This classification cleverly reflects different components of spatial skills that have been described in previous research, such as the above-mentioned categories by Linn

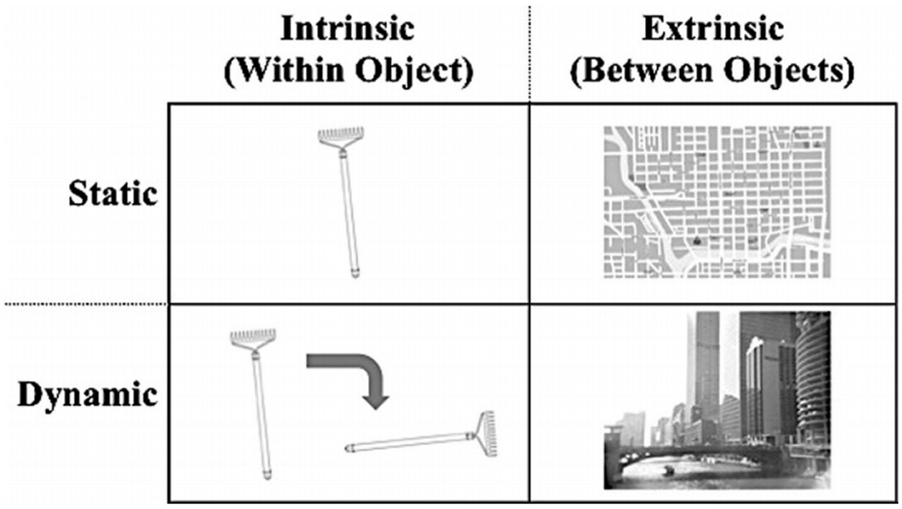

FIGURE 1
A two-by-two classification of spatial skills proposed by Uttal et al. (2013a). Reprinted with permission.

and Petersen (1985), and the additional categories, visuo-spatial perceptual speed and spatial relations, mentioned by Carroll (1993).

The two-by-two classification has been adopted and supported by multiple studies (e.g., Taylor and Hutton, 2013; Gilligan et al., 2018; Hodgkiss et al., 2018). Nevertheless, there was also a debate on whether spatial skill is unitary or multidimensional (Mix et al., 2016; Carroll, (1993)). Based on the two-by-two classification, Mix et al. (2018) conducted factor analyses on kindergarteners, third and sixth graders. The result indicated that the two-factor (intrinsic-extrinsic) model was the best fit for both kindergarteners and third graders, while the one-factor model fits better for sixth graders. This result suggested that the two-by-two framework, especially the dynamic-static dimension, may not accurately classify the latent structure underlying elementary school children's spatial performance.

2.2. Relationship between spatial ability and STEM performance

Several studies have described a relationship between spatial ability and STEM learning. For example, in the domain of math, positive correlations between spatial ability and math performance have been found among preschoolers (Kyttälä et al., 2003; Rasmussen and Bisanz, 2005), middle childhood (Casey et al., 2001; Geary et al., 2007), preadolescence and young adulthood (Shea et al., 2001; Webb et al., 2007), suggesting that people who perform better in spatial tasks also achieve better overall mathematics performance. The correlation between spatial ability and math involves different subdomains of math, including subdomains that, superficially, do not appear to be spatial (Mix et al., 2016, 2018). For example, positive correlations have been found with geometry and word problem solving (Delgado and Prieto, 2004; Kyttälä and Lehto, 2008), as well as with arithmetic (Reuhkala, 2001) and early numeracy (Kyttälä et al., 2003; Lefevre et al., 2010). Besides math, spatial ability also correlates with performance in other STEM domains, including engineering (Sorby, 2009), geoscience (Atit et al., 2015), chemistry (Stieff et al., 2012), computer programming (Jones and Burnett, 2008), design (Lin, 2016), and medical studies (Hegarty et al., 2007).

Although many studies support the relationship between spatial ability and STEM performance, it is still unclear why they are associated, making it difficult to predict to what extent improvement of spatial ability would directly benefit STEM performance. Drawing from the literature on space-math associations (e.g., Mix, 2019; Hawes and Ansari, 2020; Hawes et al., 2022), it is likely that there are multiple ways in which spatial cognition and STEM learning are related. For example, in the domain of arithmetic, Mix (2019) theorized that spatial ability may allow individuals to decode the spatial arrangement of symbols in mathematical equations (e.g., differentiate 15 from 51). It further allows individuals to form mental representations of math problems (e.g., imagine the scene of a math problem), which is a critical step to successful problem solving (Duffy et al., 2020). Lastly, the mental number line may serve as the spatial representation that helps children solve arithmetic problems (e.g., locate the first addend and count up the number of spaces for the second addend). Thus, children with strong spatial skills may find it easier to understand and implement these skills.

Hawes and Ansari (2020) proposed four explanatory accounts to explain why spatial ability and math learning are related. The *spatial representation of number account, spatial modeling account, shared neural processing account, and working memory account*. The first two accounts are similar to what Mix (2019) proposed, emphasizing the role of the number line and spatial representation. The *shared neural processing account* argues that spatial and math are related because they rely on the same brain regions. The *working memory account* assumes that individual differences in visual–spatial working memory are responsible for processing the short-term storage of visual and spatial information (Baddeley, 1993), therefore explaining the relation between spatial and math.

The above theoretical accounts highlighted potential mechanisms underlying the space-math associations. Besides, Gunderson et al. (2012) demonstrated from two longitudinal studies that children's mental transformation ability at age 5 significantly predicted their approximate calculation performance at age 8 [$\beta = 0.34$, t (151) = 2.27, $p < 0.05$], with their numerical knowledge measured by the number line estimation task at age 6 mediating this relation [$\beta = 0.51$, t (151) = 3.26, $p = 0.001$]. These empirical findings suggested that mental transformation ability may help children create a meaningful numerical representation, which allows them to better comprehend the linear number line, similar to what the above-mentioned *spatial representation of number account* suggested.

However, many questions remain, such as why and under what conditions would these mechanisms work and no causal relation could be drawn without an experimental study. Future studies are needed to examine the above-mentioned mechanisms empirically in the area of math as well as in STEM. For example, the spatial representation of number account and spatial modeling account could be tested by providing children with spatial training to see if they achieve a better estimation of the number line or are more likely to form accurate and complete schematic representations in math problems.

3. Developing students' spatial ability through STEM education

3.1. Spatial interventions and transfer to STEM performance

Empirical studies have shown that spatial skills are malleable (Uttal et al., 2013a). This malleability on the one hand and the well-established connection between spatial and STEM skills (Wai et al., 2009; Hodgkiss et al., 2018; Young et al., 2018) on the other, raise the question of whether training spatial skills will improve STEM performance. A recent meta-analysis on studies examining transfer of spatial training to mathematical skills (Hawes et al., 2022) showed from 29 studies that spatial training can have a positive effect on both spatial skills and mathematics performance. Specifically, the meta-analysis showed that spatial training improved spatial skills with a moderate average effect size [Hedges's $g = 0.49$] and enhanced mathematics performance with a small to moderate average effect size [$g = 0.28$]. However, the effect varied depending on age, type of material, and type of transfer. Given the divergence of the outcomes regarding the effectiveness of spatial training on mathematics, the present section attempts to provide an overview of studies that can suggest optimal results in the educational settings.

The section below presents spatial training programs that show transfer to STEM performance in formal K-12 education, specifically in mathematics, from various countries using various methods and factors. We organized the section based on two training approaches: "spatial interventions" and "spatializing the curriculum interventions." Spatial interventions target specific spatial skills, such as spatial visualization and mental rotation, using physical, digital, or hybrid materials and then examine the extent to which mathematics performance is improved. Whereas, spatializing the curriculum interventions seek to enrich the instructions in a classroom with spatial elements. We conclude this section by discussing how these interventions make use of formats or materials that aim to mimic authentic learning experiences.

3.2. Spatial interventions that show transfer to mathematics

3.2.1. Physical format

A study by Burte et al. (2017) is an example of using hands-on materials to improve third to sixth-grade students' spatial thinking in the U.S. with the objective to support their mathematical performance. More specifically, the aim of the program is for students to engage in practices that are relevant to mathematical concepts (e.g., dividing a paper into fractional parts) and scientific reasoning (making sense of a diagram). The hands-on materials they used were origami and pop-up paper, part of an engineering-based program called "Think3D!." The training included paper-folding and -cutting tasks for students to construct three-dimensional objects, interpret diagrams and solve real-life problems. The intermediate steps of this activity are similar to the Mental Paper Folding task and as such, allow students to train their mental rotation skills in an embodied manner. The construction activities nudged students to use spatial language, which is commonly used in math learning, such as *angle, corner, direction, line, position, shape, side, and symmetry*. These practices may also be used to spatialize the mathematics instruction, a spatial training approach we will be discussing in section 3.1.2. Results from this study revealed an interaction between grade and improved accuracy in visual representation problems [$F(3, 77) = 3.38$, $p < 0.05$, $\eta^2 = 0.04$] and abstract math problem-solving [$F(3, 79) = 5.11$, $p < 0.01$, $\eta^2 = 0.06$], with fifth and sixth graders improving but not third and fourth graders. An interaction was also found between grade and math problems that required spatial thinking [$F(3, 79) = 10.35$, $p < 0.001$, $\eta^2 = 0.11$], with only students at higher grades showing an improvement. These findings suggested that spatial and mathematical gains from this embodied spatial intervention may be moderated by the age and the developmental stage of students. We argue that although this study indicated a potential association between the effectiveness of embodied training and students' developmental stage (Burte et al., 2017), this conclusion is tentative and requires further research, including studies with a control group, or an examination of whether the activities used in the training were appropriate for younger students.

3.2.2. Digital format

In a different type of study, Gilligan et al. (2020) delivered computer-based spatial training on mental rotation or spatial scaling to 8-year-old students in the UK, following a randomized, controlled, pre-post training design. Students in the control group were trained in word reading. Students in the training groups (mental rotation; spatial scaling; word reading) were trained either by implicit or explicit video instruction within the school environment. Explicit-instruction-based training consisted of videos that instructed students on how to solve either mental rotation, spatial scaling, or word reading tasks. In the implicit-instruction-based training, students chose the answer they believed to be correct without receiving any guidance. Instead, these students received feedback on whether their answer was correct or not. Findings from this study are insightful as the brief, 3- to 6-min spatial training resulted in gains in the particular spatial skills that were trained (near transfer), the untrained spatial skills (intermediate transfer), and in mathematical tasks as well (far transfer). As far as the mathematical gains are concerned, students who received spatial scaling training significantly improved their performance on the number line estimation task [$t(79) = 2.12$, $p = 0.037$, $d = 0.236$]. Gilligan and colleagues suggested that this transfer is potentially due to the proportional reasoning demanded in both tasks. Students who received mental rotation training improved significantly on missing term problems [$t(69) = 2.73$, $p = 0.008$, $d = 0.241$], regardless of whether the instruction was explicit or implicit. This finding was in line with the results from Cheng and Mix (2014), who found that mental rotation training improved 6- to-8-year-old students' performance at missing term tasks [$t(30) = 2.79$, $p = 0.005$]. However, in the study by Hawes et al. (2015), implicit mental rotation training failed to replicate Cheng and Mix's (2014) results. In regard to why outcomes have not been consistent in these three studies, Gilligan and colleagues hypothesized that the common feature of a part-whole type of mental rotation training[1] between the two studies may have been the catalytic factor for success. Another insight from this study is the "direct causal effect of spatial skills on mathematical performance," without the authors ruling out the possibility of a "bidirectional relationship" between the two (Gilligan et al., 2020, p. 16), which requires further investigation. Meanwhile, computer-based spatial training like this provides a promising practice of spatial training both in terms of its cognitive outcomes and its feasibility, as it can be integrated into the math classroom by having students work on simple spatial activities on electronic devices.

3.2.3. Hybrid format

The classroom-based spatial training by Lowrie et al. (2019) combined the use of hands-on manipulatives with the extensive use of digital instruments. This three-week intervention was aimed at training spatial visualization of 10- to 12-year-old students in Australia. The intervention followed the Experience-Language-Pictorial-Symbolic-Application (ELPSA) pedagogical framework (see Lowrie and Patahuddin, 2015), emphasizing the Pictorial component. "Pictorial" stands for problem visualization, making sense of visual representations, and checking predictions after physically manipulating stimuli. In addition to using hands-on resources, this intervention included tasks in which students used digital applications to create symmetrical objects, predict how nets can form a cube, and explore which two-dimensional shapes can be formed from cutting three-dimensional objects. Compared to the control group, students who received the spatial training showed a moderate improvement in their mathematics performance [$t(17) = 6.95$, $p = 0.016$, $d = 0.39$], especially on geometry tasks [$t(17) = 5.92$, $p = 0.025$], and word problems [$t(17) = 6.11$, $p = 0.023$]. Importantly, this intervention can be conveniently implemented by teachers in their daily instruction after 10 hours of professional development training.

To conclude, studies have now shown that spatial skills can be trained, and that training may lead to enhanced performance in STEM subjects (e.g., Uttal et al., 2013a; Hawes et al., 2022). As it is currently unclear what the underlying mechanisms responsible for training and transfer effects are, it is still difficult to evaluate the effectiveness of spatial training. There are multiple ways in which spatial cognition and STEM performance might be related, such as the spatial representation of numbers account discussed in section 2.2, and the format and materials used in STEM tasks discussed in this section. Further research is needed to study which of those (or other) factors should be the primary target of intervention studies. Moreover, the occurrence of transfer irrespective of the length of training and the amount of spatial gains (Hawes et al., 2022) may suggest that transfer is not necessarily achieved through enhanced spatial skills or refined representations. Instead, spatial training may exert its effects by nudging students to use a spatial problem solving approach during STEM tasks. This hypothesis needs to be tested in future research. Research is also needed to compare the effectiveness of different types of spatial training and examine contextual and individual factors that facilitate training effects. In addition, the meta-analysis by Hawes et al. (2022) suggested that training gain and transfer of spatial gains to mathematics seem to be enhanced when hands-on materials are used, as opposed to computerized materials. Together, these findings suggest that efforts to reinforce the connection between spatial and STEM tasks may help to further stimulate transfer, particularly if hands-on materials are used.

3.2.4. Spatialized curriculum interventions

One way to enhance the connection between spatial cognition and STEM learning is to "spatialize the curriculum." This approach has the potential to not only develop students' spatial skills but also allow students to directly apply such skills while learning subject knowledge like mathematics. An additional advantage of spatializing the curriculum is to enhance the already existing instructions with spatial components, such as the use of spatial gestures and spatial language, without adding extra hours to the school schedule (Newcombe, 2017). Meanwhile, in order to spatialize the curriculum, new curriculum materials have to be designed and teachers need to be trained in order to spatially enhance their instruction (Newcombe, 2017).

An example where this approach was taken is the intervention by Hawes et al. (2017). This intervention focused on spatial visualization with training integrated into the routine for kindergarten to K-2 (4- to 7-year-old students in Canada) mathematics teaching during a thirty-two-week period. The researchers, using a design research paradigm, collaborated with teachers in order to develop and field test teaching plans that would provide opportunities for students to learn early geometry in a dynamic spatial manner. Moreover, these teaching plans incorporated activities for the students to gradually develop their spatial visualization skills. The goal of this intervention was to not teach geometry in a static way but provide the opportunity for students to realize geometric shapes through their spatial transformations. Through playful inquiry, students explored the possible two or three-dimensional configurations of five square tiles and cubes respectively, compared and studied the produced objects, and dynamically practiced axial reflection symmetry and area measurement all using hands-on material (see Figure 2. for an example). Compared to the control class, students in the intervention classes showed a greater improvement in spatial language, visual–spatial geometry, two-dimensional mental rotation, and in symbolic number comparison by the end of the academic school year. Another positive aspect of this particular effort to spatialize early geometry teaching is the high level of engagement demonstrated by both students and teachers.

3.2.5. Promising results: Where do the next steps lie?

In this section, we presented various modes of spatial training delivery that not only helped students improve their spatial abilities

1 Part-whole type of mental rotation means rotating an object and combining it with another element in order to produce a whole.

FIGURE 2
Example of a hands-on activity requiring visualization and mental rotation skills within a spatialized math curriculum (Hawes et al., 2017). Reprinted with permission.

but also showed transfer to STEM learning, such as mathematics. However, research examining transfer of spatial interventions to performance in STEM subjects is still relatively new (Hawes et al., 2022). As a majority of the research on transfer has been centered around the transfer to mathematics, transfer from spatial interventions to performance in other STEM subjects, such as science, engineering, and technology, warrants further investigation (Uttal and Cohen, 2012; Stieff and Uttal, 2015; Margulieux, 2020).

Spatial interventions employing hands-on materials, computer-based instruction with feedback, or a mix of both, as well as the efforts to spatialize STEM curricula, all indicate a shift of focus from developing students' spatial ability through paper and pencil task training to training that closely resembles how students learn together in classroom environments. Whether it is the incorporation of physical and virtual manipulatives, instructional feedback, or playful inquiry, we can see that spatial interventions developed in recent decades are becoming increasingly intentional in mimicking authentic learning experiences. In the following sections, we continue to explore opportunities to develop spatial ability in authentic learning environments.

3.3. Developing students' spatial ability through integrated STEM education

Looking at STEM education holistically, there is a lack of consensus on what STEM education actually entails, as the definition of STEM varies greatly across contexts and among various stakeholders such as researchers, teachers, and school administrators (Brown et al., 2011). Given this variability, it is important that we clarify the kind of STEM education we are focusing on in this section.

Instead of describing STEM as merely a cluster of science, technology, engineering, and mathematics subjects, the term, integrated STEM education, has gained increasing recognition from educational researchers in North America, Europe, and Asia (e.g., Stohlmann et al., 2012; Bryan et al., 2015; Mustafa et al., 2016; Thibaut et al., 2018). Kelley and Knowles (2016) defined integrated STEM education as "the approach to teaching the STEM content of two or more STEM domains, bound by STEM practices within an authentic context for the purpose of connecting these subjects to enhance student learning" (p. 3). Carefully designed integrated STEM education can lead to a range of desirable learning outcomes, such as increased knowledge in STEM (Kelley et al., 2022), improved test performance in subjects like math and science (Tillman et al., 2014), enhanced problem-solving skills (Netwong, 2018), a more positive attitude toward STEM (Sisman et al., 2021), and higher levels of engagement with STEM courses (Taylor and Hutton, 2013; Peng and Sollervall, 2014). Given that STEM education is, at its core, an intersection of different disciplines, we believe it is valuable to discuss where and how researchers and practitioners of integrated STEM may support students' spatial ability development.

To enhance students' spatial ability in authentic learning contexts, we need to prepare students with spatial knowledge and spatial thinking skills that address the challenges they face in different disciplines (e.g., Hinze et al., 2014; Stieff et al., 2014; Atit et al., 2020). Instead of taking the learning of spatial skills as a separate domain, the National Research Council in the U.S. recommended viewing it as a "missing link" across different subject knowledge that "permeates" different disciplines (2006, p. 7). Thus, understanding how to develop students' spatial thinking skills in integrated STEM education not only aligns with the goal of developing spatial skills with content knowledge but also renders a possible answer to how spatial thinking may tie the learning in different STEM subjects integratively.

In the conceptual framework for integrated STEM education, Kelley and Knowles (2016) highlighted four learning approaches that well accommodate integrated STEM learning, including engineering design, technological literacy, scientific inquiry, and mathematical thinking. What ground the four approaches are the shared practices across disciplines and the situated context for learning, so that "learning is authentic and relevant, therefore representative of an experience found in actual STEM practice" (Kelley and Knowles, 2016, p. 4). In the following subsections, we present an overview of how existing spatial interventions or spatialized educational practices leverage each of the four integrated STEM education approaches, and how integrated STEM problems that originate from real-world problems can be used to challenge students' spatial skills development.

3.3.1. Engaging students in spatial thinking through engineering design

Engineering design projects offer a promising platform to weave knowledge and skills needed for different disciplines together while challenging students' inquiry, analytical, and problem-solving skills (Kelley and Knowles, 2016). Several interventions that target students' spatial ability development have been delivered in the form of engineering design projects (e.g., Taylor and Hutton, 2013; Ramey and Uttal, 2017). For example, a programming robot intervention for grade 4th to 6th students led to an increase in a range of spatial reasoning skills such as building 3D objects from pictures and shape rotation (Francis et al., 2021). Francis and colleagues reported significant improvement in the spatial task performance from both the week-long short-term group (e.g., in building 3D objects from pictures, [$t(36) = -2.9$, $p < 0.05$]) and the one-year, long-term group (e.g., in shape rotation, [$t(47) = -3.0$, $p < 0.05$]).

Julià and Antolì (2018) developed a multidisciplinary STEM course on robotics to develop students' spatial skills and mechanical reasoning skills. This year-long course for sixth and seventh-grade students used educational robotics kits to promote hands-on, collaborative learning. Students were given the chance to solve mechanical modeling and building problems in groups, which activated their knowledge and skills related to engineering and technology. The classroom teacher observed that the course also helped students to learn concepts related to science, engineering, and technology in addition to mathematics. To gain a comprehensive view of students' spatial ability, this study tested students with a variety of spatial tasks before and after the course. Students' pre- and post-test scores showed that the robotics course led to significant increases in perspective-taking spatial orientation task performance of sixth-grade students [$t(24) = -2.16$, $p = 0.0356$], and seventh-grade students [$t(23) = -2.04$, $p = 0.0471$]. Other spatial tests, such as the card rotation test and the paper folding test, as well as the mechanical reasoning skills test, yielded promising, close to significant results in terms of improvement from pre- to post-test. Despite the small sample size, this study shows how a STEM course that is integrative in its nature has the potential to develop both spatial skills and STEM-related reasoning skills among students.

3.3.2. Engaging students in spatial thinking through technological literacy

Technology, as Kelley and Knowles phrased, is not just an important tool or vehicle through which students learn about science, engineering, and mathematics, but is itself a discipline consisting of knowledge and practices involved in "designing, making, and using" of technology (p. 6). As the use of technology is ubiquitous across various disciplines, opportunities exist to spatialize technology-focused or technology-enhanced learning experiences. For example, Bhaduri et al. (2021) conducted a pilot study to teach seventh and eighth-grade students 3D modeling and 3D printing with the goal of understanding and supporting students' spatial thinking during computer-aided designs. As students solved an authentic engineering design problem that required the making of a 3D model prosthetic for animals with disabilities, they developed technological literacy in computer-aided modeling tools as well as 3D printing technology. Meanwhile, students exercised their spatial skills such as mental rotation, perspective-changing, and forming mental models of their digital designs.

In another example, Peng and Sollervall (2014) developed a technology-supported, outdoor learning activity that engaged students in learning mathematics concepts, solving a real-world mathematical problem, and exercising their spatial orientation skills. Using a mobile application that informed students of the relative distance between students themselves and the physical markers on the field, the sixth-grade students oriented themselves in the outdoor environment, continuously estimated and calculated distances, and tested out different spatial orientation and coordination strategies in order to solve the math problem.

3.3.3. Engaging students in spatial thinking through scientific inquiry

Learning science through inquiry is essential to prepare students for scientific investigations. Existing inquiry-based educational practices have shown an effort to address students' spatial ability development. As an illustration, during 6 months of geometry instruction that aimed to engage students in mathematical inquiry, secondary students were able to formulate questions that reflected thoughtful mathematical reasoning and spatial reasoning (Lehrer et al., 2013).

Understanding and creating scientific diagrams also demand spatial thinking (Newcombe, 2010). In secondary chemistry classrooms, Stieff (2011) investigated the use of computer-based visualization in a guided-inquiry curriculum to develop students' competence in working with scientific representations. By exploring and visualizing the properties of different chemical substances and the dynamic chemical processes, students in the visualization-focused curriculum developed a better understanding of the content knowledge than those who received traditional lecture instruction, and also developed more competency in creating scientific representations like chemistry professionals do.

For another example, Oberle (2020) recorded sixth, seventh, and eighth-grade students' learning through the National Geographic Geo-Inquiry Process, a curriculum that presents rich opportunities for spatial thinking such as comprehending geographic representations and creating geographic representations using data or maps. While the researcher did not explicitly focus on the gains of spatial ability among the students, the learning of geographic skills often demands multiple types of spatial thinking, such as spatial orientation, spatial visualization, and complex spatial reasoning (National Research Council, 2006). Compared to the control group who received a traditional, non-inquiry-focused curriculum, those who participated in the Geo-Inquiry class showed modest improvement in tasks such as discussing spatial patterns at different scales and elaborating on spatial patterns using maps.

3.3.4. Engaging students in spatial thinking through mathematical thinking

Cohrssen et al. (2017) and Cohrssen and Pearn (2021) incorporated project-based learning, which is another important way to practice integrated STEM (Ritz and Fan, 2015), to teach soon-to-be elementary children about spatial thinking and consequently support their mathematical thinking. Children who were facing a transition from kindergarten to elementary school were challenged to learn the route to their new school and visually represent their route through map-making. The core and complementary activities provided ample opportunities for children to actively reason about spatial orientation and spatial visualization. Moreover, children were encouraged to use directional and locational language and other symbolic representations such as gesturing and sketching to represent their thinking. Taking a qualitative lens, this study allowed researchers to dive deep into how children's spatial thinking has been developed during these activities. For example, providing children with the vocabulary to describe positions, directions, and 2D and 3D shapes, together with prompting children's use of such words through questions, led to fruitful spatial and mathematical conversations between children and teachers. In addition to exercising navigation skills, using spatial memory, and making spatial representations, these children were also actively using their mathematics knowledge during the project-based learning experience.

3.3.5. Designing spatial interventions and spatial practices in integrated STEM education

The studies mentioned above situated students' development of spatial skills in the learning of multiple subject areas, resonating with the suggestion that the use of spatial skills to solve STEM problems is often context-dependent (e.g., Hegarty et al., 2007; Ormand et al., 2014; Atit et al., 2020). Whether it is using integrated STEM curricula as channels to develop students' spatial ability or training students' spatial skills alongside other academic and cognitive skills, integrated STEM education holds promise for developing both spatial and subject knowledge in different disciplines. Meanwhile, these studies presented students with real-world STEM problems to ensure an authentic learning experience. Such organic educational settings accommodate the pursuit of improving students' abilities to solve spatially-complex STEM problems, which potentially leads to transfer of desirable learning outcomes in STEM disciplines (e.g., Hawes et al., 2017). It is worth mentioning that some studies fall into more than one category of the integrated STEM approach, for example, addressing both engineering design and technology literacy. This in turn reveals the nature of integrated STEM, where ideally the four approaches (engineering design, technological literacy, scientific inquiry, and mathematical thinking) will leverage and support each other.

Many of the interventions, case studies, or design research discussed above adopted a mixed-method research approach in which researchers gathered both quantitative data, such as changes in students' spatial scores, and qualitative data, such as classroom observations and in-depth interviews. This approach allows researchers to comprehensively portray how students approached various spatial problems or the difficulties they faced when thinking spatially. The valuable pieces of information from qualitative data are often unattainable through purely quantitative data, and they might offer evidence on why and how the educational approaches lead to spatial learning outcomes.

While spatial ability studies often rely on the results of students' spatial tests, existing spatial ability tests might not fully reflect the spatial skills required in STEM courses. For example, Julià and Antolì (2018) noted that the spatial skills practiced in their educational robotics STEM course did not align well with the tasks in the spatial ability test they employed. In addition, many widely used psychometric measures of spatial ability "do not capture all of the spatial skills required to solve STEM-specific spatial problems" (Atit et al., 2020, p. 2). An example is the penetrative thinking skill, which means "visualizing spatial relations inside an object" (p. 147) and is critical in geology, where students need to understand, for example, the microstructures in minerals (Ormand et al., 2014). However, this skill is fundamentally different from some of the commonly assessed spatial skills, such as visualization or mental rotation (Ormand et al., 2014). Employing a suite of multiple psychometric measures may be one solution (Ormand et al., 2014; Schneider and McGrew, 2018). On the other hand, developing more comprehensive spatial ability measures is urgently needed (Schneider and McGrew, 2018). Future research implementing spatial intervention and spatialized educational practices through integrated STEM education needs to carefully consider how to evaluate students' spatial skills development in classrooms with the context of educational needs in mind.

3.4. Spatial ability development in informal STEM education settings

3.4.1. Informal STEM education

Much STEM learning takes place outside of school (Falk and Dierking, 2010). When out-of-school-time science activities are voluntary and intentionally designed not to be a part of a school's curriculum, they are referred to as informal STEM education. This includes "the act of delivering STEM content outside of the traditional student/teacher relationship to STEM stakeholders (students, parents, teachers, among others) in order to support and increase the understanding, awareness, and interest in STEM disciplines" (Tillinghast et al., 2020, p. 10). These informal settings provide a unique platform to reach students of all grade levels and ages, using different delivery methods (e.g., lectures, active learning, problem-based learning, workshops, camps, events), communicating different scientific areas (e.g., biology, chemistry, engineering, physics), within different settings (e.g., social media, schools, museums, cafés).

While not as often discussed as formal education, informal learning environments have been nudging students from kindergarten to 12th grade to delve deeper into STEM concepts or ideas they may or may not have experienced in their traditional school setting (Mohr-Schroeder et al., 2014). By using hands-on activities that emphasize embodied learning and creativity, informal STEM learning programs have been shown to increase students' confidence and interest in STEM topics, especially for students from underrepresented backgrounds or with limited access to STEM resources (Boone et al., 2020). Extra-curricular STEM involvement has been shown to predict sixth to 12th grade girls' interest and confidence in mathematics (Heaverlo, 2011). Moreover, informal STEM learning experience was one of the predictors of secondary students' interest in STEM careers (Stocklmayer et al., 2010; Halim et al., 2021).

Overall, informal STEM learning is an important way to provide students with authentic experiences that can increase their interest in STEM (Stocklmayer et al., 2010; Mohr-Schroeder et al., 2014; Roberts et al., 2018) and in STEM careers (Dabney et al., 2012; Kitchen et al., 2018). Considering a fair amount of research has been focusing on the role of informal STEM education in attracting students into STEM careers by enhancing their self-efficacy (Tracey, 2010; Lent et al., 2018), STEM motivation (Suter, 2016; Vennix et al., 2018), and STEM interest (Heaverlo, 2011; Dabney et al., 2013), it is worth investigating if and how informal STEM education settings can also be a tool to address another important factor that predicts STEM success: spatial ability (Shea et al., 2001; Sorby, 2009).

3.4.2. Opportunities to develop students' spatial ability through informal STEM education

Many focus areas inside STEM domains, including teaching methods, assessment tools, curriculum developments, and informal learning programs, can be identified as promoting STEM understanding. Informal STEM learning, in particular, remains one of the primary methods to promote STEM disciplines. By using various delivery methods to reach different age groups in diverse settings, informal STEM learning leads students to integrate knowledge, concepts, and methodologies from different fields in order to achieve specific goals. This integration is often difficult to achieve within traditional pedagogy (Tillinghast et al., 2020). Implicit, or indirect, spatial training occurs when spatial training becomes a part of the normal learning activities such as math-related tasks within a math lesson (Uttal et al., 2013b; Maquet et al., 2022). Such implicit training can also be adopted in informal STEM education settings. Moreover, activities in informal STEM education programs often involve hands-on, project-based learning, and have the potential to give a spatial dimension to STEM content (Newcombe et al., 2013).

A middle school summer camp for engineering education explored the role of spatial thinking in engineering learning (Ramey and Uttal, 2017). This camp consisted of a mixture of lectures and hands-on engineering activities. The researchers found that both construction kit activities – a kit containing written/diagrammatic instructions and building materials, asking the participants to follow the instructions to build a specific device – and engineering design activities – a design challenge with specific guidelines and material constraints that allowed them to take multiple creative pathways to a solution – played important and complementary roles in eliciting engineering-relevant spatial skills. The researchers also identified a series of students' actions during the engineering, design, or making activities that reflected their cognitive spatial processes, which they defined as spatial sensemaking activities. These include, for example, gesturing to represent "a dynamic spatial arrangement/process," "drawing out ideas for the purpose of design," and "discussing shape, orientation, position, or movement of objects, groups of objects, or representation" (p. 289–290). Importantly, the summer camp environment allowed the researchers to unpack how different kinds of cognitive spatial processes – typically identified in the lab or through psychometric assessments – might look like in everyday learning contexts.

In a set of technology-enhanced, STEAM-making (STEM + Arts) activities, Ramey et al. (2020) examined how spatial reasoning contributed to fifth and sixth-graders' learning. These activities integrated concepts from STEAM disciplines, such as math or science, into making and tinkering activities, rather than traditional lectures or reading materials. Moreover, they were suitable for both in-school and out-of-school learning (Stevens et al., 2016). Ramey and colleagues found that hands-on, collaborative problem-solving with spatial tools and representations improved students' spatial reasoning, and that different types of spatial reasoning were used frequently. While previous studies to improve students' spatial ability have typically taken place in psychology laboratories or instructional courses (Uttal et al., 2013c), these interventions often used spatial representations that do not reflect the ones students encounter when solving STEM problems in their daily lives (Ramey et al., 2020). Therefore, incorporating spatial training in informal STEM education settings can provide students with a more authentic experience of the spatial representations they will encounter in real-world scenarios.

3.4.3. Using knowledge from existing spatial interventions to develop informal STEM education activities

Past research joined spatial ability dimensions and informal STEM education programs by challenging students with spatially complex STEM activities (e.g., Samaroo et al., 2018) or by understanding students' cognitive spatial processes involved in these activities (e.g., Ramey and Uttal, 2017; Ramey et al., 2020). A window of opportunity is open to use the knowledge gained from past spatial ability research alongside informal STEM education to advance this field of study. By joining spatial ability studies and informal STEM education, we may come closer to filling some gaps in research. For instance, we could investigate whether students' spatial ability levels affect their enjoyment or engagement in informal STEM education programs, or whether their spatial ability levels mediate the positive effect of informal STEM education programs on their self-efficacy.

To the best of our knowledge, existing studies that involve spatial activities within informal STEM education activities, have not collected spatial reasoning data or looked specifically into how spatial reasoning occurs during these activities. Instead, researchers have utilized the spatial dimension by engaging students in spatially complex activities and examining the impact of these activities on raising awareness of STEM careers. For example, in one informal STEM learning activity (Samaroo et al., 2018), the students from sixth to eighth grades were required to (1) model the engineering design process by making blueprints of their cities; and (2) create replica models showing a block of their cities using all recyclable materials. Through visualizing and sketching their designs, as well as transferring the design from paper to three-dimensional representation, the students made a connection from 2D to 3D which promotes spatial thinking. It is worth discussing how future research might integrate spatial ability research and informal STEM education research using established frameworks, such as the two-by-two spatial skills framework (Uttal et al., 2013a), or the lists of spatial sensemaking activities (Ramey and Uttal, 2017), to understand the cognitive spatial processes experienced by students.

Future informal STEM education programs might also make use of a wide range of activities that have been shown to support individuals' spatial ability development. Toys that provide construction play (such as Legos and blocks) have been shown to engage children in thinking spatially (Casey and Bobb, 2003; Sorby, 2009; Verdine et al., 2014). Origami-based instruction helped students understand concepts that had previously been difficult for them, such as angle, geometric shapes, area, and fractions (Cakmak et al., 2014). Technology play can also promote students' spatial skills development (Newcombe, 2010; Uttal et al., 2013c). For example, playing spatially challenging video games led to increased visual–spatial attention (Newcombe et al., 2013). Furthermore, Minecraft, which provides an open digital world that allows players freedom in the way of playing and building (Canossa et al., 2013), can be used as an instructional tool for supporting spatial skill development, with gains similar to the use of other 3D applications such as Google SketchUp and augmented reality (AR) (Carbonell-Carrera et al., 2021). The wide range of learning activities and experiences offered by informal STEM education allows it to engage a diverse range of audiences with spatial activities and training, helping them to develop familiarity and proficiency with spatial skills that are critical in STEM learning.

4. The moderating effects of gender and socioeconomic status

Aside from the various interventional or educational approaches to developing students' spatial abilities we have discussed, it is crucial to recognize that the development of spatial abilities occurs at the individual level. Individual variables, including gender and socioeconomic status, may moderate the impact of spatial interventions or spatialized educational practices on individual students. To situate students' learning of spatial skills in authentic educational contexts and to ensure that students from all backgrounds are equipped with the spatial skills vital for STEM learning, it is important to investigate whether certain student populations may be in stronger need of resources. In the following two sections, we first discuss the relationship between spatial ability and gender and how learning can be structured to reduce gender differences in spatial ability. We then discuss how socioeconomic status can impact students' acquisition and application of spatial skills, and how the development of future spatial interventions and spatial practices can benefit from taking these factors into consideration.

4.1. Gender differences in spatial tests, interventions, and practices

In many countries, participation in STEM fields is still facing the elephant in the room: a significant gender gap between males and females. Data collected by PISA (Programme for International Student Assessment) in 2015 shed some light on this topic, providing two compelling facts: while male students scored higher in mathematics and science on average, female students were stronger in reading; moreover, male students displayed more confidence and interest in learning science (Stoet and Geary, 2018), while female students showed less intention to choose a STEM career than male students. The underrepresentation of women in the STEM field implies that women have less promising career prospects in this field (Ruthsatz et al., 2012). When it comes to scientific and technological development, the industry is suffering from a dearth of female talent, viewpoints, and experiences.

Given the significance of science and technology in our society, it is necessary to address the possible causes of this gender gap. As we discussed earlier, spatial ability is one of the key factors to success in STEM areas (Shea et al., 2001; Wai et al., 2009; Ruthsatz et al., 2012), and having high spatial ability is a predictor for choosing a STEM career (Wai et al., 2009). Nevertheless, this ability is not absent of gender differences. Several studies have pointed out sex-related differences in spatial skills (e.g., Gorska et al., 1998; Sorby and Veurink, 2010; Neuburger et al., 2015; Moon et al., 2016; Newcombe, 2020), such as mental rotation and spatial perception tasks (Newcombe, 2020).

4.1.1. Gender differences in spatial ability: Where do they come from?

Gender differences in spatial task performance have been widely studied. They usually appear from the age of puberty (Quaiser-Pohl et al., 2016), although some studies suggested that they can emerge at around the age of 10 (Neuburger et al., 2012). The origin of these differences is not yet clear. A recent meta-analysis on gender differences in spatial ability showed that although decades of research have tried to find biological causes, none could sufficiently explain gender differences in spatial ability (Bartlett and Camba, 2023). Instead, Bartlett and Camba suggested that gender roles and social norms may play a more decisive role in shaping these disparities. Traditionally speaking, many non-formal spatial activities, such as football (Voyer et al., 2000) and construction play (Sorby, 2009), are regarded as better suited for boys. This may be why boys have been more encouraged than girls to play with toys that promote spatial ability development, engage in sports, and STEM-related courses (Lippa et al., 2010; Moè, 2012). Research indicated that parents even tend to use more spatial language with boys than with girls (Newcombe, 2020). Having more exposure to spatially complex activities from a young age potentially gives male students an advantage over female students in spatial tasks. Moreover, it is a common societal belief that boys are more suited for and capable of spatially complex activities than girls. Such beliefs can negatively impact girls' self-concept and potentially constrain their spatial skills development (Moè, 2012; Neuburger et al., 2015).

In spatial ability task performance, gender differences may be explained by the use of holistic or analytic strategies when solving spatial tasks (e.g., Kail et al., 1979; Hsi et al., 1997; Hegarty, 2018). Using holistic strategies means visualizing and manipulating the visual information as a whole, while using analytic strategies implies focusing on one part of the visual information at a time or employing verbal descriptions, which is less effective and more time-consuming (Maresch, 2014a). Although multiple strategies can be applied when solving a task (Maresch, 2014b), women tend to apply more analytic strategies, while men are more likely to apply more holistic strategies, potentially making their performance faster and more accurate (Glück et al., 2005).

Women also tend to answer more carefully than men do and hence, more slowly (Quaiser-Pohl et al., 2016). Therefore, when there is a time limit for a test, men tend to outperform women. Research indicated that if the time variable is eliminated or enough time is given, males and females performed more evenly (Moè and Pazzaglia, 2006; Voyer, 2011; Maeda and Yoon, 2015; Wang and Degol, 2017). On the other hand, some have suggested that removing the time factor or giving abundant time to solve the tests, could lead to a less accurate measurement of spatial skills, as they might be using more analytic strategies (Shepard and Metzler, 1971; Linn and Petersen, 1985).

All in all, while it is difficult to pinpoint the origin of gender differences in spatial ability, it seems that nurture, as well as the environments in which children develop, play essential roles (Newcombe, 2020). As discussed in previous sections, individuals' spatial abilities can be improved by training (Uttal et al., 2013a). In the following section, we explore how spatial interventions have the promise to reduce gender differences in spatial task performance (Newcombe, 2010).

4.1.2. Addressing gender differences through spatial interventions

As one of the possible causes for gender differences in spatial performance is male students having more exposure to spatial activities than female students, spatial interventions can potentially support female students' spatial ability development and confidence (Sorby et al., 2005). However, certain intervention designs may be more beneficial than others. Neuburger et al. (2012) explored how social expectations and self-concept influence fifth-grade students' spatial task performance. In their research, they created three conditions for an intervention, instructing students that: (a) girls outperform boys in mental rotation tasks; (b) boys outperform girls in such tasks; and (c) both girls and boys are equally skilled. In the "girls better" and "no gender difference" conditions, girls' performance on spatial tasks improved while boys' performance worsened. Moè (2012) found that assuming that performance depends on effort instead of innate abilities improved the spatial task performance of both women and men. These results revealed that it is important for spatial intervention to address psychological factors, such as their beliefs, self-concept, and self-efficacy (Lent et al., 2010; Tracey, 2010), that may heavily affect students' performance.

Practicing with feedback has also been found to enhance female students' performance in spatial thinking. Feedback, defined as information provided by an agent about one's performance or understanding (Hattie, 2008), is a powerful tool through which the learning process can be enhanced (Hattie and Timperley, 2007). However, its efficacy depends on various factors, with some related to the characteristics of the feedback, such as the person who gives the feedback, the timing, and the content, and some concerning students' individualities, such as their motivation levels, learning goals, prior knowledge, or gender (Maier et al., 2016). In fact, previous research showed that, in general, female students benefited more than male students from practicing with feedback, resulting in more gains in spatial ability development (Narciss et al., 2014). Moreover, the reaction time of females improved after receiving feedback, as it seemed to make them feel more confident about their skills and answers (Rahe et al., 2019). Therefore, practicing with feedback might translate into a motivational strategy that reduces the gender gap by improving women's performance in spatial tasks (Kass et al., 1998). In order to enhance female students' spatial ability and consequently, their presence in STEM activities, it is fundamental to provide them with interventions that attend to their needs.

Lastly, creating a collaborative learning environment also shows promise in reducing the gender gap in spatial task performance (Phelps and Damon, 1989). Hoskyns-Staples and Blackmore (2020) noticed that when girls were constructing brick buildings in mixed-age groups, they were able to share their expertise and support each other with the spatial skills and mathematical knowledge needed to build and arrange the brick buildings. In addition, these girls made use of their social knowledge to ensure that the brick town reflected the facilities equipped in a real-life town. Therefore, it is worth considering adding the element of collaboration as well as using real-life problem-solving scenarios in spatial interventions, so that both girls and boys can have the opportunity to capitalize on their collaborative skills and social knowledge.

4.1.3. Developing interventions that support spatial ability development of both genders

After decades of research in spatial ability, there are still many unanswered questions, such as how gender differences impact individuals' spatial ability at different developmental stages, and how gender may moderate the relationship between spatial ability and STEM learning. Even though the gender gap in spatial ability cannot be completely erased yet, there are several steps that can be taken in order to minimize them.

First, both girls and boys should be equally encouraged to play with spatial toys (e.g., construction bricks, tangrams, puzzles, etc.) and to engage in spatial activities (e.g., playing sports, attending science courses, etc.), as these set the path for developing spatial skills (Newcombe and Frick, 2010). Special attention should also be paid to the language used with children. This not only refers to using more spatial language but also to reassuring children of their spatial abilities. It is also worth investigating if playing materials that have been traditionally gendered can be utilized to promote girls' spatial skills development. Crafting, sewing, and textile design, while typically regarded as feminine, actually demand a range of spatial skills (Newcombe et al., 1983; Workman and Ling Zhang, 1999). For example, E-textiles, which are fabric artifacts that require knowledge of computing and electronics (Berzowska, 2005), have been found to be especially engaging for girls (Buechley et al., 2013) and stimulated their leadership in making (Buchholz et al., 2014). Innovative technology like this may be especially valuable to challenge girls' spatial thinking while developing their interdisciplinary knowledge and skills.

Feedback, which has been shown to positively impact female students' performance in spatial tasks, should also be considered in future spatial ability assessments, as well as in teaching and learning in authentic environments. Further research is needed to determine when and how feedback can be provided during spatial interventions to best support both female and male students. Lastly, as Bartlett and Camba (2023) concluded, many of the traditionally-used spatial tests were built upon the criteria of "maximizing gender differences in favor of males" (p. 17). Thus, future research must be cautious in selecting approaches to measure students' spatial ability in ways that do not reinforce the gender gap.

4.2. Socioeconomic status and its effect on students' spatial ability

Aside from the gender gap in STEM, the underrepresentation of students from low socioeconomic status (SES), ethnic minorities, and other marginalized groups in STEM fields (National Research Council, 2011; MacPhee et al., 2013; Saw et al., 2018; Rosenthal, 2021) remains a particular concern that requires attention. Underrepresentation can restrict the pool from which skillful individuals can be selected to follow careers in STEM and negatively impact the self-efficacy in STEM learning of marginalized groups, making them particularly vulnerable to dropping out of STEM programs (Marginson et al., 2013). Rather alarmingly, students who suffer from more than one social disadvantage, such as being from an ethnic minority group and having a low SES background, face even greater challenges in developing an interest and building self-efficacy in learning STEM, as well as in entering and persisting in the STEM fields compared to those who face only one or none social disadvantage (MacPhee et al., 2013; Saw et al., 2018).

Tracking more than 500 primary school children in the United States for 2 years and assessing their spatial ability at four time points, Levine et al. (2005) found that students from the high- and middle-SES groups consistently performed better than those from the low-SES group on both the mental rotation task and the aerial-maps task. Casey et al. (2011) derived similar findings from their fourth-grade sample that children from low-income communities performed worse on spatial reasoning than children from affluent communities. Such discrepancy between the high- & middle-SES groups' and the low-SES groups' performance on spatial tasks has been seen as early as in preschool (Jirout and Newcombe, 2015; Bower et al., 2020). Multiple factors, such as the lack of access to spatially challenging toys or games (Levine et al., 2005; Jirout and Newcombe, 2015), lower quality of spatial play (Bower et al., 2020), and limited learning opportunities and environmental stimuli that promote spatial practices from their immediate environment (Casey et al., 2011), may have led to disparities in spatial abilities.

4.2.1. Spatial interventions and their indications for students from underprivileged backgrounds

Bower et al. (2020) carried out a five-week, high-quality, playful spatial task training that incorporated feedback, gestures, and spatial language from trainers to help children correct errors during their play. Contrary to children without training, children from low-SES backgrounds who received spatial training showed enhanced performance in 2D spatial tasks [$\beta = 0.28$, $p < 0.001$]. A moderation effect was found between SES and the transfer of training effect from trained 2D spatial task to untrained 3D spatial task. Through the playful training with 2D tasks, only children from low-SES backgrounds improved in their performance of solving part of the 3D spatial tasks [$\beta = 0.19$, $p = 0.027$] but not those from high-SES backgrounds. Overall, children from low-SES backgrounds benefited substantially from spatial training compared to their high-SES counterparts. As Bower and colleagues explained, spatial training might have given children from low-SES backgrounds the spatial skills and tools that are essential to solving spatial and math problems. Without spatial training, the experiences of practicing and applying these skills may not have been easily accessible to children from low-SES backgrounds compared to their high-SES counterparts.

To understand how SES can moderate spatial skills' relationship to STEM performance, Casey et al. (2011) analyzed how students' performance on spatial tasks relates to their performance on math tasks. They found that the association between spatial reasoning scores and math problem-solving performance among students from affluent communities was strong with a large effect size (from 0.69 to 0.83), whereas such an association was not statistically significant for students from low-income communities. This study highlighted the need for future research on educational spatial interventions to consider not only the spatial ability levels of students from low-SES backgrounds but also their abilities to utilize acquired spatial skills to solve academic or real-world problems.

Another way to address the needs of underprivileged students is to unpack the challenges they face when confronted with spatially demanding tasks. Bhaduri et al. (2021) conducted a study with 397 secondary students from rural schools, who were mostly from low-SES and ethnic-minority backgrounds, to introduce them to 3D design and 3D printing, as well as to understand the challenges in spatial thinking faced by these students when using online 3D interfaces. Using a combination of quantitative and qualitative data, such as interviews and screen recordings from the design interface, the research team was able to investigate several difficulties students experienced when building spatially complex models on the virtual 3D platform. The insights gained from this research are helpful for educators and future researchers to understand the practical and technical barriers that may hinder students' spatial skills development. Meanwhile, the in-depth qualitative data shed light on different types of scaffolds that can be provided to underprivileged students during future 3D design processes.

4.2.2. Making spatial interventions inclusive for students from underprivileged backgrounds

While it can be difficult to conduct studies solely focused on a minority group, due to reasons such as a limited number of participants and other demographic concerns, it is necessary for future researchers to consider the potential moderating effect of socioeconomic status aside from, and along with, gender when intervening to develop students' spatial ability.

Some researchers have been making more efforts to involve students from diverse backgrounds in their sample, such as recruiting participants from a wide range of SES backgrounds (e.g., Casey et al., 2011; Ramful et al., 2017; Lowrie and Jorgensen, 2018). This approach not only increases the representativeness of the studies but also ensures that research findings are relevant to previously understudied groups who could potentially benefit the most from these interventions. Yet, despite these efforts, there is still a dearth of studies that address the relationship between SES and spatial ability (Carr et al., 2018) and the problem of underrepresentation of socially disadvantaged students in STEM persists. Innovative practices in interventions need to be geared toward all SES groups, with a focus on developing spatial interventions to equip students with the necessary spatial skills they may not have been able to gain from their day-to-day environment (Bower et al., 2020). Meanwhile, special attention needs to be paid to the issue of intersectionality through the recruitment of groups who experience multiple social disadvantages, such as female students from low-SES backgrounds and/or ethnic minorities, as they could be the ones who benefit the most from spatial interventions. In addition, when analyzing and interpreting data from educational spatial interventions, future research needs to factor in the moderating effect of variables such as SES, ethnicity, and gender.

Various social disadvantages might interfere with students' performance in spatial tasks as well as in STEM. From a research point of view, a mixed-method research design may help researchers gain a richer understanding of the needs of and the barriers faced by students from underprivileged groups. For example, semi-structured interviews allowed researchers to identify difficulties faced by students when working in the 3D space (Bhaduri et al., 2021). Qualitatively analyzing the robotics design work of underprivileged students from rural areas, Leonard and colleagues found that these students incorporated "elements of culture and place into game design" (2016, p. 873). The in-depth information obtained from qualitative data in addition to quantitative data allowed researchers to design future authentic learning experiences that foster an interest in STEM among underprivileged students (Leonard et al., 2016).

As we mentioned earlier, informal STEM education makes authentic and engaging learning experiences available to a wide range of audiences. With the increasing number of summer camps and afterschool programs that target minority students or students from underprivileged backgrounds (Repenning et al., 2010), informal STEM learning may also be one of the solutions to supply these students with learning and playing resources that are spatially stimulating.

5. Conclusion

One theme that stood out in our paper was the progress made, and steps to be taken, in situating spatial interventions, spatial activities, and spatialized educational practices, in authentic learning contexts.

Revisiting the relationship between spatial ability and STEM learning from a cognitive science perspective, we highlighted the need to understand the cognitive mechanisms that explain how different spatial skills are related to different aspects of STEM learning. A further question we can ask from this, similar to a point of deliberation raised by existing studies (Stull et al., 2012; Stieff et al., 2014; Atit et al., 2020), is whether excelling at spatial task performance necessarily leads to effective applications of spatial skills to solve real-world STEM problems.

Among spatial interventions that have been delivered in physical, digital, or hybrid format, as well as the increasing efforts to spatialize STEM curricula, we noticed that one tendency in the design of recent spatial interventions is to align training materials and training modes more closely with students' daily learning experiences. Whether it is using hands-on manipulatives, providing digital feedback when students solve spatial tasks, or adapting existing student-centered pedagogies to support students' spatial development needs, they represent the intentional recruitment of authentic, contextually-relevant teaching and learning methods.

To spotlight additional approaches to developing K-12 students' spatial ability through authentic learning experiences, we turned to integrated STEM education and informal STEM education, which both play important roles in general STEM education but have rarely been explored by researchers in the field of spatial ability. Integrated STEM education encompasses the interdisciplinary essence of STEM and the interplay among the four elements, engineering design, technological literacy, scientific inquiry, and mathematical thinking. We suggested that spatializing more aspects of integrated STEM education would be desirable and feasible given how existing studies have made use of one or more of these elements to develop students' spatial abilities. Despite that most of the preceding discussions have focused on formal education, informal STEM education presents a promising avenue that offers untapped resources to engage and immerse students in spatial thinking. While some of the spatial representations in existing spatial training do not necessarily reflect the types of spatial representations encountered by students when solving real-world problems (e.g., Julià and Antolì, 2018; Atit et al., 2020; Ramey et al., 2020), both integrated STEM education and informal STEM education have the potential to familiarize students with spatial representations they will see and use in authentic learning environments.

Weighing various approaches to developing students' spatial ability, while emphasizing the importance of situating learning in authentic contexts, we propose a conjecture that the overall relationship between developing students' spatial thinking and enhancing their abilities to solve real-world STEM problems is bidirectional and can be conveniently categorized as shown in Figure 3. It is clear that spatial ability plays a vital role in STEM learning (e.g., Sorby, 2009; Mix et al., 2021; Hawes et al., 2022), and that one of the key objectives of developing students' spatial abilities is to increase the enrollment and performance in STEM (e.g., Newcombe, 2010; Uttal et al., 2013a; Stieff and Uttal, 2015). Meanwhile, researchers have underscored the need for spatial training that authentically addresses the spatial skills students use to solve real-world STEM problems (e.g., Ormand et al., 2014; Atit et al., 2020; Ramey et al., 2020). Therefore, we anticipate that directly using spatially complex STEM problems to target students' spatial skills development may be a desirable option and is readily applicable in conventional classroom teaching, integrated STEM practices, as well as in informal STEM activities. Additionally, spatially complex STEM problems may have the benefit of developing students' spatial ability along with their content knowledge and skills in multiple disciplines (e.g., Peng and Sollervall, 2014; Burte et al., 2017; Julià and Antolì, 2018). Explicitly encouraging students to use spatial skills when solving spatially demanding problems is expected to support their understanding of scientific and technological concepts and practices. For example, having secondary school students create visualizations of mechanical or chemical systems not only solidified their understanding of these systems but also developed their spatial thinking skills (Bobek and Tversky, 2016). More research is needed to further examine this interplay.

Finally, we discussed how gender and SES may moderate the effect of spatial interventions on students' spatial ability development. While the reasons contributing to such moderation effects are often complicated, we believe it is worthwhile for future research to take factors such as gender, SES, and the interaction between these two factors into careful consideration when planning research, drawing samples, and interpreting data. In addition, while a large number of spatial intervention studies relied on quantitative measures, mixed-method research designs and qualitative data collection methods may yield more nuanced insights into educational contexts and differential individual experiences. By obtaining a comprehensive understanding of the effect of spatial interventions, researchers can then design scaffolds that support the spatial ability development of students from underprivileged social groups or those who have been underrepresented in the STEM field. Overall, extrapolating spatial intervention findings with gender, socioeconomic status, and other potential moderating factors in mind can help translate interventions into sustainable educational practices (Casey et al., 2011; Bower et al., 2020) and make research findings more informative for both educational practitioners and policymakers.

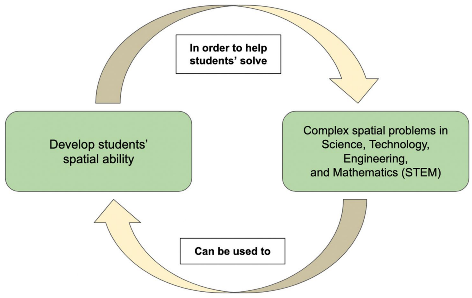

FIGURE 3
The conjectured relationship between developing students' spatial ability and enhancing their abilities to solve spatially complex STEM problems.

Author contributions

CZ contributed to drafting and final editing. CL, EL, MV, and NS-C contributed equally to drafting and editing. DJ, GD, GM, MP, and RK contributed to editing, providing feedback, and proofreading. All authors contributed to the manuscript and approved it for submission.

Funding

This research is part of the SellSTEM - Spatially Enhanced Learning Linked to STEM - Marie Skłodowska-Curie Innovative Training Network to investigate the role of spatial ability in and for STEM learning. It has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 956124.

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.

Publisher's note

References

Atit, K., Gagnier, K., and Shipley, T. F. (2015). Student gestures aid penetrative thinking. *J. Geosci. Educ.* 63, 66–72. doi: 10.5408/14-008.1

Atit, K., Uttal, D. H., and Stieff, M. (2020). Situating space: Using a discipline-focused lens to examine spatial thinking skills. *Cogn. Res.* 5, 1–16. doi: 10.1186/s41235-020-00210-z

Baddeley, A. (1993). "Working memory or working attention?" in *Attention: Selection, awareness and control. A tribute to Donald Broadbent*. eds. A. Baddeley and L. Weiskrantz (Oxford: Oxford University Press), 152–170.

Bartlett, K. A., and Camba, J. D. (2023). Gender differences in spatial ability: A critical review. *Educ. Psychol. Rev.* 35, 1–29. doi: 10.1007/s10648-023-09728-2

Berzowska, J. (2005). Electronic textiles: Wearable computers, reactive fashion, and soft computation. *Text* 3, 58–75. doi: 10.2752/147597505778052639

Bhaduri, S., Biddy, Q. L., Bush, J., Suresh, A., and Sumner, T. (2021). 3DnST: A framework towards understanding children's interaction with Tinkercad and enhancing spatial thinking skills. Proceedings of the Interaction Design and Children (IDC'21), Association for Computing Machinery, Athens, Greece.

Bobek, E., and Tversky, B. (2016). Creating visual explanations improves learning. *Cogn. Res.* 1, 27–14. doi: 10.1186/s41235-016-0031-6

Boone, A., Vanderwall, J., Klitsner, M., and Spyridakis, I. (2020). STEM outreach in underrepresented communities through the lens of play, creativity, and movement. Proceedings of the 2020 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, USA.

Bower, C., Zimmermann, L., Verdine, B., Toub, T. S., Islam, S., Foster, L., et al. (2020). Piecing together the role of a spatial assembly intervention in preschoolers' spatial and mathematics learning: Influences of gesture, spatial language, and socioeconomic status. *Dev. Psychol.* 56, 686–698. doi: 10.1037/dev0000899

Brown, R., Brown, J., Reardon, K., and Merrill, C. (2011). Understanding STEM: Current perceptions. *Technol. Eng. Teach.* 70, 5–9.

Bryan, L. A., Moore, T. J., Johnson, C. C., and Roehrig, G. H. (2015). "Integrated STEM education" in *STEM road map: A framework for integrated STEM education*. eds. C. C. Johnson, E. E. Peters-Burton and T. J. Moore (New York, NY: Routledge Taylor & Francis Group)

Buchholz, B., Shively, K., Peppler, K., and Wohlwend, K. (2014). Hands on, hands off: Gendered access in crafting and electronics practices. *Mind Cult. Act.* 21, 278–297. doi: 10.1080/10749039.2014.939762

Buckley, J., Seery, N., and Canty, D. (2018). A heuristic framework of spatial ability: A review and Synthesis of spatial factor literature to support its translation into STEM education. *Educational Psychology Review* 30, 947–972. doi: 10.1007/s10648-018-9432-z

Buechley, L., Peppler, K. A., Eisenberg, M., and Kafai, Y. B. (Eds.). (2013). *Textile messages: Dispatches from the world of e-textiles and education*. New York, NY: Peter Lang.

Burte, H., Gardony, A. L., Hutton, A., and Taylor, H. A. (2017). Think3d!: Improving mathematics learning through embodied spatial training. *Cogn. Res.* 2, 13–18. doi: 10.1186/s41235-017-0052-9

Cakmak, S., Isiksal, M., and Koc, Y. (2014). Investigating effect of origami-based instruction on elementary students' spatial skills and perceptions. *J. Educ. Res.* 107, 59–68. doi: 10.1080/00220671.2012.753861

Canossa, A., Martinez, J. B., and Togelius, J. (2013). Give me a reason to dig Minecraft and psychology of motivation. Proceedings of the 2013 IEEE Conference on Computational Inteligence in Games (CIG), Niagara Falls, ON, Canada.

Caprile, M., Palmén, R., Sanz, P., and Dente, G. (2015). Encouraging STEM studies labour market situation and comparison of practices targeted at young people in different member states. *Publications Office*. 1–38. doi: 10.2861/519030

Carbonell-Carrera, C., Jaeger, A. J., Saorín, J. L., Melián, D., and de la Torre-Cantero, J. (2021). Minecraft as a block building approach for developing spatial skills. *Entertain. Comput.* 38:100427. doi: 10.1016/j.entcom.2021.100427

Carroll, J. B. (1993). Human cognitive abilities: a survey off actor analytic studies (Cambridge: Cambridge University Press).

Carr, M., Alexeev, N., Wang, L., Barned, N., Horan, E., and Reed, A. (2018). The development of spatial skills in elementary school students. *Child Dev.* 89, 446–460. doi: 10.1111/cdev.12753

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

Casey, B. M., and Bobb, B. (2003). Early childhood corner: The power of block building. *Teach. Child. Math.* 10, 98–102. doi: 10.5951/TCM.10.2.0098

Casey, B. M., Dearing, E., Vasilyeva, M., Ganley, C. M., and Tine, M. (2011). Spatial and numerical predictors of measurement performance: The moderating effects of community income and gender. *J. Educ. Psychol.* 103, 296–311. doi: 10.1037/A0022516

Casey, M. B., Nuttall, R. L., and Pezaris, E. (2001). Spatial-mechanical reasoning skills versus mathematics self-confidence as mediators of gender differences on mathematics subtests using cross-national gender-based items. *J. Res. Math. Educ.* 32, 28–57. doi: 10.2307/749620

Chatterjee, A. (2008). The neural organization of spatial thought and language. *Semin. Speech Lang.* 29, 226–238. doi: 10.1055/s-0028-1082886

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

Cohrssen, C., de Quadros-Wander, B., Page, J., and Klarin, S. (2017). Between the big trees: A project-based approach to investigating shape and spatial thinking in a kindergarten program. *Australas. J. Early Childhood* 42, 94–104. doi: 10.23965/AJEC.42.1.11

Cohrssen, C., and Pearn, C. (2021). Assessing preschool children's maps against the first four levels of the primary curriculum: Lessons to learn. *Math. Educ. Res. J.* 33, 43–60. doi: 10.1007/s13394-019-00298-7

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

Dabney, K. P., Tai, R. H., Almarode, J. T., Miller-Friedmann, J. L., Sonnert, G., Sadler, P. M., et al. (2012). Out-of-school time science activities and their association with career interest in STEM. *Int. J. Sci. Educ. Part B* 2, 63–79. doi: 10.1080/21548455.2011.629455

Davies, S. R., and Horst, M. (2016). *Science communication: Culture, identity and citizenship, 1st Edn*. Berlin: Springer.

Delgado, A. R., and Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Int. Intell.* 32, 25–32. doi: 10.1016/S0160-2896(03)00061-8

Dortch, D., and Patel, C. (2017). Black undergraduate women and their sense of belonging in STEM at predominantly white institutions. *NASPA J. Women Higher Educ.* 10, 202–215. doi: 10.1080/19407882.2017.1331854

Duffy, G., Sorby, S. A., and Bowe, B. (2020). An investigation of the role of spatial ability in representing and solving word problems among engineering students. *J. Eng. Educ.* 109, 424–442. doi: 10.1002/jee.20349

Falk, J. H., and Dierking, L. D. (2010). The 95 percent solution. *Am. Sci.* 98, 486–493. doi: 10.1511/2010.87.486

Fatourou, P., Papageorgiou, Y., and Petousi, V. (2019). Women are needed in STEM: European policies and incentives. *Commun. ACM* 62, 52–57. doi: 10.1145/3312565

Francis, K., Rothschuh, S., Poscente, D., and Davis, B. (2021). Malleability of spatial reasoning with short-term and long-term robotics interventions. *Technol. Knowl. Learn.* 27, 927–956. doi: 10.1007/s10758-021-09520-7

Freeman, B., Marginson, S., and Tytler, R. (2019). "An international view of STEM education" in *STEM Education, 2.0* (Leiden, Netherlands: Brill Sense)

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., and Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Dev.* 78, 1343–1359. doi: 10.1111/j.1467-8624.2007.01069.x

Gilligan, K. A., Hodgkiss, A., Thomas, M. S., and Farran, E. K. (2018). The use of discrimination scalingtasks: A novel perspective on the development of spatial scaling in children. *Cognitive Development*. 47, 133–145. doi: 10.1016/j.cogdev.2018.04.001

Gilligan, K. A., Thomas, M. S. C., and Farran, E. K. (2020). First demonstration of effective spatial training for near transfer to spatial performance and far transfer to a range of mathematics skills at 8 years. *Dev. Sci.* 23:e12909. doi: 10.1111/desc.12909

Glück, J., Kaufmann, H., Dünser, A., and Steinbügl, K. (2005). Geometrie und Raumvorstellung – Psychologische Perspektiven. *Informationsblätter der Geometrie (IBDG)* 24, 4–11.

Gorska, R., Sorby, S. A., and Leopold, C. (1998). Gender differences in visualization skills-an international perspective. *Eng. Des. Graph. J.* 62, 9–18.

Gough, A. (2015). STEM policy and science education: Scientistic curriculum and sociopolitical silences. *Cult. Stud. Sci. Educ.* 10, 445–458. doi: 10.1007/s11422-014-9590-3

Gunderson, E. A., Ramirez, G., Beilock, S. L., and Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Dev. Psychol.* 48, 1229–1241. doi: 10.1037/a0027433

Halim, L., Mohd Shahali, E. H., and Iksan, H. Z. (2021). Effect of environmental factors on students' interest in STEM careers: The mediating role of self-efficacy. *Res. Sci. Technol. Educ.* doi: 10.1080/02635143.2021.2008341, [Epub ahead of print]

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. *1st Edn*. New York: Routledge.

Hattie, J., and Timperley, H. (2007). The power of feedback. *Rev. Educ. Res.* 77, 81–112. doi: 10.3102/003465430298487

Hawes, Z., and Ansari, D. (2020). What explains the relationship between spatial and mathematical skills? A review of evidence from brain and behavior. *Psychon. Bull. Rev.* 27, 465–482. doi: 10.3758/s13423-019-01694-7

Hawes, Z. C. K., Gilligan-Lee, K. A., and Mix, K. S. (2022). Effects of spatial training on mathematics performance: A meta-analysis. *Dev. Psychol.* 58, 112–137. doi: 10.1037/dev0001281

Hawes, Z., Moss, J., Caswell, B., Naqvi, S., and MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: Effects of a 32-week intervention. *Cogn. Instr.* 35, 236–264. doi: 10.1080/07370008.2017.1323902

Hawes, Z., Moss, J., Caswell, B., and Poliszczuk, D. (2015). Effects of mental rotation training on children's spatial and mathematics performance: A randomized controlled study. *Trends Neurosci. Educ.* 4, 60–68. doi: 10.1016/j.tine.2015.05.001

Heaverlo, C. (2011). *STEM development: A study of 6th-12th grade girls' interest and confidence in mathematics and science*. [Doctoral Dissertation]. Ames, IW: Iowa State University.

Hegarty, M. (2014). Spatial thinking in undergraduate science education. *Spat. Cogn. Comput.* 14, 142–167. doi: 10.1080/13875868.2014.889696

Hegarty, M. (2018). Ability and sex differences in spatial thinking: What does the mental rotation test really measure? *Psychon. Bull. Rev.* 25, 1212–1219. doi: 10.3758/s13423-017-1347-z

Hegarty, M., Keehner, M., Cohen, C., Montello, D. R., and Lippa, Y. (2007). "The role of spatial cognition in medicine: Applications for selecting and training professionals" in *Applied spatial cognition: From research to cognitive technology*. ed. G. L. Allen (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 285–315.

Herrington, J., Reeves, T. C., and Oliver, R. (2014). "Authentic learning environments" in *Handbook of research on educational communications and technology*. eds. J. Spector, M. Merrill, J. Elen and M. Bishop (New York, NY: Springer)

Hinze, S. R., Williamson, V. M., Shultz, M. J., Deslongchamps, G., Williamson, K. C., and Rapp, D. N. (2014). "Spatial ability and learning from visualizations in STEM disciplines" in *Space in mind: Concepts for spatial learning and education*. eds. D. R. Montello, K. Grossner and D. G. Janelle, The MIT Press 99–118.

Hodgkiss, A., Gilligan, K. A., Tolmie, A. K., Thomas, M. S., and Farran, E. K. (2018). Spatial cognition and science achievement: The contribution of intrinsic and extrinsic spatial skills from 7 to 11 years. *Br. J. Educ. Psychol.* 88, 675–697. doi: 10.1111/bjep.12211

Hoskyns-Staples, L., and Blackmore, K. (2020). "Collaborative problem-solving in primary mathematics: Developing shape and spatial awareness" in *Social and learning relationships in primary schools. Bloomsbury monograph series* eds. A. Kington and K. Blackmore (London: Bloomsbury Academic), 9–30. doi: 10.5040/9781350096097.ch-001

Hsi, S., Linn, M. C., and Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. *J. Eng. Educ.* 86, 151–158. doi: 10.1002/j.2168-9830.1997.tb00278.x

Jirout, J. J., and Newcombe, N. S. (2015). Building blocks for developing spatial skills: Evidence from alarge, representative US sample. *Psychologicalscience* 26, 302–310. doi: 10.1177/0956797614563338

Jones, S., and Burnett, G. (2008). Spatial ability and learning to program. *Hum. Technol.* 4, 47–61. doi: 10.17011/ht/urn.200804151352

Joyce, A., and Dzoga, M. (2011). *Intel white paper: Science, technology, engineering and mathematics education–overcoming challenges in Europe*. Brussels: European SchoolNet-Intel Educator Academy EMEA.

Julià, C., and Antolì, J. Ò. (2018). Enhancing spatial ability and mechanical reasoning through a STEM course. *Int. J. Technol. Des. Educ.* 28, 957–983. doi: 10.1007/S10798-017-9428-X

Kail, R., Carter, P., and Pellegrino, J. (1979). The locus of sex differences in spatial ability. *Percept. Psychophys.* 26, 182–186. doi: 10.3758/BF03199867

Kass, S. J., Ahlers, R. H., and Dugger, M. (1998). Eliminating gender differences through practice in an applied visual spatial task. *Hum. Perform.* 11, 337–349. doi: 10.1207/s15327043hup1104_3

Kearney, C. (2011). *Efforts to increase students' interest in pursuing science, technology, engineering and mathematics studies and careers. National Measures*. Brussels, Belgium: European Schoolnet.

Kell, H. J., Lubinski, D., Benbow, C. P., and Steiger, J. H. (2013). Creativity and technical innovation: Spatial ability's unique role. *Psychol. Sci.* 24, 1831–1836. doi: 10.1177/0956797613478615

Kelley, T. R., and Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *Int. J. STEM Educ.* 3, 1–11. doi: 10.1186/s40594-016-0046-z

Kelley, T. R., Sung, E., Han, J., and Knowles, J. G. (2022). Impacting secondary students' STEM knowledge through collaborative STEM teacher partnerships. *Int. J. Technol. Des. Educ.* 1–22. doi: 10.1007/s10798-022-09783-w

Kitchen, J. A., Sonnert, G., and Sadler, P. M. (2018). The impact of college- and university-run high school summer programs on students' end of high school STEM career aspirations. *Sci. Educ.* 102, 529–547. doi: 10.1002/sce.21332

Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., and Hautamäki, J. (2003). Visuospatial working memory and early numeracy. *Educ. Child Psychol.* 20, 65–76. doi: 10.53841/bpsecp.2003.20.3.65

Kyttälä, M., and Lehto, J. E. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence. *Eur. J. Psychol. Educ.* 23, 77–94. doi: 10.1007/BF03173141

Lefevre, J. A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., et al. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Dev.* 81, 1753–1767. doi: 10.1111/j.1467-8624.2010.01508.x

Lehrer, R., Kobiela, M., and Weinberg, P. J. (2013). Cultivating inquiry about space in a middle school mathematics classroom. *ZDM* 45, 365–376. doi: 10.1007/s11858-012-0479-x

Lent, R. W., Sheu, H.-B., Gloster, C. S., and Wilkins, G. (2010). Longitudinal test of the social cognitive model of choice in engineering students at historically black universities. *J. Vocat. Behav.* 76, 387–394. doi: 10.1016/j.jvb.2009.09.002

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

Leonard, J., Buss, A., Gamboa, R., Mitchell, M., Fashola, O. S., Hubert, T., et al. (2016). Using robotics and game design to enhance children's self-efficacy, STEM attitudes, and computational thinking skills. *J. Sci. Educ. Technol.* 25, 860–876. doi: 10.1007/s10956-016-9628-2

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

Lin, H. (2016). Influence of design training and spatial solution strategies on spatial ability performance. *Int. J. Technol. Des. Educ.* 26, 123–131. doi: 10.1007/s10798-015-9302-7

Linn, M. C., and Petersen, A. (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., Collar, M. L., and Peters, M. (2010). Sex differences in mental rotation and line angle judgements are positively associated with gender equality and economic development across 53 nations. *Arch. Sex. Behav.* 39, 990–997. doi: 10.1007/d10508-008-9460-8

Lord, T. R. (1987). A look at spatial abilities in undergraduate women science majors. *J. Res. Sci. Teach.* 24, 757–767. doi: 10.1002/tea.3660240808

Lowrie, T., and Jorgensen, R. (2018). Equity and spatial reasoning: Reducing the mathematical achievement gap in gender and social disadvantage. *Math. Educ. Res. J.* 30, 65–75. doi: 10.1007/s13394-017-0213-7

Lowrie, T., Logan, T., and Hegarty, M. (2019). The influence of spatial visualization training on students' spatial reasoning and mathematics performance. *J. Exp. Educ.* 20, 729–751. doi: 10.1080/15248372.2019.1653298

Lowrie, T., Logan, T., and Ramful, A. (2017). Visuospatial training improves elementary students' mathematics performance. *Br. J. Educ. Psychol.* 87, 170–186. doi: 10.1111/bjep.12142

Lowrie, T., and Patahuddin, S. M. (2015). ELPSA as a lesson design framework. *J. Math. Educ.* 6, 77–92. doi: 10.22342/JME.6.2.2166.77-92

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

Maeda, Y., and Yoon, Y. S. (2015). Are gender differences in spatial ability real or an artifact? Evaluation of measurement invariance on the revised PSVT:R. *J. Psychoeduc. Assess.* 34, 397–403. doi: 10.1177/0734282915609843

Maier, U., Wolf, N., and Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Comput. Educ.* 95, 85–98. doi: 10.1016/j.compedu.2015.12.002

Maquet, L., Dunbar, R., Buckley, J., Seery, N., and Sorby, S. (2022). A review of the literature to inform the efficacy of a spatial skills intervention for secondary level STEM education. In the 39th pupils attitude towards technology (PATT) conference proceedings, Canada.

Maresch, G. (2014a). Spatial ability–the phases of spatial ability research. *J. Geometry Graphics* 17, 237–250.

Maresch, G. (2014b). Strategies for assessing spatial ability tasks. *J. Geometry Graphics* 18, 125–132.

Marginson, S., Tytler, R., Freeman, B., and Roberts, K., (2013). *STEM: Country comparisons: International comparisons of science, technology, engineering and mathematics (STEM) education*. Final Report. Australian Council of Learned Academies. Melbourne, VIC.

Margulieux, L. E. (2020). Spatial encoding strategy theory: The relationship between spatial skill and stem achievement. *ACM Inroads* 11, 65–75. doi: 10.1145/3381891

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence* 37, 1–10. doi: 10.1016/j.intell.2008.08.004

Mix, K. S. (2019). Why are spatial skill and mathematics related? *Child Dev. Perspect.* 13, 121–126. doi: 10.1111/cdep.12323

Mix, K. S., Hambrick, D. Z., Satyam, V. R., Burgoyne, A. P., and Levine, S. C. (2018). The latent structure of spatial skill: A test of the 2× 2 typology. *Cognition* 180, 268–278. doi: 10.1016/j.cognition.2018.07.012

Mix, K. S., Levine, S. C., Cheng, Y. L., Stockton, J. D. S., and Bower, C. (2021). Effects of spatial training on mathematics in first and sixth grade children. *J. Educ. Psychol.* 113, 304–314. doi: 10.1037/edu0000494

Mix, K. S., Levine, S. C., Cheng, Y. L., Young, C., Hambrick, D. Z., Ping, R., et al. (2016). Separate but correlated: The latent structure of space and mathematics across development. *J. Exp. Psychol. Gen.* 145, 1206–1227. doi: 10.1037/xge0000182

Moè, A. (2012). Gender difference does not mean genetic difference: Externalizing improves performance in mental rotation. *Learn. Individ. Differ.* 22, 20–24. doi: 10.1016/j.lindif.2011.11.001

Moè, A., and Pazzaglia, F. (2006). Following the instructions! *Learn. Individ. Differ.* 16, 369–377. doi: 10.1016/j.lindif.2007.01.002

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: STEM camp. *Sch. Sci. Math.* 114, 291–301. doi: 10.1111/ssm.12079

Moon, Y., Jo, H., Kim, J., and Ryu, J. (2016). Exploring gender differences in spatial orientation ability on representing cognitive map. *Int. J. Psychol. Behav. Sci.* 6, 91–98. doi: 10.5923/j.ijpbs.20160602.09

Murphy, M. C., Gopalan, M., Carter, E. R., Emerson, K. T., Bottoms, B. L., and Walton, G. M. (2020). A customized belonging intervention improves retention of socially disadvantaged students at a broad-access university. *Sci. Adv.* 6:eaba4677. doi: 10.1126/sciadv.aba4677

Mustafa, N., Ismail, Z., Tasir, Z., and Mohamad Said, M. N. H. (2016). A meta-analysis on effective strategies for integrated STEM education. *Adv. Sci. Lett.* 22, 4225–4228. doi: 10.1166/asl.2016.8111

Narciss, S., Sosnovsky, S., Schnaubert, L., Andrès, E., Eichelmann, A., Goguadze, G., et al. (2014). Exploring feedback and student characteristics relevant for personalizing feedback strategies. *Comput. Educ.* 71, 56–76. doi: 10.1016/j.compedu.2013.09.011

National Research Council (2006). *Learning to think spatially*. Washington, DC: The National Academies Press.

National Research Council. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Washington, DC: The National Academies Press.

Netwong, T. (2018). Development of problem solving skills by integration learning following stem education for higher education. *Int. J. Inform. Educ. Technol.* 8, 639–643. doi: 10.18178/IJIET.2018.8.9.1114

Neuburger, S., Jansen, P., Heil, M., and Quaiser-Pohl, C. (2012). A threat in the classroom: Gender stereotype activation and mental-rotation performance in elementary-school children. *Z. Psychol.* 220, 61–69. doi: 10.1027/2151-2604/a000097

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. Differ.* 37, 169–175. doi: 10.1016/j.lindif.2014.09.003

Newcombe, N. S. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *Am. Educ.* 34:29.

Newcombe, N. S. (2017). *Harnessing spatial thinking to support stem learning. OECD Education Working Papers*. OECD Publishing: Paris.

Newcombe, N. S. (2020). The puzzle of spatial sex differences: Current status and prerequisites to solutions. *Child Dev. Perspect.* 14, 251–257. doi: 10.1111/cdep.12389

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

Newcombe, N. S., and Frick, A. (2010). Early education for spatial intelligence: Why, what, and how. *Mind Brain Educ.* 4, 102–111. doi: 10.1111/j.1751-228X.2010.01089.x

Newcombe, N. S., and Shipley, T. F. (2014). "Thinking about spatial thinking: New typology, new assessments" in *Studying visual and spatial reasoning for design creativity*. ed. J. Gero (Dordrecht: Springer Netherlands)

Newcombe, N. S., Uttal, D. H., and Sauter, M. (2013). "Spatial development" in *Oxford handbook of developmental psychology, Vol 1: Body and mind*. ed. P. Zelazo (Oxford: Oxford University Press), 564–590.

Oberle, A. (2020). Advancing students' abilities through the geo-inquiry process. *J. Geogr.* 119, 43–54. doi: 10.1080/00221341.2019.1698641

Ormand, C. J., Manduca, C., Shipley, T. F., Tikoff, B., Harwood, C. L., Atit, K., et al. (2014). Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study. *J. Geosci. Educ.* 62, 146–154. doi: 10.5408/13-027.1

Palmer, S. (1978). "Fundamental aspects of cognitive representation" in *Cognition and categorization*. ed. E. Rosch (Hillsdale, NJ: Erlbaum), 259–303.

Peng, A., and Sollervall, H. (2014). Primary school students' spatial orientation strategies in an outdoor learning activity supported by mobile technologies. *Int. J. Educ. Math. Sci. Technol.* 2, 246–256. doi: 10.18404/IJEMST.61603

Phelps, E., and Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *J. Educ. Psychol.* 81, 639–646. doi: 10.1037/0022-0663.81.4.639

Quaiser-Pohl, C., Jansen, P., Lehmann, J., and Kudielka, B. M. (2016). Is there a relationship between the performance in a chronometric mental-rotations test and salivary testosterone and estradiol levels in children aged 9-14 years?: Mental rotation and salivary hormones in (pre-)puberty. *Dev. Psychobiol.* 58, 120–128. doi: 10.1002/dev.21333

Rahe, M., Ruthsatz, V., Jansen, P., and Quaiser-Pohl, C. (2019). Different practice effects for males and females by psychometric and chronometric mental-rotation tests. *J. Cogn. Psychol.* 31, 92–103. doi: 10.1080/20445911.2018.1561702

Ramey, K. E., Stevens, R., and Uttal, D. H. (2020). In-FUSE-ing STEAM learning with spatial reasoning: Distributed spatial sensemaking in school-based making activities. *J. Educ. Psychol.* 112, 466–493. doi: 10.1037/edu0000422

Ramey, K. E., and Uttal, D. H. (2017). Making sense of space: Distributed spatial sensemaking in a middle school summer engineering camp. *J. Learn. Sci.* 26, 277–319. doi: 10.1080/10508406.2016.1277226

Ramful, A., Lowrie, T., and Logan, T. (2017). Measurement of spatial ability: Construction and validation of the spatial reasoning instrument for middle school students. *J. Psychoeduc. Assess.* 35, 709–727. doi: 10.1177/0734282916659207

Rasmussen, C., and Bisanz, J. (2005). Representation and working memory in early arithmetic. *J. Exp. Child Psychol.* 91, 137–157. doi: 10.1016/j.jecp.2005.01.004

Repenning, A., Webb, D., and Ioannidou, A. (2010). Scalable game design and the development of a checklist for getting computational thinking into public schools. In Proceedings of the 41st ACM technical symposium on Computer science education, New York, NY, United States

Reuhkala, M. (2001). Mathematical skills in ninth-graders: Relationship with visuo-spatial abilities and working memory. *Educ. Psychol.* 21, 387–399. doi: 10.1080/01443410120090786

Ritz, J. M., and Fan, S. C. (2015). STEM and technology education: International state-of-the-art. *Int. J. Technol. Des. Educ.* 25, 429–451. doi: 10.1007/s10798-014-9290-z

Roberts, 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

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walberg-Henriksson, H., and Hemmo, V. (2007). Science education NOW: A renewed education for the future of Europe. European Commission Directorate-General for Research. Publications Office.

Rosenthal, K. (2021). Is gender equality still an issue? Gender (Im) balances in STEM. *Chemie Ingenieur Technik* 93, 1207–1209. doi: 10.1002/cite.202000216

Ruthsatz, V., Neuburger, S., and Quaiser-Pohl, C. (2012). The social relevance and the socio-cultural origins of gender differences in spatial abilities. *Acta Universitatis Lodziensis Folia Sociol* 43, 17–32.

Samaroo, D., Villatoro, M., Narine, S., Iqbal, A., and Natal, K. (2018). "A multitier approach to integrating STEM education into a local elementary school" in *Science Education and Civic Engagement: An International Journal*. 10, 35–42.

Saw, G., Chang, C. N., and Chan, H. Y. (2018). Cross-sectional and longitudinal disparities in STEM career aspirations at the intersection of gender, race/ethnicity, and socioeconomic status. *Educ. Res.* 47, 525–531. doi: 10.3102/0013189X18787818

Schneider, W. J., and McGrew, K. S. (2018). "The Cattell–Horn–Carroll theory of cognitive abilities," in *Contemporary intellectual assessment: Theories, tests, andissues*. eds. D. P. Flanagan and E. M. McDonough (The Guilford Press), 73–163.

Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. Science. *Science.* 171, 701–703. doi: 10.1126/science.171.3972.701

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

Sisman, B., Kucuk, S., and Yaman, Y. (2021). The effects of robotics training on children's spatial ability and attitude toward STEM. *Int. J. Soc. Robot.* 13, 379–389. doi: 10.1007/s12369-020-00646-9

Sorby, S. A. (2009). Educational research in developing 3-D spatial skills for engineering students. *Int. J. Sci. Educ.* 31, 459–480. doi: 10.1080/09500690802595839

Sorby, S., Drummer, T., Hungwe, K., and Charlesworth, P. (2005). "Developing 3-D spatial visualization skills for non-engineering students" in *Proceedings of the 2005 American Society for Engineering Education Annual Conference & Exposition* Portland, Oregon: American Society for Engineering Education. doi: 10.18260/1-2--15370

Sorby, S., and Veurink, N. (2010). Are the visualization skills of first year engineering students changing? Paper presented at 2010 Annual Conference & Exposition, Louisville, Kentucky

Sorby, S. A., and Veurink, N. (2019). Preparing for STEM: Impact of spatial visualization training on middle school math performance. *J. Women Minorities Sci. Eng.* 25, 1–23. doi: 10.1615/JWomenMinorScienEng.2018024516

Stevens, R., Jona, K., Penney, L., Champion, D., Ramey, K. E., Hilppö, J., et al. (2016). FUSE: An alternative infrastructure for empowering learners in schools. Proceedings of the 12th international conference of the learning sciences, Singapore.

Stieff, M. (2011). Improving representational competence using molecular simulations embedded in inquiry activities. *J. Res. Sci. Teach.* 48, 1137–1158. doi: 10.1002/tea.20438

Stieff, M., Dixon, B. L., Ryu, M., Kumi, B. C., and Hegarty, M. (2014). Strategy training eliminates sex differences in spatial problem solving in a stem domain. *J. Educ. Psychol.* 106, 390–402. doi: 10.1037/a0034823

Stieff, M., Ryu, M., Dixon, B., and Hegarty, M. (2012). The role of spatial ability and strategy preference for spatial problem solving in organic chemistry. *J. Chem. Educ.* 89, 854–859. doi: 10.1021/ed200071d

Stieff, M., and Uttal, D. (2015). How much can spatial training improve STEM achievement? *Educ. Psychol. Rev.* 27, 607–615. doi: 10.1007/s10648-015-9304-8

Stocklmayer, S. M., Rennie, L. J., and Gilbert, J. K. (2010). The roles of the formal and informal sectors in the provision of effective science education. *Stud. Sci. Educ.* 46, 1–44. doi: 10.1080/03057260903562284

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

Stohlmann, M., Moore, T. J., and Roehrig, G. H. (2012). Considerations for teaching integrated STEM education. *J. Pre-College Eng. Educ. Res.* 2, 28–34. doi: 10.5703/1288284314653

Stull, A. T., Hegarty, M., Dixon, B., and Stieff, M. (2012). Representational translation with concrete models in organic chemistry. *Cogn. Instr.* 30, 404–434. doi: 10.1080/07370008.2012.719956

Suter, L. E. (2016). Outside school time: An examination of science achievement and non-cognitive characteristics of 15-year olds in several countries. *Int. J. Sci. Educ.* 38, 663–687. doi: 10.1080/09500693.2016.1147661

Talmy, L. (2000). *Towards a cognitive semantics: Concept structuring system*. Cambridge, MA: MIT Press.

Taylor, H. A., and Hutton, A. (2013). Think3d!: Training spatial thinking fundamental to STEM education. *Cogn. Instr.* 31, 434–455. doi: 10.1080/07370008.2013.828727

Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., et al. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *Eur. J. STEM Educ.* 3:2. doi: 10.20897/ejsteme/85525

Tillinghast, R. C., Appel, D. C., Winsor, C., and Mansouri, M. (2020). STEM outreach: A literature review and definition. Proceedings of the 2020 IEEE integrated STEM education conference (ISEC), Princeton, NJ, USA.

Tillman, D. A., An, S. A., Cohen, J. D., Kjellstrom, W., and Boren, R. L. (2014). Exploring wind power: Improving mathematical thinking through digital fabrication. *J. Educ. Multimedia Hypermedia*. Waynesville, NCUSA: Association for the Advancement of Computing in Education (AACE). 23, 401–421.

Tracey, T. J. G. (2010). Relation of interest and self-efficacy occupational congruence and career choice certainty. *J. Vocat. Behav.* 76, 441–447. doi: 10.1016/j.jvb.2009.10.013

Uttal, D. H., Amaya, M., del Rosario Maita, M., Hand, L. L., Cohen, C. A., O'Doherty, K., et al. (2013b). It works both ways: Transfer difficulties between manipulatives and written subtraction solutions. *Child Dev. Res.* 2013, 1–13. doi: 10.1155/2013/216367

Uttal, D. H., and Cohen, C. A. (2012). Spatial thinking and STEM education. *Psychol. Learn. Motiv.* 57, 147–181. doi: 10.1016/B978-0-12-394293-7.00004-2

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

Uttal, D. H., Miller, D. I., and Newcombe, N. S. (2013c). Exploring and enhancing spatial thinking: Links to achievement in science, technology, engineering, and mathematics? *Curr. Dir. Psychol. Sci.* 22, 367–373. doi: 10.1177/0963721413484756

Vennix, J., den Brok, P., and Taconis, R. (2018). Do outreach activities in secondary STEM education motivate students and improve their attitudes towards STEM? *Int. J. Sci. Educ.* 40, 1263–1283. doi: 10.1080/09500693.2018.1473659

Verdine, B. N., Golinkoff, R. M., Hirsh-Pasek, K., Newcombe, N. S., Filipowicz, A. T., and Chang, A. (2014). Deconstructing building blocks: Preschoolers' spatial assembly performance relates to early mathematical skills. *Child Dev.* 85, 1062–1076. doi: 10.1111/cdev.12165

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., Nolan, C., and Voyer, S. (2000). The relation between experience and spatial performance in men and women. *Sex Roles* 43, 891–915. doi: 10.1023/a:1011041006679

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

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

Webb, R. M., Lubinski, D., and Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *J. Educ. Psychol.* 99, 397–420. doi: 10.1037/0022-0663.99.2.397

Workman, J. E., and Zhang, L. (1999). Relationship of general and apparel spatial visualization ability. *Cloth. Text. Res. J.* 17, 169–175. doi: 10.1177/0887302X9901700401

Young, C. J., Levine, S. C., and Mix, K. S. (2018). The connection between spatial and mathematical ability across development. *Front. Psychol.* 9:755. doi: 10.3389/fpsyg.2018.00755

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 19 May 2023
DOI 10.3389/feduc.2023.1168459

Check for updates

OPEN ACCESS

EDITED BY
Subramaniam Ramanathan,
Nanyang Technological University, Singapore

REVIEWED BY
Watcharee Ketpichainarong,
Mahidol University, Thailand
José Cravino,
University of Trás-os-Montes and Alto Douro, Portugal

*CORRESPONDENCE
Lisbet Rønningsbakk
✉ lisbet.ronningsbakk@uit.no

RECEIVED 17 February 2023
ACCEPTED 05 May 2023
PUBLISHED 19 May 2023

CITATION
Blackmore K and Rønningsbakk L (2023) Let us explain everything: pupils' perspectives of the affordances of mobile technology during primary science inquiry.
Front. Educ. 8:1168459.
doi: 10.3389/feduc.2023.1168459

Let us explain everything: pupils' perspectives of the affordances of mobile technology during primary science inquiry

Karen Blackmore[1] and Lisbet Rønningsbakk[2]*

[1]Institute of Education, Primary Initial Teacher Education Department, University of Worcester, Worcester, United Kingdom, [2]Department of Education, Faculty of Humanities, Social Sciences and Education, UiT the Arctic University of Norway, Tromsø, Norway

This three-year longitudinal case study focused on the deployment of mobile technology in the form of tablet computers (iPads), during Inquiry Based Science Education (IBSE). The research took place in a larger than average primary school in the West Midlands, UK, which showed a strong commitment to Technology Enhanced Learning (TEL) resulting in iPads being used as an integral learning tool, across the entire curriculum. During the research, pupils in Upper Key Stage Two (10–11 year olds) were observed taking part in science weeks which consisted of intense periods of science inquiry, much of which was child-led. The impact of the embedded use of iPads was monitored by scrutinising pupils' work in the form of multimedia presentations and experimental reports. Pupils' learning behaviours and attitudes to mobile technology were explored through observations and paired interviews. The embedded use of iPads during IBSE was shown to increase science knowledge acquisition and support scientific literacy, recording of processes and aid understanding of working scientifically. Furthermore, iPads were shown to afford opportunities for personalisation of scientific learning experiences and foster collaboration at several levels, factors which were highly valued by the pupils. The outcomes of this study can be used to further inform the refinement of m-learning strategies in primary science and illuminate opportunities for developing the practice of science pedagogues.

KEYWORDS
Inquiry Based Science Education, primary education, iPads, app, scientific literacy, multimedia

Introduction

Pupils who are being educated now in primary schools will contribute to a legacy of scientific knowledge and societal judgements in the 22nd century, and as such will require strong scientific literacy skills and a deep understanding of all three disciplines. This aspiration is somewhat problematic in England given that the profile of primary science in the curriculum is undoubtedly diminished (Wellcome Trust, 2011, 2014; OFSTED, 2011, 2014). Hence, UK stakeholders have argued that this situation requires urgent remediation in terms of policy change in order to improve opportunities within primary science education

(Economic and Social Research Council, 2013; Science Community Representing Education [SCORE], 2013; Wellcome Trust, 2013) and encourage learners to see themselves as potential scientists of the future. Since science capital (based on social capital theory and defined as the tendency of pupils to believe they can make a scientific contribution to society), is thought to build during the formative primary stages of education (STEM Learning, 2019); it is essential to commit to further refine the effectiveness of science pedagogy throughout early formal education for all learners (Harlen, 2010; Lievesley, 2014). It is argued (Duschl et al., 2007) that in a technologically advanced society, learning should not be focused merely on the recall of facts but rather the development of a deep understanding of the nature of science and its associated methodology. Therefore, the emphasis should be placed firmly on exploration and research skill development, enabling future generations of scientists to create global solutions (United Nations Education Scientific and Cultural Organisation [UNESCO], 2017).

There is a consensus of opinion that in order for science education to be effective it needs to be undertaken from the perspective of pupils' experiences of the world around them (Worth, 2010) and to some degree be responsive to the findings that pupils value both guided and independent means to study (Lau et al., 2017). In this way, pupils' natural curiosity can be harnessed during exploratory inquiry, which can lead to more formalised learning. Following the initial phase of scientific curiosity, it is necessary for pupils to plan, record, analyse and share their inquiries, traditionally this was achieved using paper and pen methods. In the last decade there has been a change in the way pupils communicate with each other and teachers in the primary classroom, with the prevalent use of mobile technology (m-technology). Whilst many studies across the primary curriculum including mathematics (Hilton, 2018), literacy (Lynch and Redpath, 2014; Browder et al., 2017; Bergeson and Rosheim, 2018) and the humanities (Monem et al., 2018) have highlighted the possibilities of deploying m-technology, in the majority of cases these have focused on realising learning outcomes and teacher perspectives (Boon et al., 2021).

Studies specifically focusing on the pedagogical affordances of m-technology during science inquiry are less prevalent and generally do not concern the attitudes of the pupils themselves to this mode of learning (Wang et al., 2022). Hence, this study focuses on primary age pupils in upper Key Stage Two (years 5 and 6, pupils aged 10–11 years old) where they are establishing the beginnings of scientific inquiry, and in addition experiencing a crucial formative stage during which they may start to regard themselves as scientists of the future.

Pedagogical framework: Inquiry Based Science Education (IBSE) and an inquiry-based approach

The key focus of this study relies on the use of an established pedagogy, namely Inquiry Based Science Education (IBSE), supplemented by the deployment of mobile technology (m-technology). Specifically, the research aims to explore the potential synergies between IBSE as facilitated by m-technology, to enable deeper understanding of scientific inquiry in the primary classroom, from the perspective of the learner.

There is a large body of work devoted to IBSE, and in general the evidence suggests it is a highly effective method of engaging learners with scientific discovery and key scientific concepts (Berg et al., 2003; Leonard et al., 2009; Furtak et al., 2012; van Uum et al., 2016). However, closer examination of the literature reveals that the situation is not totally clear-cut (Rönnebeck et al., 2016). In his meta-analysis, Anderson (2002) highlights whilst the effectiveness of this type of pedagogy is generally agreed, researchers tend to define IBSE with respect to their own particular research context. Therefore, there can be subtle differences in what researchers believe constitutes inquiry-based learning opportunities. For this study the definition honed by Linn et al. (2004), p. 4 is adopted, since it concisely summarises the key features of IBSE as "The intentional process of diagnosing problems, critiquing experiments, and distinguishing alternatives, planning investigations, researching conjectures, searching for information, constructing models, debating with peers, and forming coherent arguments."

This is an extensive delineation which adheres to the epistemology of advanced scientific study and research. Initial contemplation might suggest this as an overly ambitious approach for primary education but in fact, both the nature and process associated with science have already been demonstrated to be effectively explored within inquiry-based approaches (Bianchini and Colburn, 2000). In breaking down the definition further, it can be seen that the first four components: problematising, critiquing experiments, distinguishing alternatives and subsequent planning investigations, are all integral elements of the "working scientifically" portion of the National Curriculum for Science in England (DfE, 2013, 2015).

Such are the benefits of this pedagogic approach, that inquiry-based approaches have been adopted by curriculum designers in Australia, USA, Middle East, Asia, and Europe (Abd-El-Khalick, 2003; National Curriculum Board of Australia, 2009; National Research Council, 2012), as a means of fostering and developing scientific knowledge and skills. Inquiry-based science teaching affords learners opportunities to not only engage in valuable *hands-on/minds on* learning experiences (Minner et al., 2010) but also develop science specific content knowledge (Sandoval, 2005) and cultivate a far-reaching appreciation of the nature of scientific discovery (Schwartz et al., 2004).

For an inquiry based approach to be effective, scientific discovery needs to be made accessible to all pupils and not perceived to be only undertaken by people of certain cultural or educational backgrounds (Gibbs, 2014). Learners also need to be able to engage with a range of development opportunities to hone their scientific skills or in simplistic terms work like a scientist (Archer et al., 2010; DfE, 2013, 2015). A range of inquiries, including fair or comparative testing, pattern recognition, secondary research and problem-solving should be undertaken to allow pupils to improve these key skills which represent "working scientifically" (Turner et al., 2011) in a similar way to a scientist. García-Carmona (2020, p. 448) citing National Research Council (1996, pp. 137–138), develops this idea further and advocates positioning inquiry-based approaches as being explicitly linked to "what scientists do" and highlights "students should evaluate their own results or solutions to problems, as well as those of other children" and consider "alternative explanations" both of which, mirror the processes that scientists undertake during dissemination of their ideas and peer review. Encouraging children to see themselves as scientists, also confers the added advantage of increasing pupils' science capital which can be a positive indicator for sustained science study in post-16 education (Archer et al., 2015).

A significant number of components of the outlined process, also represent typical constructivist-based learning activities within the primary classroom (Hackling and Prain, 2008; OFSTED, 2011). In addition, proponents of this approach point out that an inquiry based science is highly accessible to primary school pupils and can be carried out almost everywhere: in non-specialist classrooms, outdoors and even at home with parental supervision at a low cost (Blacklock, 2012; Lee, 2012). It is important to note however that practical experience alone is not enough to secure deep understanding of methodology or scientific principles (Bransford et al., 2000) but rather "wrap around" thinking as a result of initial research, trialing, analysis and discussion is required to substantiate learning during inquiry (Pedaste et al., 2015).

Potential barriers during investigational primary science

In England, upper KS 2 pupils (10–11-year olds) are introduced to the scientific method through the concept of "fair testing" which encourages the development of a logical and systematic approach (DfE, 2013) as a component of IBSE. Teachers can act as mediators of this approach by encouraging careful observation and the formulation of authentic child-led questions (Keys and Bryan, 2001; OFSTED, 2011, p. 15). They can also model a collaborative approach to inquiry in their classrooms, supported by extensive opportunities for in-depth discussion and peer review (OFSTED, 2011, p. 638). Teachers arguably have a pivotal role in sustaining pupil interest and engagement by prompting deep thinking using extensive classroom discourse (Crawford, 2000).

Whilst IBSE teaching approaches have been judged to be efficacious (Chang and Mao, 1999; Anderson, 2002; Wilson et al., 2009; Minner et al., 2010) there are inherent practical problems associated with such methodologies with young learners. Central to these challenges is the fact that primary aged pupils are often engrossed in the practical elements of inquiries and hence reluctant to pause and document their observations. To facilitate effective learning, these outcomes need to be captured in a time efficient manner, so that pupils do not miss key events. They also need to assimilate the process of planning and carrying out investigations, so they can develop confidence in setting up future inquiries. Often in the overriding quest to see what happens, the implementation of a systematic and fair testing regime is overlooked and potentially forgotten. Mobile technologies that are able to capture in-the-moment scientific phenomena and processes (Clark and Luckin, 2013) could therefore be valuable tools in improving effective metacognition of operational scientific experimentation. Additionally, they afford learners endless opportunities for research and analysis, by collating and processing a range of data including multi-media outputs (Clark and Luckin, 2013) and acting as a vital conduit for the requirement to communicate their ideas confidently (van der Graaf, 2020). M-technology deployment during science inquiry may also afford teachers the opportunity to share key experimental outcomes with children after the investigation itself, when there is less time pressure (Burden et al., 2016).

Technology enhanced learning using mobile technologies

Classic TEL approaches have been used to support scientific inquiry for some time according to Kim et al. (2007, p. 1017) who describe their potential as being able to provide: 'technology-enhanced, student-centred, flexible opportunities' during inquiry processes. However, due to the physical space required by stand-alone computers and laptops, they are challenging to integrate closely into scientific investigations, resulting in their use being restricted to initial research phases or subsequent analysis. In contrast, there is the more recent phenomenon of rapid adoption of mobile learning (m-learning) in primary schools, much of it mediated by the use of tablet computers (Boon et al., 2021). Research in different age phases including early years education, Lynch and Redpath (2014); Burden and Kearney (2016); Song and Wen (2018) suggests that due to their portability, multi-modal nature and ease of use, m-technologies may give rise to additional unique learning opportunities. Such learning prospects have been described by the iPAC (i Personalise, Authenticity and Collaboration) framework as developed by Burden and Kearney (2016). This model describes how m-technologies can be used by learners to personalise experiences, improve or make explicit the authenticity of operation and encourage effective collaboration during experiential learning.

Each of the three main constructs (P, A, and C) are characterised by sub constructs e.g., personalisation of the learning experience, which may give rise to an increase in agency; where pupils record and later discuss outcomes in their own way. This positive psychological aspect in conjunction with the improved motivation associated with inquiry-based approaches (Justice et al., 2009) may act as a powerful enabler in science learning. Through multimedia experiences (documentaries, YouTube video-clips, virtual reality laboratories) pupils are becoming more aware of what real everyday science looks like. Arguably they are beginning to be critical and make judgements about the legitimacy of scientific experimentation and as such becoming more discerning in experimental design. By giving pupils the choice of embedding m-technologies during the scientific processes, they may consider their learning experiences more credible and from their perspective, authentic.

M-technology as a means of fostering scientific collaboration

In terms of collaboration, scientific discovery and development has always been associated with discourse and peer review. Even at a young age, learners are capable and open to social constructivist learning mechanisms via the processes of data gathering and sharing in science (Dunn et al., 2016 and Furman et al., 2018). Several science pedagogues (Ford and Forman, 2006; Allen, 2010; Harlen, 2010) highlight how crucial it is for learners to be afforded opportunities to engage in high quality learning conversations with their peers and teachers about their scientific discoveries. During, and as a result of these conversations misconceptions may be revealed and challenged, resulting in improved understanding and remodelling of conceptually difficult ideas (Allen, 2010). A systematic review of the use of m-technology during science learning by Afikah et al. (2022) highlights the increased social interactions evident using such an approach and enhanced opportunities for pupils to develop their problem-solving skills.

Research questions

Since the aforementioned literature suggests that pupils are able to take the lead in the classroom, collaborate together and socially construct their emerging scientific ideas, this study seeks to explore in depth the enactment of IBSE using m-technology from a learner perspective and focuses on the following research questions:

(i) In what ways may iPads be used to support inquiry-based learning including planning, recording and working scientifically?
(ii) What learning behaviours and attitudes do primary aged pupils display toward iPad use during IBSE?

Materials and methods

Research context

The research setting was a larger than average primary school (400+ pupils, taught by approximately 40 teaching and support staff) in the West Midlands region of England. The school was newly built prior to the study and as a result, pupils had access to a wide range of high-quality information and communication technology (ICT) equipment, including tablet computers (iPads). The research focused on Upper Key Stage 2 (UKS2) pupils (10–11 year olds) where each of the 60 pupils in year 5 (10-year olds) had access to their own iPad tablet computer and each of the 60 pupils in year six (11-year olds) shared an iPad between two learners. Since the pupils used iPads every day in a range of lessons, they were very familiar with operating them and aware of the options to use different apps for writing, video capture and drawing etc.

The study took place during science weeks, which were characterised by periods of intense scientific inquiry by all pupils. During these weeks the normal school timetable was suspended and rather than it represent the usual mix of curriculum subjects including core subjects (English and mathematics) and foundations subject (e.g., geography, history or religious education), the pupils exclusively studied science for approximately 25 h per week. These science weeks involving at least eight different inquiries, took place in addition to the usual science lessons at six monthly intervals over the space of three consecutive academic years. This initiative aimed to enable pupils to improve their scientific skills and knowledge bases, in a subject which has seen a reduced profile in primary schools since the removal of statutory testing in 2009 (Wellcome Trust, 2014).

The analysis of pupils' work was aided by the school's sophisticated data management system, which allowed tracking of attainment within a large database of pupils' work. Due to the integrated learning and assessment infrastructure using m-technology, this school was judged to be highly suitable for this case study.

Expectations of pupils' learning during the science weeks

In addition to learning appropriate content knowledge during their science study across all three science disciplines, pupils were required to acquire disciplinary knowledge in terms of *working scientifically*. According to the requirements of the National Curriculum for England: Science Programmes of Study (DfE, 2015, p. 25) pupils learn to follow the scientific method. Hence the pupils worked logically and sequentially through the scientific processes of planning, taking measurements, recording data and results, making predictions and reporting and presenting findings. In addition, with respect to their developing scientific literacy, they were required to use appropriate scientific language and illustrations to communicate their ideas. In summary, during the science weeks pupils were expected to engage with all the science inquiries by working practically and documenting and sharing their investigational outcomes using the iPads.

Expectations of teachers and leaders during science weeks

Teachers and science/phase leaders adopted an inclusive, child led (Siry et al., 2012) stance within their classrooms, in keeping with the ethos of the school. Specifically, they aimed to articulate and model that anyone can act like and become a scientist, and that all perspectives on approaches and modes of IBSE were valued. Teachers encouraged pupils to ask questions of themselves and others, discuss their points of view and work carefully and methodically. If a pupil required help or guidance that was given by a teacher but they strived to enable progress whilst encouraging independent learning. During the science weeks the teachers and leaders liaised extensively with the researchers to exchange their ideas on what was working well and what if anything, needed development. They also gave feedback to pupils whilst showcasing their work, using interactive white boards to share inquiry reports and presentations.

Research ethics

The research was undertaken from a perspective of minimising the researcher footprint and according to best practice as defined by British Educational Research Association (2018). In short, participants were approached via the deputy head of school who acted as a gatekeeper for the study and issued effective communication of the purpose and expectations of the study. Both parental written consent and on-going pupil assent to take part was sought. Participants were regularly reminded of their right to withdraw from the study, although none chose to do so. Confidentiality and anonymity were preserved at all times.

Participants and sampling

During each phase of the study across three academic years, samples of work were taken from thirty pupils from each year group (sixty pupils in total), which represented data from half of all learners in UKS2. The purposive sample represented a mix of gender, ethnicity and academic attainment. In addition, twelve pairs of pupils were interviewed from each cohort, each year to determine their perspectives and attitudes to learning. The data from pupils who worked both in pairs or singly were used for data analysis purposes.

Data collection instruments

(i) Semi-structured observations of pupils and teachers during the scientific inquiries

Semi-structured observations took place during all phases of scientific inquiry for both year groups, resulting in over forty hours of observational material being acquired. The schedule was designed using key learning facilitators identified by the seminal work of Burden and Younie (2014) from the Mapping Educationalist Specialist Know-How (MESH) guide in using iPads in the classroom. Such facilitators were: undertaking scientific processes and social interaction, initiating own enquiry strategies, exploiting the mobility of iPads and engagement and motivation during the scientific inquiry. Descriptions of each type of learning behaviour were produced and discussed at length prior to observation to ensure reliability of the measure. Piloting in another school with similar aged pupils and class sizes was used to check for validity.

(ii) Semi-structured interviews with pupils

These were designed according to the strategy outlined by Kvale (1996). Initial general questions were asked about the scientific enquiry as a means of creating rapport and ensuring pupil comfort, for example: "Can you tell me about your work in science this week?" These questions were followed by more probing questions exploring in-depth the use of iPads by the pupils. For example: "Did the iPad help you understand the science? Can you give me an example please?" Pupils were also asked if they used more traditional methods to process and record data, for example handwritten tables.

The interviews which lasted approximately 30 min, were completed by asking questions to discern attitudes toward using iPads during work in science. Frequently during the interviews, responses were read back to the pupils to check for meaning and understanding on the behalf of the researcher. In addition, the majority of pupils chose to bring their iPad to the interviews and illustrated their answers by demonstrating their learning activities recorded as multimodal documents. Notes were taken to record these responses and screen shots taken of the pupils' work.

Data analysis

The pupil interviews (with accompanying notes detailing when pupils had referred to specific examples of their work on the iPads) were recorded on an additional iPad, transcribed verbatim and, together with the classroom observations, the data were analysed separately using thematic analysis. Systematic analysis was achieved by using NVivo 11 software to support a five-stage process according to the premises of Braun and Clarke (2006). The first phase involved familiarisation of the data by reading and re-reading multiple times. This stage was followed by the generation of initial codes (termed nodes in NVivo 11) from across the entire dataset. Memos were made to ensure the contexts of the codes were retained. Cluster analysis was then undertaken to create thematic maps and highlight potential themes which were then reviewed. A round of confirming codes and then re-coding was undertaken to confirm emergent themes. In some cases, themes were collated into sub themes or merged leading to the definition of final themes. Internal reliability of the data was improved by another independent researcher being involved in the analysis and rechecking codes and themes. The trustworthiness of the findings was enhanced by triangulation (Guba and Lincoln, 1994) across the data from observations, pupils' work and paired interviews. This process is summarised in **Figure 1**.

Results

Using the data from the pupil and teacher observations and interviews, three key themes emerged with respect to the first research question as follows:

- *Use of apps on iPads to increase science knowledge acquisition and scientific literacy*
- *Use of iPads to support planning and recording of inquiry processes*
- *Use of iPads to increase understanding of working scientifically*

Each theme will be described in detail in turn using illustrative observations, examples of pupils' work and quotes in the following section.

Use of apps on iPads to increase science knowledge acquisition and scientific literacy

One of the key strategies pupils used to increase their science knowledge during science weeks and beyond, was the use of small stand-alone applications (apps) on the iPads, with the types of apps used, falling into two broad categories. The first type was science specific and encompassed background knowledge on each topic for example plant and human biology. The second type of app was more generic and allowed processing and reporting of the investigational data, examples include *Popplet*, *iMovie*, *Comic Strip*, *Explain Everything*, and *Book Creator*. Pupils were given the opportunity to choose from a range of apps prior

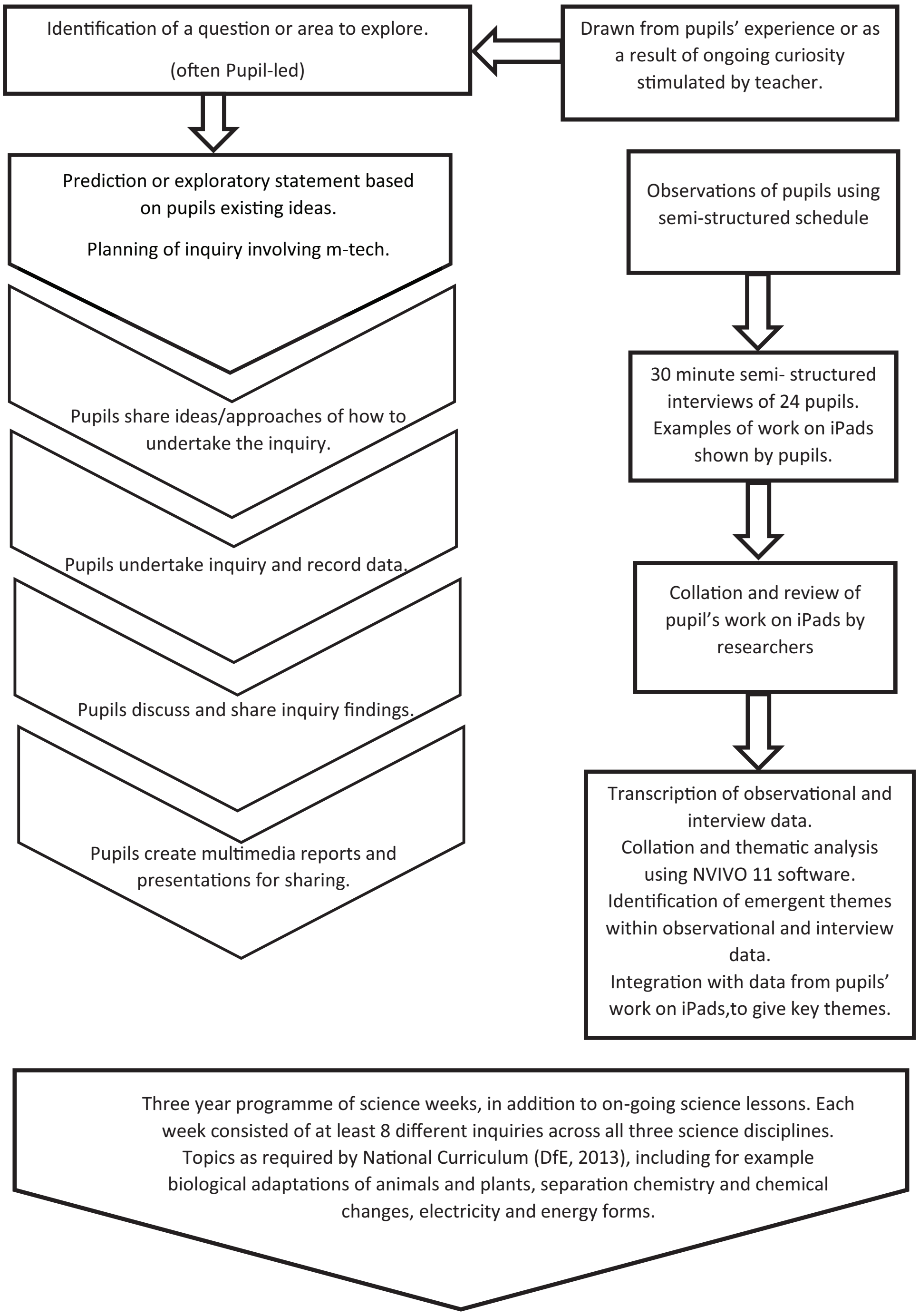

FIGURE 1
The methodological flow and timeline of the research (informed by Harlen, 2021).

to the science weeks and were able to access the apps at any time during their learning. Interview data revealed approximately one third valued this opportunity to address their own learning needs, as illustrated by the following quote from a year 5 girl:

> I sometimes like to use science apps when I don't understand something, I ask my teacher if that is OK, and I use time after or before lessons. I liked the *Trees are Best* app as it helped me learn a lot of interesting things about trees. I also used *Plant Parts* to help me become more confident with naming the parts of a flower.

All pupils confirmed at interview that they were self-reliant in terms of accessing information using science specific apps, they were often observed before and after investigational sessions sharing science-based facts and demonstrating different aspects of the apps to each other.

Whilst apps containing science specific material were undoubtedly useful for a significant proportion of learners, the vast majority of app usage centred around the use of generic apps. The first of these apps, *iMovie*, was used extensively to video capture real time recordings of specific experimental outcomes. The material was then reviewed and edited by the pupils prior to inclusion in their reports. For example, in the case of producing a Poly Vinyl Acetate (PVA) polymer, termed by the pupils "Flubber," they were observed to video record the bouncing of "Flubber" balls alongside a 1 m ruler, in order to determine both the height of bounce and the number of bounces. Pupils were then observed to calculate mean averages of both variables which demonstrated links with their scientific knowledge of measurements.

Generic apps were also used at a deeper level to collate and process information from the scientific investigations. Both *Book Creator* and *Comic Strip* were used in a narrative manner to detail and record experimental methodology and results. High quality chronicling of both the experimental processes and results was evident. For example, analysis of an experimental write-up using the app *Comic Strip* contained photographs of the red cabbage pigment extraction step followed by filtration and the subsequent use of the extract as a pH indicator. Similarly, pupils used the app *Explain Everything* to produce an e-book logically and systematically detailing an investigation into the action of plant enzymes, on a jelly protein layer. Video clips obtained using the app *iMovie* were embedded into the experimental report to supplement the pupils' observations.

It was observed on multiple occasions that the pupils appeared confident to manipulate and process a wide range of material in order to construct mind-maps, e-books, and comic strips to illustrate their scientific knowledge acquisition. One aspect of this knowledge acquisition was strong scientific literacy, supported by online resources. It was frequently observed that during the initial research phases of the inquiry, pupils often used online dictionaries during the construction of their investigational reports. For example the significance of the phrase "the liver detoxifies and metabolises" was explored for meaning by a year 6 girl, who immediately shared her understanding with the whole class. She emphasised the meaning of both the words *detoxify* and *metabolism* in a manner which suggested she understood two of these key functions of the liver.

Another good example of how pupils approached brand new terminology, was the use of the scientific term "enzymes" which several pupils found problematic. In this case pupils engaged with their teachers in modelling the actions of digestive enzymes on fat and protein molecules. They photographed the simplified cardboard molecules the teacher had created of the food molecules and how they were broken down by enzymes (modelled by scissors). This resulted in extensive discussion between groups of pupils who went on to use other multimedia apps to explore the meaning of new term. As a result, one year five pupil drew a cartoon of an enzyme molecule breaking down a long polymer to illustrate his understanding of the catabolic nature of enzymes.

The pupils also understood the importance of using text to describe the key scientific evidence contained within the multimedia and frequently ascribed each legend to a particular graphic, e.g. "The hummingbird does not have to suck up liquid into its beak, instead there are two grooves called troughs which draw the liquid through them."

Similarly, key words such as "magma", "pressure", "lava," and "eruption" were used to correctly annotate and narrate an animation showing volcanic activity in a pupil's e-book.

Use of iPads to support planning and recording of inquiry processes

Three quarters of the pupils routinely used the mind mapping app *Popplet* to plan and analyse their data. This tool allows the user to build up the mind map by quickly creating many separate "bubbles" (termed *popples*) capable of storing data or ideas. The pupils were observed to create highly detailed planning frameworks consisting of multiple steps and large arrays of data organised in tables, using this app. For example, during the water transpiration investigation using celery and dyed water, pupils created a table using popples to record each celery sample in each environmental condition (low temperature/light, room temperature, and moderate temperature/light and darkness). The following day, when transpiration had taken place, pupils were observed to photograph each celery sample and place the image in their pre-prepared recording grid.

Pupils also described at interview how they referred to and refined mind maps during the whole inquiry, as articulated by a year 5 boy as follows:

> I really like Popplet, it helps me plan what I am going to do. If I forget, I go back and check through the Popples. . ..Let me show you, this is the plan I did for my science experiment and here are all the Popples I used to make my results table.

Another feature of app usage by the pupils was the very proficient use of photographs, graphics and video capture material to illustrate scientific reports. Over three quarters of the pupils demonstrated proficient use of the app *Book Creator* to record in great detail experimental outcomes. For example, this was seen in a digestion investigation, where the e-books created by the pupils contained predictions pertaining to how the fruit slices would affect the consistency of the jelly. In order to fully record the digestion process over the course of the experiment, pairs of pupils used the iPad time lapse photography feature to record the experimental outcomes. An edited video clip of this process was then observed to be embedded in the resultant e-book alongside a well formulated narrative of the experiment. The pupils were abundantly aware of the need to be systematic in their recording as illustrated by the following exchange:

> Researcher: "How did you know what to take the photographs of?"
> Pupil: "Well some of our friends told us and others were just common sense basically. If you're going to put it in the method, take a picture of it. This is the red cabbage at the start; these are all the pictures we took when we went through the process to get the purple juice out."

Observational and interview data confirmed that pupils were able to effectively use the apps on the iPads to record a range of static and dynamic experimental outcomes and capture data in a way that they could then include in their scientific reports and presentations.

Use of iPads to increase understanding of working scientifically

It was clear that the use of iPads during inquiry had been facilitative for underpinning the scientific inquiry process. Analysis of results and scientific discussion seemed to be particularly supported by this approach. Pairs of pupils described how important it was when analysing a chemical sample to compare the result with that of a control and they achieved this by using the camera app. This was illustrated by the following exchange during interview:

> Pupil: "The one at the back is a control one" [shows researcher photograph of test tube containing plant extract].
> Researcher: "Oh okay, why do you need a control tube?"
> Pupil: "To see if the extract has changed colour or not."

In addition, the pupils were observed to compare their experimental outcomes embedded within multimedia files with other pupils, prompting scientific dialogue. Subsequent discussion appeared to prompt a host of higher order questions for example:

> Pupil: "Mine dissolved faster than yours, did you stir faster?"
> Pupil: "Did you use the same amount of water as us?. Yours is darker red?"

During the seed inquiry it was evident that the pupils were comparing their prediction for each fruit in terms of number and appearance of the seeds (which they had inserted in a single *popple*), with the actual findings following cutting open the fruit. This demonstrated they knew it was important to compare and contrast scientific predictions with actual results.

In terms of integrated iPad use facilitating in-depth understanding of scientific inquiry, the following example is illustrative:

> Researcher: "How did you use the iPad to learn about animal adaptations?"
> Pupil: "I've done a *Popplet* about the bird beaks" [shows the researcher a mind map showing the initial planning of the bird beak investigation together with Popples to record how much food was collected over which time periods, for each beak].
> Researcher: "Then what did you do?"
> Pupil: "Then we got to do an investigation . . . with tools that were like actual beaks . . . and then we had some stars in the water we had to time ourselves . . .. to see how many we could get into the cup with different tools. We then watched a video clip about the adaptation of the bird beaks."
> Pupil: "After our experiment I researched different bird diets and beaks and found out about humming birds having long beaks so they can get nectar out of plants."
> Researcher: "And what have you learnt about bird beaks?"
> Pupil: "How the birds would like survive, with the limited food resources."

These exchanges illustrate the strong understanding pupils had derived of working scientifically using iPads during their inquiries and subsequently carrying out additional research using online resources.

Learning behaviours and attitudes to the use of iPads in science inquiry

Three themes emerged in response to the second research question from the semi-structured observations, interviews and pupil work as follows:

- *Personalisation of the science learning experience*
- *Positive attitudes to the benefits of using the affordances of mobile technology*
- *Collaboration with peers and teachers during scientific inquiries*

Personalisation of the science learning experience

The majority of pupils took the opportunities afforded them by iPads to customise their learning experience during the science weeks. One of the pair of year five pupils described how they worked together initially when researching and planning and then went on to document their own inquiries individually:

Pupil: "At the beginning we made a little mind-map about what we would like to know about this particular project. We did lots of research on the internet about different plants in different places, this was interesting."

The pair then proceeded to describe how they had each customised their own ebook by using different backdrops, fonts, colours and arrangements of images and texts to create their own unique record. They then revealed when they had finished, they used the iPads to peer-review.

Some pupils made unusual app choices; one girl described how she used *PicCollage*, an app usually reserved for social sharing of photographs, to compare images from the detergent investigation:

> Pupil: "It can help you because you can screenshot it [test sample] . . . you can write a caption to it as well. . . then you'll have a page of pictures."

She then proceeded to explain how she carefully compared each test sample after washing with the control sample, to give a score to each detergent in terms of cleaning capacity.

The pupils were also articulate in describing how they liked to demonstrate their agency as learners. They were honest and reflective in expressing their preferences to use mobile technology as exemplified by the following year 6 pupil:

"I know some schools or some places they don't use iPads.. . . you have to write it all down on paper and I'm really slow at writing all of it down on paper."

Another year 6 girl commented how much easier it was to edit her work efficiently using the iPads:

"It was a lot easier to make changes when I reviewed the writing, and I could add pictures and make a really good document."

Pupils were highly organised during the analysis stage with pairs critiquing the overall shape and content of the mind maps during learning. For example, during the sheep's lung dissection, one pair arranged the *popples* in a zigzag formation in order to put the respiratory system organs in a logical space and add more information. They described how this choice had enabled them to effectively handle their data:

Pupil: "We found it easier to spread our ideas out, instead of just writing one massive paragraph about it."

Overall, these quotes illustrate how personal learning preferences could be accommodated by using the m-technology.

Positive attitudes to the benefits of using the affordances of mobile technology

Often during interviews, the pupils described how they used the mobility of the iPads to great effect, for example during experimentation outside:

Pupil: "I took pictures when we were outside for science . . . we brought them inside and I could use the pictures to remind me what we had found."

Others commented on how the ease with which the iPad could be positioned to take photographs was very helpful:

Pupil: "I took it from that angle. . .because you could see it had sunk instead of being on top."

Due to the portability of the iPads, photographs of scientific processes were taken quickly and efficiently. One group used the mobility of the iPads to photograph up close, how the Poly Vinyl Acetate (PVA) glue polymer was used to fill a mould, and changed in shape over the course of the investigation. The pupils also used the apps to improve the accuracy of their observations, e.g., when polymer testing, they video-captured a ruler alongside the sample to record the snapping point.

This was also the case when recording outcomes which were technically difficult to see in great detail, as in the case of the static electricity investigation. Several observations indicated the pupils appeared fascinated to see the small plastic pieces attracted to the charged rod and used *iMovie* to film the pieces being lifted into the air. They then used the iMovie clip to illustrate their experimental results.

The only rarely reported dissatisfaction was pupils describing being "disappointed" when work was lost due to software updating processes or network problems.

Collaboration of peers and teachers during scientific inquiries

Collaboration was a strong theme identified within the pupils' mode of learning when using iPads. Throughout the entire science weeks, they used material stored on the iPads and the class cloud server to initiate scientific discussion, gather and share data.

Conversations were often initiated by pupils exchanging research or their own experimental data. Overall, the collaborative approach was highly valued by the pupils as illustrated from the following quotes from year 5 and 6 Pupils.

Year 5 pupil: "We share ideas. Sometimes when I am struggling, I ask someone else and they show me what they did with the iPad."
Year 5 pupil: "We use them all the time; we show our teacher our data or others.

And from a year 6 girl:

"I love sharing the work when it is mirrored on the board. When my work is shown I feel very happy, and I like to get ideas from other people."

Some pupils appreciated the support of their partner during data gathering, a year 5 girl, new to the school said:

"We used the iPads to record scores for the bird beak experiment. I am not confident in working by myself, I like working with someone just to make sure that I've got all my stuff right."

It was also clear that the pupils enjoyed viewing each other's *iMovie* video clips of experimental findings and discussing outcomes:

Pupil: "I like listening to my friend's recordings on their iPads and then we share ideas."

During the explorations into circuits, lots of small group interactions were observed, for example between three pairs of pupils who showed each other video clips of trial and error type experiments. The pupils appeared happy to accept differences in approaches and outcomes. Indeed, some went as far as saying they valued critique from their peers.

The pupils saw data sharing in this way as a win-win situation and termed the phrase "Magpie" to indicate using someone else's approach when they liked it:

Pupil: "When you've done some work, you can show it to everyone. If you're stuck, you can Magpie . . . you can be shown what to do, not told."

In the case of formative assessment, the teachers used Apple TV on the interactive white board [IWB] to share the pupil reports containing photographs, videos and diagrams. The pupils were observed to frequently collaborate together during this time of lesson plenaries and to maintain focus by discussing their results and approaches of others.

In terms of collaboration with the teachers and leaders, scientific inquiry reports were deposited by pupils on the shared cloud, used for cohort assessment purposes. The pupils valued the resulting additional feedback on their experimental work.

> Year 5 boy: "Sometimes, the year five leader organises our work and gives us feedback so then we can get better and learn a bit more."

This quote demonstrates the pupils understood how sharing their work facilitated understanding of their learning not just with their peers but with teachers and year group leaders.

Conclusion to findings

In summary, integration of the qualitative data from observations, interviews, and pupils' work confirmed that the use of iPads within the science inquiries appeared to significantly enhance learning opportunities. Pupils were seen to move seamlessly from initial research, to planning, to experimental phases, whilst simultaneously adding to and refining their ideas. They used the large capacity of iPads to store, analyse data and present their findings, using an array of multimedia modes. Overall, the pupils exhibited highly positive attitudes and collaborative learning behaviours when using m-learning approaches.

Discussion

Reviewing the findings, it is clear that the researchers, teachers and pupils believed there were significant advantages to m-learning during scientific inquiry. This approach undoubtedly gave rise to innovative opportunities to undertake authentic scientific learning across all three science disciplines. Echoing the future scenarios of science learning and skill acquisition, proposed by Kearney et al. (2022) and originally advocated by Duschl et al. (2007) and United Nations Education Scientific and Cultural Organisation [UNESCO] (2017), this study foregrounds both the autonomous and connective affordances conferred by m-learning. Specifically, it highlights the capacity to deploy m-learning strategies during *all* phases of scientific inquiry in contrast to the use of classical stand-alone computers, which were restricted to the initial phases of research and planning and then analysis (Kim et al., 2007).

Bespoke learning resources in the form of specific apps were shown to be effective in affording pupils' opportunities to research and develop their own science subject knowledge in a pupil-centred manner as advocated by Worth (2010). This echoes the work of Pedaste et al. (2015) who asserted that practical inquiry alone could not ensure deep learning but rather initial research was an important component of pupil metacognition and sense making in science.

The findings show clear resonances with the work of Hilton (2018) with respect to iPads facilitating engagement with mathematical processing. There were several occasions where pupils were observed to use the unique capabilities of the m-technology; for example, in videorecording followed by in-depth measurements, to refine the accuracy and analysis of their inquiries. Since the future of many aspects of scientific work rely on the deployment of mobile and remote technologies for measurements, this bodes well for the skills of the upcoming workforce.

In terms of overall effectiveness, an important point of interest was the large positive impact on the development of pupils' digital scientific literacy through the use of both generic and specific apps, during all phases of the scientific inquiry. The pupils were able to move easily from one mode of scientific representation (e.g., text, graphics, or video material) to another and incorporate a range of data into their reports. Prain and Waldrip (2010):2 argue that it is a great advantage for learning, when students are able to interpret "modal diversity in representations of science concepts." Indeed, it is not just different modes of representation that are of importance in fostering understanding, but rather the development of multiple ways of communicating and believing in science (Moje, 2008).

In order to (re)-construct scientific beliefs, learners need opportunities to engage in discussion with teachers, in addition to learning through sharing and critically evaluating their own and peers' scientific inquiry (Ford and Forman, 2006; Yacoubian and BouJaoude, 2010). This involves the development of scientific literacy skills (Lee et al., 2013; Bergeson and Rosheim, 2018) and the correct usage of key science terminology (Shanahan and Shanahan, 2008), a component highlighted in this study, by the sharing and meaning making of key scientific terms using on-line dictionaries and the extensive deployment of appropriate vocabulary in the pupil presentations and reports. Overall, the use of iPads during inquiry afforded the pupils opportunities to develop their scientific literacy toward the end of KS 2, prior to transition to secondary education. It could be argued that this is of a particular significance, since the confident use of scientific language during argument making has been shown to be pivotal (van der Graaf, 2020).

From an inclusive practice perspective, in agreement with several studies (Miller et al., 2013; Knight and Davies, 2016; Browder et al., 2017), the use of iPads during scientific inquiry fostered accessibility for a range of learners, especially those with specific literacy needs. Such personalised strategies appeared to benefit learners who voiced the opinion that they liked to learn in their own way. Pupils with specific literacy needs described how m-learning allowed them to approach the documentation of the inquiries positively and enable them to make a prompt start and not lag behind the rest of the class. The science specific apps also allowed for effective differentiation of learning opportunities; with more able pupils accessing detailed information on specific topics e.g., evolutionary change, whilst others used the apps to review and reinforce their learning of conceptually challenging concepts (Allen, 2010) e.g., electrical circuits. Several pupils highlighted how the iPads allowed them to keep up with the pace of the lesson. This is wholly appropriate given evidence that pupils appreciate *both* independent and guided means of study (Lau et al., 2017) and echoes the study by Furman et al. (2018), who revealed even younger learners were pleased by opportunities to display autonomy, when using tablet computers during IBSE.

The pupils' attitudes to using tablet computers during IBSE were overwhelmingly positive, one pair described how working with iPads enabled them to work at a different level and how they became excited when designing their own multimedia reports. Several pairs expressed how they preferred using iPads as they helped them learn more efficiently and registered disappointment that their diminished use at the culmination of upper KS2. These positive attitudes are largely in agreement with the findings of a systematised literature review into m-learning using iPads, undertaken by Boon et al. (2021), which explored the experiences of primary age pupils using m-technology to access the curriculum overall. In the specific case of science learning, the findings confirm those of others (Looi et al., 2011; Dunn et al., 2016; van Deursen et al., 2016) that m-technology is a valuable tool to facilitate ISBE. When utilising iPads during inquiry, pupils particularly valued using them to exchange ideas and undertake collaborative research together in agreement with Rønningsbakk (2020), arguably two of the most important aspects of the ways scientists work. Since pupils are motivated to use iPads, an important consideration during IBSE (Justice et al., 2009) and this pedagogical approach has been shown to be broadly effective (Clark and Luckin, 2013; Burden and Kearney, 2016; Hong et al., 2017), it would be advisable for teachers and curriculum designers to integrate m-technology into future scientific inquiry. Indeed, it is advocated that pupils' voices should be given primacy even at the expense of teachers' concerns when designing effective m-learning opportunities (Dunn et al., 2016), since they are agents of their own learning both in and beyond the classroom.

The exact role of teachers in inquiry based approaches is flexible however it has been suggested that teachers should afford their learners opportunities to explore "conceptual, epistemic, social and/or procedural domains of scientific knowledge" (van Uum et al., 2016, p. 450). In other words, pupils need to access the key scientific ideas, understand how scientific knowledge is constructed and appreciate how the nature of working scientifically and scientific debate, supports the process. Key to all these domains is the process of questioning, a skill highly represented in primary science teaching practice (Earle, 2014). This study has shown m-learning can be used to support pupils formulate and share questions with their pers and teachers. Teachers can also provide further exciting stimuli for pupils to illicit their ideas and prompt further questioning and discussion, which can be part of a formative assessment strategy (Constantinou et al., 2018). This could well be achieved by deploying mobile technology as a means of accessing engaging resources. In terms of effective teacher practice during IBSE, it has been suggested that primary school teachers may not see themselves substantively as "science teachers" and consequently may exhibit concerns about their science content knowledge (Harlen, 2021). However, research (Mellander and Svärdh, 2018) has shown that teachers who display strong pedagogic content knowledge of inquiry, deploy approaches associated with improved pupil learning. These arguments have clear implications for IBSE to be explicitly explored in teacher education programmes (Strat and Jegstad, 2022) in conjunction with the affordances of m-learning, as part of an inquiry based approach.

Since pupils articulated eloquently their preference for close collaboration and sharing of their scientific findings and expertise with peers and teachers alike, these findings can be used to inform future pedagogical approaches. They resonate significantly with the metastudy undertaken by Afikah et al. (2022) across all phases of the science curriculum, where collaboration was shown to be a key element of engagement and shared problem-solving. Specifically, this paper argues for the empowerment of individual pupils to realise their full potential within scientific inquiry and build their science capital, by using m-learning to enable them to collaborate with their peers and in turn, contribute to wider scientific discourse.

This socio-cultural potential of m-learning echoes Dunlop et al. (2015) arguments for building "communities of enquiry" during IBSE and builds on the work of Felix et al. (2005) who advocates addressing a question of common concern by harnessing different points of view. The added advantage of social constructivist learning is that it models many of the processes adopted during authentic scientific discovery in the real world (Burden and Kearney, 2016). Indeed, many features of the science learning opportunities afforded by this approach are compatible with the creation of technology enhanced multi-modalities as advocated by Zhang et al. (2010) and Lynch and Redpath (2014). These result in learners being made more aware of their metacognition and allow potential scientists of the future, to see for themselves how knowledge is constructed in a scientifically robust and collaborative manner.

Limitations

There were some limitations, arguably the most prominent being the use of a single, longitudinal, school case study. It is accepted that this school was at the forefront of good practice with respect to the use of iPads across the curriculum and therefore findings cannot be considered to be wholly generalisable. However, with the proliferation of Apple Distinguished Schools and Multi Academy Trusts in England, where pedagogical approaches are often shared within clusters of schools (Apple, 2021; Baxter and John, 2021), this study aimed to illustrate what is possible in terms of the affordances of m-learning supported IBSE and hence share effective practice. With respect to inherent limitations of the data collection tools, these are well known; pupils may behave differently when they know they are being observed and during interview may seek to provide answers which they believe the researcher will value (Einarsdóttir, 2007). However, the researchers feel both these limitations were reduced by the fact that an insider/outsider positionality (Moore, 2012) was adopted due to the strong relationships developed with the staff and learners over a three-year period. The limitations of work scrutiny are that it is partly subjective and open to bias. However, in this case, since the pupils brought their work on iPads to the interviews and illustrated their evolving ideas, it is believed the researchers gained a robust understanding of pupils' thinking processes as advocated by Kellett (2006).

Conclusion

It is highly evident that m-learning is becoming wider spread in primary schools (Boon et al., 2021) and it is hoped that the key findings of this study will inform discourse around the use of m-technology to enhance primary science pedagogy from the perspectives of the learners. Further research would include the exploration of key findings, for example improved scientific recording and attitudes of pupils to their own metacognition, in a range of schools both within the U.K and internationally. In a global scientific community where an increasingly diverse range of modalities are being used to document and analyse during scientific enquiry (Asch et al., 2018), this approach could make a significant contribution to the engagement and education of the scientists of the future.

Data availability statement

The datasets presented in this article are not readily available because of confidentiality. Requests to access the datasets should be directed to k.blackmore@worc.ac.uk.

Ethics statement

The studies involving human participants were reviewed and approved by College for Arts, Humanities and Education Ethics Committee, University of Worcester. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

KB wrote the first draft of the manuscript. LR discussed research and publications as a lens through which to view the manuscript, critiqued and wrote specific sections of the second draft, and contributed additional current literature. Both authors edited subsequent drafts together and contributed to manuscript revision.

Acknowledgments

We thank all the participants in this study who generously shared their knowledge and views of learning in primary science.

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.

Publisher's note

References

Abd-El-Khalick, F. (2003). Inquiry in science education: International Perspectives. *Cult. Comp. Stud.* 88, 397–419.

Afikah, A., Astuti, S., Suyanta, S., Jumadi, J., and Rohaeti, E. (2022). Mobile Learning in Science Education to Improve Higher-Order Thinking Skills (HOTS) and Communication Skills: A Systematic Review. *Int. J. Adv. Comp. Sci. Applic.* 13, 698–704.

Allen, M. (2010). *Misconceptions in primary science*. Berkshire: McGraw-Hill.

Anderson, R. (2002). Reforming Science Teaching: What Research Says about Inquiry. *J. Sci. Teach. Educ.* 13, 1–12.

Apple (2021). *Apple Distinguished Schools: Centres of Leadership and Educational Excellence 2021–2024 Overview*. Cupertino, CA: Apple.

Archer, L., Dawson, E., DeWitt, J., Seakins, A., and Wong, B. (2015). "Science capital": A conceptual, methodological, and empirical argument for extending bourdieusian notions of capital beyond the arts. *J. Res. Sci. Teach.* 52, 922–948.

Archer, L., DeWitt, J., Osbourne, J., Dillon, J., Willis, B., and Wong, B. (2010). 'Doing' science versus 'being' a scientist: Examining 10/11-year-old school pupils' constructions of science through the lens of identity. *Sci. Educ.* 94, 617–639.

Asch, M., Moore, T., Badia, R., Beck, M., Beckman, P., Bidot, T., et al. (2018). Big data and extreme-scale computing: Pathways to Convergence-Toward a shaping strategy for a future software and data ecosystem for scientific inquiry. *Int. J. High Perform. Comput. Applic.* 32, 435–479.

Baxter, J., and John, A. (2021). Strategy as learning in multi-academy trusts in England: strategic thinking in action. *Sch. Leadersh. Manage.* 41, 290–310.

Berg, C., Bergendahl, V., Lundberg, B., and Tibell, L. (2003). Benefiting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus an open-inquiry version of the same experiment. *Int. J. Sci. Educ.* 25, 351–372.

Bergeson, K., and Rosheim, K. (2018). Literacy, equity, and the employment of iPads in the classroom: A comparison of secure and developing readers. *Int. J. Educ. Math. Sci. Technol.* 6, 173–181.

Bianchini, J., and Colburn, A. (2000). Teaching the nature of science through inquiry to prospective elementary teachers: A tale of two researchers. *J. Res. Sci. Teach.* 37, 177–209.

Blacklock, K. (2012). Science on a tight budget. *Primary Sci.* 121, 5–7.

Boon, H., Boon, L., and Bartle, T. (2021). Does iPad use support learning in students aged 9–14 years? A systematic review. *Aust. Educ. Res.* 48, 525–541. doi: 10.1007/s13384-020-00400-0

Bransford, J., Brown, A., and Cocking, R. (eds) (2000). *How people learn: brain, mind, experience, and school*. Washington, DC: National Academy Press.

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101.

British Educational Research Association (2018). *Ethical Guidelines for Educational Research*. London: British Educational Research Association.

Browder, D., Root, J., and Wood, L. (2017). Effects of a story-mapping procedure using the iPad on the comprehension of narrative texts by students with autism spectrum disorder. *Focus Autism Other Dev. Disabil.* 32, 243–255. 611387 doi: 10.1177/1088357615

Burden, K., Aubusson, P., Brindley, S., and Schuck, S. (2016). Changing knowledge, changing technology: Implications for Teacher Education Futures. *J. Educ. Teach.* 42, 4–16.

Burden, K., and Kearney, M. (2016). Future Scenarios for Mobile Science Learning. *Res. Sci. Educ.* 46, 287–308.

Burden, K., and Younie, S. (2014). *Using iPads effectively to enhance learning in schools*. London: MESHGuides.

Chang, C., and Mao, S. (1999). Comparison of Taiwan science students' outcomes with inquiry- group versus traditional instruction. *J. Educ. Res.* 92, 340–346. doi: 10.1186/s13054-016-1208-6

Clark, W., and Luckin, R. (2013). *What the research says: iPads in the classroom. London Knowledge Lab*. London: Institute of Education University of London.

Constantinou, C., Tsivitanidou, O., and Rybska, E. (2018). "What Is Inquiry-Based Science Teaching and Learning?," in *Professional Development for Inquiry-Based Science Teaching and Learning*, eds O. E. Tsivitanidou, P. Gray, E. Rybska, L. Louca, and C. P. Constantinou (Cham: Springer), doi: 10.1007/978-3-319-91406-0_1

Crawford, B. (2000). Embracing the essence of inquiry: New roles for science teachers. *J. Res. Sci. Teach.* 37, 916–937. doi: 10.1002/1098-2736(200011)37:9<916::AID-TEA4>3.0.CO;2-2

DfE (2013). *National curriculum in England: science programmes of study*. England: DfE.

DfE (2015). *National Curriculum in England: Science Programmes of Study: Statutory guidance*. England: DfE.

Dunlop, L., Compton, K., Clarke, L., and McKelvey-Martin, V. (2015). Child-led enquiry in primary science. *Education* 43, 462–481. doi: 10.1080/03004279.2013.822013

Dunn, J., Gray, C., Moffett, P., and Mitchell, D. (2016). 'It's more funner than doing work': pupils' perspectives on using tablet computers in the early years of school. *Early Child Dev. Care* 188, 819–831.

Duschl, R., Schweingruber, H., and Shouse, A. (eds) (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.

Earle, S. (2014). Formative and summative assessment of science in English primary schools: evidence from the Primary Science Quality Mark. *Res. Sci. Technol. Educ.* 32, 216–228.

Economic and Social Research Council (2013). *What influences participation in science and mathematics?' A briefing paper from the Targeted Initiative on Science and Mathematics Education (TISME)*. London: Economic and Social Research Council.

Einarsdóttir, J. (2007). Research with pupils: methodological and ethical challenges. *Eur. Early Child. Educ. Res. J.* 15, 197–211. doi: 10.1136/bmjopen-2020-044143

Felix, V., Mena, L., and Fischer, R. (2005). *Teaching Pupils to Think*, 2nd Edn. Cheltenham: Nelson.

Ford, M., and Forman, E. (2006). Refining disciplinary learning in classroom contexts. *Rev. Res. Educ.* 30, 1–33. doi: 10.1002/bmb.21387

Furman, M., De Angelis, S., Dominguez Prost, E., and Taylor, I. (2018). Tablets as an educational tool for enhancing preschool science. *Int. J. Early Years Educ.* 27, 6–19.

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

García-Carmona, A. (2020). From Inquiry-Based Science Education to the Approach Based on Scientific Practices. *Sci. Educ.* 29, 443–463. doi: 10.1007/s11191-020-00108-8

Gibbs, K. (2014). *Diversity in STEM: What It Is and Why It Matters*. Berlin: Scientific American.

Guba, E., and Lincoln, Y. (1994). "Competing paradigms in qualitative research," in *Handbook of qualitative methods*, eds N. Denzin and Y. Lincoln (London: Sage).

Hackling, M., and Prain, V. (2008). *Impact of Primary Connections on Students' processes literacies of science and science processes*. Canberra: Australian Academy of Science.

Harlen, W. (2010). *Principles and Big Ideas of Science Education*. Hatfield: Association for Science Education.

Harlen, W. (2021). *The Case for Inquiry-based Science Education – IBSE*. Trieste: The Inter Academy Partnership (IAP).

Hilton, A. (2018). Engaging primary school students in mathematics: Can iPads make a difference? *Int. J. Sci. Math. Educ.* 16, 145–165.

Hong, J., Hwang, M., Tai, K., and Tsai, C. (2017). An exploration of students' science learning interest related to their cognitive anxiety, cognitive load, self-confidence and learning progress using inquiry-based learning with an iPad. *Res. Sci. Educ.* 47, 193–212.

Justice, C., Rice, J., Roy, D., Hudspith, B., and Jenkins, H. (2009). Inquiry-based learning in higher education: Administrators' perspectives on integrating inquiry pedagogy into the curriculum. *High. Educ.* 58, 841–855. doi: 10.1007/s10734-009-9228-7

Kearney, M., Schuck, S., and Burden, K. (2022). Digital pedagogies for future school education: promoting inclusion. *Irish Educ. Stud.* 41, 117–133.

Kellett, M. (2006). "Pupils as active researchers: Using engagement with research process to enhance creativity and thinking skills in 10-12 year-olds," in *Paper presented at the British Educational Research Association Annual Conference*, Warwick.

Keys, C., and Bryan, L. (2001). Co-constructing inquiry-based science with teachers: *Essential* research for lasting reform. *J. Res. Sci. Teach.* 38, 631–645.

Kim, M., Hannafin, M., and Bryan, L. (2007). Technology-Enhanced Inquiry Tools in Science Education: An Emerging Pedagogical Framework for Classroom Practice. *Sci. Educ.* 91, 1010–1030.

Knight, K., and Davies, R. (2016). Using a mobile dichotomous key iPad application as a scaffolding tool in a museum setting. *Interact. Learn. Environ.* 24, 814–828.

Kvale, S. (1996). *Interviews: An Introduction to Qualitative Research Interviewing*. Thousand Oaks, CA: Sage Publications.

Lau, W., Lui, V., and Chu, S. (2017). The use of wikis in a science inquiry-based project in a primary school. *Educ. Technol. Res. Dev.* 65, 533–553.

Lee, A. (2012). Development of a parent's guide for the Singapore primary science curriculum: Empowering parents as facilitators of their pupils' science learning outside the formal classrooms. *Asia Pac. For. Sci. Learn. Teach.* 13, 1–27.

Lee, O., Quinn, H., and Valdés, G. (2013). Science and language for English language learners in relation to next generation Science standards and with implications for common core state standards for English language arts and mathematics. *Educ. Res.* 42, 223–233.

Leonard, J., Boakes, N., and Moore, C. (2009). Conducting Science Inquiry in Primary Classrooms: Case Studies of Two Preservice Teachers'. *Inquiry Based Pract. J. Element. Sci. Educ.* 21, 27–50.

Lievesley, T. (2014). What does Ofsted say? *Editor. Primary Sci.* 133, 22–23.

Linn, M., Davis, E., and Bell, E. (2004). *Internet environments for science education*. London: Lawrence Erlbaum.

Looi, C., Zhang, B., Chen, W., Seow, P., Chia, G., Norris, C., et al. (2011). 1:1 mobile inquiry learning experience for primary science students: a study of learning effectiveness. *J. Comput. Assist. Learn.* 27, 269–287. doi: 10.1111/j.1365-2729.2010.00390.x

Lynch, J., and Redpath, T. (2014). 'Smart' technologies in early years literacy education: A meta-narrative of paradigmatic tensions in iPad use in an Australian preparatory classroom. *J. Early Child. Liter.* 14, 147–174. doi: 10.1177/1468798412453150

Mellander, E., and Svärdh, T. (2018). Inquiry-based learning put to the test: medium-term effects of a science and technology for children programme. *Rev. Educ.* 6, 103–141. doi: 10.1002/rev3.3109

Miller, B., Krockover, G., and Doughty, T. (2013). Using iPads to teach inquiry science to students with a moderate to severe intellectual disability: A pilot study. *J. Res. Sci. Teach.* 50, 887–911. doi: 10.1002/tea.21091

Minner, D., Levy, A., and Century, J. (2010). Inquiry-based science instruction-what is it and does it matter? Results from a research synthesis years 1984 to 2002. *J. Res. Sci. Teach.* 47, 474–496. doi: 10.1002/tea.20347

Moje, E. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *J. Adolesc. Adult Liter.* 52, 96–107.

Monem, R., Bennett, K., and Barbetta, P. (2018). The Effects of Low-Tech and High-Tech Active Student Responding Strategies during History Instruction for Students with SLD. *Learn. Disabil.* 16, 87–106.

Moore, J. (2012). A personal insight into researcher positionality. *Nurse Res.* 19, 11–14. doi: 10.7748/nr2012.07.19.4.11.c9218

National Curriculum Board of Australia (2009). *Shape of the Australian Curriculum: Science*. Canberra: National Curriculum Board of Australia.

National Research Council (1996). *Science education standards*, Washington, DC: The National Academies Press.

National Research Council (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.

OFSTED (2011). *Successful science: an evaluation of science education in England 2007–2010*. England: OFSTED.

OFSTED (2014). *Maintaining curiosity: a survey into science education in schools*. England: OFSTED.

Pedaste, M., Mäeots, M., Siiman, L., de Jong, T., van Riesen, S., Kamp, E., et al. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educ. Res. Rev.* 14, 47–61. doi: 10.1016/j.edurev.2015.02.003

Prain, V., and Waldrip, B. (2010). Representing Science Literacies: An Introduction. *Res. Sci. Educ.* 40, 1–3. doi: 10.1007/s11165-009-9153-x

Rönnebeck, S., Bernholt, S., and Ropohl, M. (2016). Searching for a common ground – A literature review of empirical research on scientific inquiry activities. *Stud. Sci. Educ.* 52, 161–197.

Rønningsbakk, L. (2020). "Digital Natives and Educational Traditions. What Changes When Exchanging Textbook Content with Internet Search?," in *Innovative Technologies and Learning. ICITL 2020. Lecture Notes in Computer Science*, eds T. Huang, T. Wu, J. Barroso, F. E. Sandnes, P. Martins, and Y. Huang (Cham: Springer).

Sandoval, W. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Sci. Educ.* 89, 634–656. doi: 10.1002/sce.20065

Schwartz, R., Lederman, N., and Crawford, B. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Sci. Teach. Educ.* 88, 610–645. doi: 10.1002/sce.10128

Science Community Representing Education [SCORE] (2013). *Resourcing Practical Science in Primary Schools*. London: SCORE.

Shanahan, T., and Shanahan, C. (2008). Teaching Disciplinary Literacy to Adolescents: Rethinking Content-Area Literacy. *Harvard Educ. Rev.* 78, 40–61.

Siry, C., Ziegler, G., and Max, C. (2012). "Doing science" through discourse-in-interaction: Young children's science investigations at the early childhood level. *Sci. Educ.* 96, 311–326. doi: 10.3310/hta25220

Song, Y., and Wen, Y. (2018). Integrating Various Apps on BYOD (Bring Your Own Device) into Seamless Inquiry-Based Learning to Enhance Primary Students' Science Learning. *J. Sci. Educ. Technol.* 27, 165–176. doi: 10.1007/s10956-017-9715-z

STEM Learning (2019). *Science Capital: making science relevant*. Mumbai: STEM Learning.

Strat, T. T. S. and Jegstad, K. M. (2022). Norwegian teacher educators' reflections on inquiry-based teaching and learning in science teacher education. *J. Sci. Teacher Educ*. doi: 10.1080/1046560X.2022.2125623

Turner, J., Keogh, B., Naylor, S., and Lawrence, L. (2011). *It's Not Fair - Or is It? A Guide to Developing Children's Ideas Through Primary Science Enquiry*. Richmond: Millgate House.

United Nations Education Scientific and Cultural Organisation [UNESCO] (2017). *Science and technology centres: reducing the gap between knowledge and action*. Paris: UNESCO.

van der Graaf, J. (2020). Inquiry-based learning and conceptual change in balance beam understanding. *Front Psychol.* 11:1621. doi: 10.3389/fpsyg.2020.01621

van Deursen, A., Ben Allouch, S., and Ruijter, L. (2016). Tablet use in primary education: Adoption hurdles and attitude determinants. *Educ. Inf. Technol.* 21:971. doi: 10.1007/s10639-014-9363-3

van Uum, M., Verhoeff, R., and Peeters, M. (2016). Inquiry based science education: towards a pedagogical framework for primary school teachers. *Int. J. Sci. Educ.* 38, 450–469.

Wang, J., Tigelaar, D., Zhou, T., and Admiraal, W. (2022). The effects of mobile technology usage on cognitive, affective, and behavioural learning outcomes in primary and secondary education: A systematic review with meta-analysis. *J. Comput. Assist. Learn.* 39:12759. doi: 10.1111/jcal.12759

Wellcome Trust (2011). *Primary Science Survey Report*. London: Wellcome Trust.

Wellcome Trust (2013). *The Deployment of Science and Maths Leaders in Primary Schools*. London: Wellcome Trust.

Wellcome Trust (2014). *Primary Science: Is It Missing Out? Recommendations for reviving primary science*. London: Wellcome Trust.

Wilson, D., Joseph, A., Taylor, J., Kowalski, S., and Carlson, J. (2009). The relative effects and equity of inquiry-based and commonplace science teaching on students' knowledge, reasoning, and argumentation. *J. Res. Sci. Teach.* 47, 276–301. doi: 10.1002/tea.20329

Worth, K. (2010). *Science in Early Childhood Classrooms: Content and Process. STEM in Early Education and Development Conference*. Cedar Falls: University of Northern Iowa.

Yacoubian, H., and BouJaoude, S. (2010). The Effect of Reflective Discussions Following Inquiry-Based Laboratory Activities on Students' Views of Nature of Science. *J. Res. Sci. Teach.* 47, 1229–1252. doi: 10.1002/tea.20380

Zhang, B., Chee-Kit, L., Seow, P., Chia, G., Wong, L.-H., Chen, W., et al. (2010). Deconstructing and reconstructing: transforming primary science learning via a mobilised curriculum. *Comput. Educ.* 55, 1504–1523. doi: 10.1016/j.compedu.2010.06.016

frontiers | Frontiers in Psychology

TYPE Original Research
PUBLISHED 02 June 2023
DOI 10.3389/fpsyg.2023.1145795

OPEN ACCESS

EDITED BY
Karen Blackmore,
University of Worcester, United Kingdom

REVIEWED BY
Antonio Luque,
University of Almeria, Spain
Isain Zapata,
Rocky Vista University, United States

*CORRESPONDENCE
Dongil Kim
✉ dikimedu@snu.ac.kr

RECEIVED 16 January 2023
ACCEPTED 16 May 2023
PUBLISHED 02 June 2023

CITATION
Kim T and Kim D (2023) Chilly climate perceived by female engineering undergraduates: an exploratory study using concept mapping.
Front. Psychol. 14:1145795.
doi: 10.3389/fpsyg.2023.1145795

Chilly climate perceived by female engineering undergraduates: an exploratory study using concept mapping

Tanhui Kim and Dongil Kim*

Department of Education, Seoul National University, Seoul, Republic of Korea

Introduction: Women still being a minority in engineering majors, they are reported to face discriminatory treatment in a collegiate environment. "Chilly climate," referring to such a sexist environment, may have a negative impact on women's mental health, academics, and careers. But, what exactly is it that female students in engineering perceive as chilly, and how chilly is it? This study aimed to explore the chilly campus climate perceived by female undergraduate engineering students in South Korea using the concept mapping method.

Methods: Semi-structured interviews were conducted with 13 participants enrolled for more than four semesters at four-year coeducational universities. After extracting 52 representative statements, the participants were asked to classify them according to content similarity and rate the influence of each statement on their perception of the chilly climate. For concept mapping analysis, multidimensional scaling analysis (ALSCAL), hierarchical cluster analysis (Ward's method), and non-hierarchical cluster analysis (K-means method) were performed.

Results: Fifty-two statements were extracted under the following four clusters: (i) "Exclusion and alienation inherent in the culture (Cluster 1)," (ii) "Sexual objectification and lack of gender sensitivity (Cluster 2)," (iii) "Male-centered academic situations (Cluster 3)," and (iv) "Prejudice and generalization (Cluster 4)." A concept map was two-dimensional: an X-axis named "context dimension," with "task: academic" and "non-task: social" at both ends, and a Y-axis named "sexism dimension", having "explicit" and "implicit" at both ends. The order of higher scores in the influence rating is as follows: Cluster 2, Cluster 3, Cluster 1, and Cluster 4.

Discussion: This study is significant because it conceptualizes the subjective experience of minorities in a collegiate environment and provides influence rating results for prioritized measures. The findings will be helpful in formulating educational policies, psychological counseling, and social advocacy activities. Future research should target larger populations, and cover more diverse cultures, majors, and age groups.

KEYWORDS

chilly climate, female engineering students, sexism, gender microaggression, concept mapping, South Korea

1. Introduction

Engineering has traditionally been a male-dominated field, and this has never changed. The percentage of female students enrolled in undergraduate engineering programs has steadily increased from 18.25% (2010) to 22.67% (2018) (National Science Foundation and National Center for Science and Engineering Statistics, 2021), but compared to men, they are still extremely underrepresented. In South Korea, only 21.4% of women enter the engineering department, which is much smaller than that in natural science (50.9%), social and human sciences (57.6%), and medicine (67.7%) (Korea Foundation for Women in Science, Engineering and Technology, 2021). As women are scarce, masculine culture and favoritism toward men are prevalent in the engineering field (Min and Lee, 2005; Duberley and Cohen, 2010; Richman et al., 2011; Hatmaker, 2013; Kim et al., 2016). For example, faculties have lower expectations for female students, provide less academic encouragement or support, and make sexist remarks (Min and Lee, 2005; Park, 2019; Roper, 2019). In Korea, the representative engineering culture includes a military-like hierarchy, obscenity, sexual harassment, and high-intensity drinking (Min and Lee, 2005; Kim et al., 2016; Park, 2019).

"Chilly climate," a term introduced by Hall and Sandler (1982), refers to receiving such sexist treatment in the collegiate environment. It denotes how traditionally masculine fields are unwelcoming or hostile to women owing to differential treatment by professors or peers in the school environment. Examples of chilly climate include verbal and overt aspects such as disparaging comments about women or sexist jokes and nonverbal and subtle aspects such as paying more attention to men's comments or waiting longer for men than women to answer a question (Hall and Sandler, 1982). Discriminatory treatment of women exists not only in the classroom but also in various situations such as academic and career counseling, laboratory or fieldwork, group projects, internships, school safety, student autonomy and cultural activities, economic support, and curricula (Hall and Sandler, 1984; Janz and Pyke, 2000). This male-centered atmosphere in schools induces female students to perceive a chilly climate, and has been steadily reported (Goldman, 2012; Cabay et al., 2018; Jensen and Deemer, 2019; Park, 2019; Roper, 2019).

This climate reinforces the marginalization of female engineering students and may have a negative impact on their mental health, academics, and careers. Female students experience negative emotions such as alienation, helplessness, anger, frustration, depression, anxiety, and stress (Hall and Sandler, 1982; Cammaert, 1985; Janz and Pyke, 2000; Walton et al., 2015), and worry about sexism, stereotype threats, and joining the group (O'Brien et al., 2015). Additionally, it causes reduced satisfaction and confidence, deflection, identity confusion, and limited self-expression (Min and Lee, 2005; Chu, 2008; Park, 2019).

Academically, such an environment makes female students doubt their abilities and internalize their devaluation, making them hesitant to participate in academic activities or build relationships with the faculty (Hall and Sandler, 1982; Hall, 2016). Moreover, it has a negative impact on cognitive development (Pascarella et al., 1997), major satisfaction, and self-efficacy (Jeong et al., 2008).

As the chilly environment for women continues to exist in engineering-related workplaces (Makarem and Wang, 2020), it acts as a career barrier for female students (Settles et al., 2006). Female students' anticipation of career barriers was related to reduced career aspirations and planning (Cardoso and Moreira, 2009; Schuster and Martiny, 2017), and low confidence in future employment, job maintenance, and promotion (Do, 2008). This leads to a "leaky pipeline"; that is, women switch out of Science, Technology, Engineering, and Math (STEM) fields, causing a vicious cycle in which women exist as minorities, leaving fewer female role models to follow (Min and Lee, 2005; Young et al., 2013; Renn, 2014).

Despite the negative effects of the chilly climate on women, men who form the majority negate that sexism is a concern or are indifferent, and some even believe that they have a high sense of gender sensitivity (Hall, 2016; Roper, 2019). This implies that discriminatory words and actions against female students may occur in a subtle or benevolent manner. This phenomenon can be regarded as a "microaggression"; that is, something not particularly intentional but done subtly to make minorities feel uncomfortable (Sue et al., 2007). As microaggression is often perceived as a minor problem, it can lead to feelings of self-doubt, confusion, or alienation (Nadal and Haynes, 2012). In addition, because it is difficult to deal with microaggression promptly or actively, it can be more harmful than overt discrimination (Sue et al., 2007; Dumont et al., 2010). Continuous exposure to microaggression can lead to feelings of low self-esteem, anxiety, and depression (Nadal and Haynes, 2012).

Given that female engineering students experience the negative impact of a chilly climate on their mental health or realization of their potential, it is necessary to thoroughly understand the experiences of these women for more effective psychological intervention and prevention. However, little in-depth research has been conducted on what they perceive as unfavorable to women and the exact circumstances that affect them negatively. Some recent studies on subtle sexism in STEM have focused more on ethnicity or race (Lee et al., 2020; Miles et al., 2020; Marshall et al., 2021), or gender-based microaggression experienced by faculties or professionals (Yang and Carroll, 2018; Makarem and Wang, 2020; Kim and Meister, 2022). In South Korea, research on chilly climates is rare. According to Hwang (2020), 75% of the research on female engineering students deals with career- or academic-related topics, and 75% of the research methods used are quantitative. Allan and Madden (2006) stated that an appropriate research method is crucial when studying a subtle subject, and quantitative methods, such as surveys, may not yield a complete picture.

Therefore, the present study aimed to investigate the chilly climate perceived by female undergraduate engineering students using the concept mapping method, which is useful for identifying the cognitive structure underlying a specific phenomenon, and can present results with minimal researcher subjectivity (Trochim, 1989; Paulson et al., 1999). This study assessed the subjective experience of female engineering students using a qualitative method, processed the data using quantitative statistical methods, and presented the results with a concise pictorial representation. Moreover, the degree of influence of the reported content on the perception of chilly climate was measured to draw more meaningful conclusions.

2. Materials and methods

2.1. Participants

The participants of this study were female undergraduates enrolled for four semesters or more in the science and engineering department at a 4 year coeducational university. To ensure diversity, students from seven universities, including universities in Seoul, the metropolitan area, and national universities in provincial areas, were selected. A list of science and engineering majors with a female enrollment rate under 30%, based on statistics from the Ministry of Education (2019), was presented. This study was approved by the Institutional Review Board (IRB) of Seoul National University (IRB No. 2012/001-023), and all participants provided written informed consent.

According to Jeong et al. (2008), the longer female students have studied, the more they are influenced by the environment in variables such as major satisfaction and career. Moreover, when research participants were recruited from December 2020 to January 2021, their classes and academic life were shifted to a non-face-to-face manner due to the coronavirus disease 2019 (COVID-19) pandemic. Therefore, students' experiences with less than four enrollment periods were considered difficult to generalize and were excluded.

Thirteen students from the engineering department were selected as study participants. The number of study participants was determined by referring to Trochim (1989), who stated that there is no strict guideline for concept mapping and 10 to 20 participants are workable. The average age of the participants was 21.4 years old (SD = 1.7 years) and the average semester enrollment period was 5.8 semesters (SD = 1.8 semesters).

2.2. Procedure

This study utilized the concept mapping method (Kane and Trochim, 2007) to explore the chilly climate perceived by female undergraduate students in the engineering field. Figure 1 shows the research procedure.

2.2.1. Step 1: preparation

In the preparation stage, a focus question was developed for the interviews. Sample questions were created through literature research, and three experts in qualitative research reviewed the questions. The final focus question selected and used in interviews was as follows: "In what situation (e.g., atmosphere, experience) were you treated differently from men or felt subtle sexism in school as a female undergraduate student in the science and engineering field?"

After obtaining IRB approval, research participants were recruited via recruitment posts on each school's website.

2.2.2. Step 2: generation of statements

Individual interviews ($n = 13$) were conducted to generate statements. The interviews were conducted between January 5 and 24, 2021. Before the interview, the purpose of the study and interview method were explained in detail through phone calls and notices. If they agreed to participate in the study voluntarily, they were requested to sign the consent form and send it via e-mail before the interview. One day before the interview date, they were sent a focus question via e-mail to allow them to think about the topic in advance.

The interviews were conducted non-face-to-face using Zoom video conference due to COVID-19. After a simple greeting and self-introduction, the participants were allowed to freely talk about the focal question. The researcher tried to promote the production of ideas by asking additional questions related to the topic along with the parts that need to be embodied or clarified. The interview was recorded with the consent of the participants, and took approximately 50–60 min per person.

A total of 183 ideas were extracted from the interviews. According to Kane and Trochim (2007), fewer than 100 statements were appropriate for concept mapping. Thus, the number of statements was abbreviated according to the method of integrating duplicate statements and removing elements mentioned by less than two people (Bedi, 2006). This repeated grouping and reclassifying process was performed by the researcher and two other Ph.D. graduates, and supervised by two professors in the educational counseling department. Fifty-two representative statements were derived using the terms used by the study participants as much as possible.

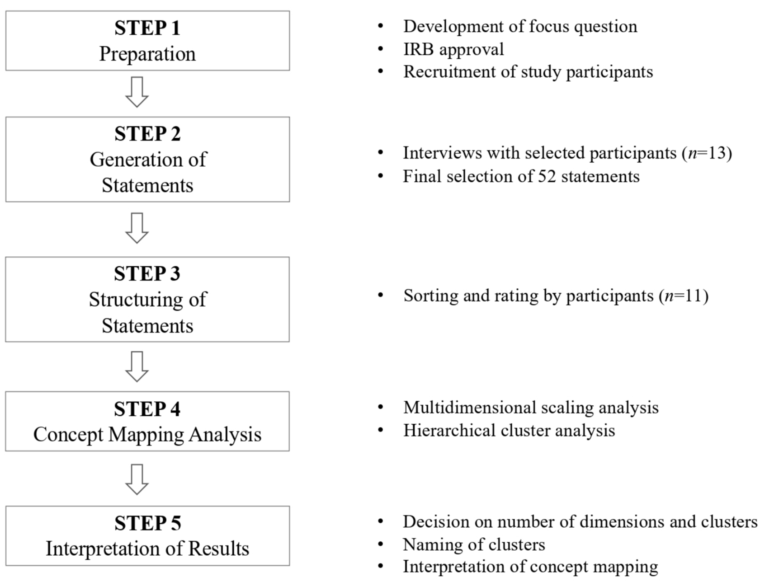

FIGURE 1
Concept mapping research process.

2.2.3. Step 3: structuring of statements

In this stage, the participants were asked to classify 52 representative statements according to content similarity and rate the influence of each statement on the perception of chilly climate. Owing to personal reasons, only 11 out of 13 participants performed these two tasks from February 25 to March 3, 2021.

Paper cards are generally utilized in similarity classification for concept mapping in offline face-to-face settings. However, this study was conducted when social distancing was recommended due to the COVID-19 pandemic, and there were difficulties in conducting face-to-face interviews with participants from provincial areas. Thus, Microsoft PowerPoint (PPT) was devised as a suitable alternative owing to its familiarity and accessibility to the study participants. Before conducting Step 3, the researcher tested with two participants whether the task instructions were easy to understand and whether tasks using the PPT were easy to perform.

In the similarity classification task, participants were given a PPT file with all 52 representative statements written on the first slide. Each statement was written in a small square card and laid out without overlap so that all statements could be seen at a glance. The participants were then asked to classify these statements by considering a single slide as one group. They were to create as many groups or slides as desired. Referring to Trochim (1989), participants were instructed to place each statement once, and not to place all statements on a single slide or to place one statement on one slide. After classification, participants were asked to name each group.

In the influence rating stage, participants were asked to evaluate on a 7-point Likert scale how much the content of each statement affects their perception of the chilly climate on and off campus (1 = *very little*, 7 = *very much*).

2.2.4. Step 4: concept mapping analysis

For concept mapping analysis, a Similarity Matrix was first created from the classification data. A 52 × 52 binary square matrix was created; 0 was coded for two statements placed in the same group, and 1 was coded for those that did not. A group similarity matrix was produced by summing 13 individual similarity matrices. A multidimensional scale analysis (ALSCAL) on SPSS 22.0 statistical program was used to analyze the number and meaning of the appropriate dimension.

Next, the coordinate values derived from the ALSCAL were used for hierarchical cluster analysis (Ward's method) and non-hierarchical cluster analysis (K-means method). Ward's method is suitable for interpreting clusters in conceptual diagrams to classify clusters based on distance (Kane and Trochim, 2007). To secure validity, two-stage clustering was conducted (Hair and Black, 2000). The number of adequate clusters derived from Ward's method was used to conduct the K-means method.

2.2.5. Step 5: interpretation of results

In Step 5, a concept map was presented on a two-dimensional graph. The group names given by the participants in Step 3 were reflected in naming the clusters. In addition, the impact of each statement and cluster on the perception of a chilly climate was understood from the rating results. Additional explanations are provided in the Results section.

3. Results

3.1. Multidimensional scaling analysis

In ALSCAL, the stress value is used to determine the appropriate dimensions of the graph. According to Kane and Trochim (2007), the stress value measures the degree to which the distances on the map are discrepant from the values in the input similarity matrix and 0.205–0.376 is an adequate range of the stress value for concept mapping. From the results of ALSCAL analysis, the associated stress value for each dimension was as follows: one dimension, 0.55 ($R^2 = 0.38$), two dimensions, 0.33 ($R^2 = 0.60$), three dimensions, 0.22 ($R^2 = 0.72$), four dimensions, 0.16 ($R^2 = 0.79$), and five dimensions, 0.13 ($R^2 = 0.83$). The lowest possible dimensions should be selected because interpretability and simplicity decrease as the number of dimensions increases (Borg and Groenen, 2005). In addition, considering the increase in R^2, the largest increase was observed in two dimensions. Thus, a two-dimensional model was chosen as the most suitable model for this study.

The researcher examined statements distributed on both the *X*- and *Y*-axes (Figure 2). First, in the positive direction of the *X*-axis, academic situations, such as school classes, team projects, and careers, were prominent. In the negative direction of the *x*-axis, phenomena in non-task contexts, such as interpersonal relationships and perception, were included. Therefore, the *X*-axis was named "context dimension" and considered to represent "task–non-task" contexts on both extremes. For clarity, the subtitles "task: academic" and "non-task: social" were added.

In the case of the *Y*-axis, the positive direction includes more open and direct gender discrimination content found in the words and actions of professors and other male students. Furthermore, the negative direction includes situations such as alienation and a lack of opportunities experienced by female students in the school environment or engineering culture. Therefore, the *Y*-axis was named "sexism dimension," and "explicit-implicit" sexism was named for both ends.

3.2. Hierarchical cluster analysis

As suggested by Yim and Ramdeen (2015), dendrograms and agglomeration schedules were used to determine the appropriate number of clusters. According to the dendrogram presented in SPSS, the number of clusters can be up to six, and according to the scree plot using the coefficient from the agglomeration schedule, the appropriate number of clusters is four. By setting the number of clusters to four, a K-means analysis was performed to form the final cluster.

Clusters were named after scrutinizing the statement contents in each cluster and the attributes of both the *X*- and *Y*-axes. Participants' feedback on the naming groups was also reflected. Cluster 1 was named "Exclusion and alienation inherent in the culture" (12 statements), Cluster 2 "Sexual objectification and lack of gender sensitivity" (15 statements), Cluster 3 "Male-centered academic situations" (18 statements), and Cluster 4 "Prejudice and generalization" (seven statements). Table 1 presents the statements for each cluster.

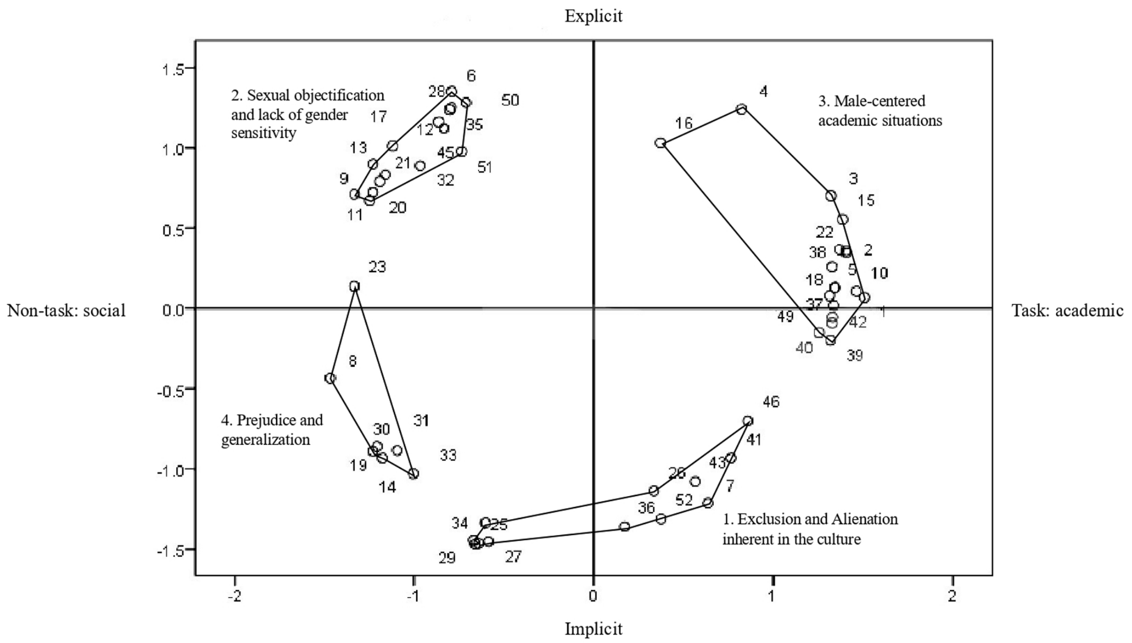

FIGURE 2
Concept map of the chilly climate perceived by female engineering students.

3.3. Interpretation of concept mapping

Figure 2 presents a concept map of the chilly climate perceived by female undergraduate engineering students. The small circles on the graph indicate the positions of the statements, reflecting the frequency with which participants categorized them into similar groups. Therefore, the closer the positions of the statements, the more similar the participants were considered.

Based on these two dimensions, the statements are distributed in four quadrants. First, the upper-right quadrant corresponds to explicit sexism in the task:academic contexts, which include direct remarks that favor men in academic or career-related situations. Most statements in Cluster 3 (Male-centered academic situations) are included here. Second, the upper-left quadrant shows explicit sexism in non-task:social contexts, such as derogatory remarks against women, sexual harassment, and molestation. The statements of Cluster 2 (Sexual objectification and lack of gender sensitivity) are densely distributed in this quadrant. Third, the lower-left quadrant is for implicit discrimination in non-task:social contexts, including prejudice and generalization of negative perspectives on women. Most statements from Cluster 4 (Prejudice and generalization) and some from Cluster 1 (Exclusion and alienation inherent in the culture) are located here. Fourth, the lower-right quadrant corresponds to implicit discrimination in the task:academic contexts such as marginalization or lack of opportunities for women. Some statements from Cluster 1 (Exclusion and alienation inherent in the culture) and Cluster 3 (Male-centered academic situations) are presented in this quadrant.

3.4. Influence ratings results

Table 1 (right-hand side) shows the mean value and standard deviation (SD) of the participants' rating results for each statement. Except for Statement 23, the average value of every statement was above the median of four on a 7-point Likert scale. Statements 50, 43, and 39 scored the highest average (6.09/7.00).

Cluster 2 (Sexual objectification and lack of gender sensitivity) had the highest mean value M = 5.37, followed by Cluster 3 (Male-centered academic situations) M = 5.26, Cluster 1 (Exclusion and alienation inherent in the culture) M = 5.08, and Cluster 4 (Prejudice and generalization) M = 4.92. Interestingly, the order of the mean value increase and that of the standard deviation decrease were the same. Cluster 2 had the smallest SD value SD = 0.41, followed by Cluster 3 SD = 0.47, Cluster 1 SD = 0.48, and Cluster 4 SD = 0.71.

4. Discussion

The purpose of this study was to investigate situations perceived as chilly by female engineering students. To this end, the study was conducted with 13 female students enrolled in engineering majors for over 4 semesters at seven four-year coeducational universities. After one-on-one interviews, 52 statements were extracted. Through the analysis, the study discovered four clusters on a conceptual map. The map was two-dimensional: an *X*-axis named "context dimension," with "task: academic" and "non-task: social" at both ends, and a *Y*-axis named "sexism dimension," having "explicit" and "implicit" at both ends. The four clusters, in the order of higher scores in the influence rating, are as follows: (i) "Exclusion and alienation inherent in the culture," (ii) "Sexual objectification and lack of gender sensitivity," (iii) "Male-centered academic situations," and (iv) "Prejudice and generalization."

First, female engineering students perceived a chilly climate the most by sexual objectification and lack of gender sensitivity. This cluster covers explicit sexism in non-task:social contexts such as blatant hostility toward feminism, sexual harassment or molestation,

TABLE 1 Statements by cluster on perceived chilly climate.

Cluster/statement	M	SD
Cluster 1: Exclusion and alienation inherent in the culture	5.08	0.48
43. School events in the department consist of male-dominated sports (soccer, basketball, e-sports, etc.).	6.09	1.14
44. To become close, female students should make efforts to meet male students' interests (game, drinking, sports, etc.).	5.73	1.27
7. Given that female students are a minority, inconveniences in a male-dominated culture or environment are not improved easily.	5.27	1.42
27. Even though female students make male friends, it is difficult to continue the relationship for many reasons (army[a], etc.).	5.27	1.79
36. There is a strong backlash from male students against programs only for female science and engineering students.	5.27	2.10
29. There is a lack of people with whom female students can openly talk about their experiences and feelings in school.	5.09	1.38
46. The more female students in the group, the more powerfully the female students can express their opinions.	5.00	1.18
25. There is a lack of opportunities for female colleagues to interact.	4.82	1.33
41. There is a lot of news about male seniors who landed nice jobs, but it is hard to hear the same about female seniors.	4.73	1.85
26. Male students' faults are tolerated more than those of female students.	4.64	1.96
34. Given that there are few female students, it is hard to choose friends.	4.64	1.75
52. Even though there are scholarships for female (engineering) students, the conditions for receipt are strict.	4.45	1.57
Cluster 2: Sexual objectification and lack of gender sensitivity	5.37	0.41
50. There is an antipathy among male students against feminism.	6.09	0.94
28. Male students rate or rank female students' appearance.	6.00	0.89
9. Men prefer female students not making their experience of sexual harassment or sexist remarks public.	5.73	1.90
13. Male students exclude and ridicule students who are considered feminists.	5.73	1.35
20. As men lack sensitivity or empathy for gender discrimination and sexual harassment, female students choose to tolerate themselves.	5.55	0.93
24. Female students had experienced or witnessed sexual harassment at a drinking party.	5.45	1.75
21. Even if they are at the scene of sexual harassment or molestation, male students do not interfere or help female students.	5.36	1.96
51. While emphasizing on them being a woman, male students help female students excessively.	5.36	1.36
6. There is little opportunity for open discussion on topics that can cause tension between men and women (e.g., feminism).	5.27	1.68
17. Male students consider and treat female students as potential girlfriends.	5.27	1.27
45. There is a drinking culture where women are split and seated at different tables.	5.18	1.47
35. Male students make comments about appearance to female students freely.	5.09	1.76
32. Male students speak freely about contents or language expressions that may be unpleasant or sensitive to female students.	5.00	1.90
11. Male students think that female students have it easier because they are "women."	4.73	1.27
12. Male students offer unsolicited help or advice to female students.	4.73	1.27
Cluster 3: Male-centered academic situations	5.26	0.47
39. There is a small number of or no female professors in the major.	6.09	1.51
3. When a male professor talks about career in class, he speaks mainly to a male audience.	5.82	0.60
5. In the team work of handling machines and tools, the key part is taken by men and the auxiliary part by women.	5.82	0.75
18. Even in the same position or grade level, the opinion of older male students carries more weight than that of female students.	5.82	1.08
1. The position of heads of department or projects are mostly held by male students.	5.64	1.21
2. A professor prefers male students to female students as his graduate students or undergraduate researchers.	5.64	2.06
40. Information and opportunities are centered on male students.	5.64	1.03
16. Although male students insist on doing the hard jobs, they judge female students if they do not participate in the work.	5.18	1.40

(Continued)

TABLE 1 (Continued)

Cluster/statement	M	SD
48. Female students are generally responsible for auxiliary tasks in a team project.	5.18	1.17
4. A professor makes sexist remarks about ability during class.	5.09	2.26
10. To be recognized on the same level as male students, female students work harder.	5.09	1.38
42. More information and opportunities are provided to male students in career-related school support or consultations with professors.	5.00	1.90
15. The professor uses subjects limited to male students (military, etc.) as an example in the class.	4.91	1.76
49. A group project is naturally led by male students from the beginning.	4.91	1.92
22. As there are only a few female students in the class, the ability or presence of female students is more noticeable to the professor.	4.82	1.54
47. In group project situations, male students are more responsive to men's opinions than those of women.	4.82	1.99
38. Regarding the lack of female professors in the major, the professor says, "It is because female students do not work hard enough."	4.64	2.25
37. Feedback on work or idea is provided better to male than female students.	4.55	2.02
Cluster 4: Prejudice and generalization	4.92	0.71
33. As there is a small number of female students, they are often under the spotlight and easily become the subject of backbiting.	5.82	0.75
31. When a female student is the subject of gossip, false rumors or misunderstandings are more easily generated and spread.	5.73	1.10
30. When a female student talks about the discomfort she feels, male students respond that she is too sensitive.	5.18	1.47
8. Men think women are emotional and cause disturbances (e.g., catfights, and factions).	4.91	1.92
14. Male students consider the words, actions and thoughts of individual female students as those of female (engineering) students as a whole.	4.64	1.36
19. The faults of female students are easily generalized to negative perceptions of women as a whole, even if they are of individuals.	4.27	1.79
23. In private situations such as drinking parties, boys treat girls with extreme caution or do not invite them at all.	3.91	1.14

[a]Korean men are obligated to go to the army.

comments on physical appearance, rude verbal expressions, and being treated as weak or potential girlfriends.

In particular, hostility toward feminism was among the most influential factors in the perception of a chilly climate. Overt antipathy about and ridicule of feminism by male peers and lack of opportunities to openly discuss this matter greatly impacted female students' perception of the engineering environment as hostile to women. Similarly, in Hall's (2016) study, men who supported women in STEM were opposed by other men for exhibiting a threat to masculinity. That is, men are considered the dominant gender, and advocating for women is not acceptable in engineering or STEM fields.

In addition, tolerating sexual harassment or sexist incidents rather than making them public had a significant effect on female students' perception of a chilly climate. This is because most men do not show much support to the victim and prefer not to stir up trouble. Therefore, "self-silencing" is common among women in STEM (Pololi and Jones, 2010), which is not a healthy way to cope with problems.

Moreover, male students evaluating or ranking female students on their looks also had a considerable impact on their perception of the chilly climate. As the remarks are about physical appearance rather than intellect or personality (Hall, 2016) and owing to the assumption that women are okay to be judged, they can be disrespectful and demeaning. In addition, situations such as treating women as potential girlfriends, positioning each woman at a different table at drinking parties, and showing excessive kindness with emphasis on their being women could be considered "subjective objectification" as women are treated as sex objects or possessions (Hall, 2016).

Second, female engineering students perceived a chilly climate in male-centered academic situations, which was found to have the second largest influence. This cluster includes both explicit and implicit sexism in task:academic contexts, and is closest to the definition of chilly climate mentioned by Hall and Sandler (1982). For instance, a professor making remarks about favoring male students in class is explicit sexism in this study and corresponds to the verbal display of chilly climate suggested by Hall and Sandler. In addition, professors' preference for male students as assistants or the tendency to provide them with more detailed feedback is implicit sexism and falls into the nonverbal aspect of the chilly climate in Hall and Sandler's study.

This cluster contained the highest number of statements with a high influence on the perception of a chilly climate. Among these, the presence of no or few female professors in the department had the greatest impact. According to previous studies, when providing academic or career-related advice, male professors have a tendency to hold low expectations from female students and encourage them to pursue management instead of technical careers (Kim and Lim, 2011; Cardador, 2017). Female mentors also positively affect undergraduate women in terms of aspiration, achievement (Young et al., 2013), self-concept, grant funding, and promotion (Richman et al., 2011). That is, female professors are not only career role models but also authoritative figures who can understand and help female students. Accordingly, the lack of female professors can further marginalize them.

In addition, in line with previous studies (Hall and Sandler, 1982; Ahn, 2017; Park, 2019), situations such as male students being the leaders most of the time, female students taking auxiliary tasks, information, and opportunities centered on male students were also confirmed in this study and largely affected the perception of a chilly climate.

Third, female engineering undergraduates perceived a chilly climate from "Exclusion and alienation inherent in the culture." This cluster is related to implicit sexism in various contexts. Examples in the non-task:social context include having difficulties in forming and maintaining relationships with male peers, lack of interaction among female friends, and small options for choosing friends. In the task:academic context, examples include statements about male-centered department events and culture, male students' opposition to programs only for women, and low availability of these support systems. Previous studies have reported women experiencing hardships in forming a relationship with male colleagues, rare instances of success stories of female graduates, and lack of school support for career development and employment (Han et al., 2010).

In particular, official department events focused only on male students' participation, which had the greatest impact on female students' perceptions of the chilly climate. This finding is distinctive and meaningful because it implies that department officials other than male professors or peers can cause a chilly climate. Female students experience a sense of alienation in a male-dominated engineering environment (Hall and Sandler, 1982; Chu, 2008; Park, 2019), and being considered incidental even for official events inevitably makes them feel more marginalized. As a sense of belonging has a direct relationship with feeling how fairly one is treated (Richman et al., 2011), female students strongly perceive a chilly climate. Furthermore, having to adapt to male students' interests to become close greatly influences female students' perception of a chilly climate, and limited interaction among female students appeared in several statements.

Fourth, female undergraduates in the engineering field perceived a chilly climate due to prejudice and generalization of women. This cluster includes implicit sexism in non-task:social contexts, such as prejudice about women being emotional and sensitive, rumors about female students, and overgeneralization derived from opinions or faults of a few female students. Prejudice against female students in STEM has been reported in many studies as follows: "Engineering is suitable for men, not women" (Chu, 2008; Blackburn, 2017; Cardador, 2017; Park, 2019), "It should be easier for female students because they are women" (Hall, 2016; Park, 2019), and "Women are not good at math and science" (Walton et al., 2015; Blackburn, 2017; Park, 2019). According to Hall's (2016) study, women conformed to the social patterns of men to avoid being labeled as "emotional" or a "bitch," which indicates the prevalence of prejudice and coping skills of women in STEM.

As minorities, being easy targets of gossip and readily created false rumors was found to have a significant effect on the perception of chilly climate. Usually, gossip occurs in informal settings where female students are absent, so there is no opportunity for them to explain or verify the truth. Minorities who are subject to negative stereotypes experience high levels of stress, stereotype threats, and daily adversities (Walton et al., 2015), and female students in STEM experience identity confusion and alienation due to stereotype threats (Johns et al., 2005; Rice et al., 2015; Thackeray, 2016). In other words, female students may experience stress for fear of being or becoming a target, and may be negatively affected due to the fear of confirming prejudice. This is unfair and discriminatory because male engineering students do not necessarily experience the same.

Thus far, the clusters have been mentioned in the order of greatest influence on the perception of chilly climate by female undergraduate students in the engineering field. "Sexual objectification and lack of gender sensitivity" and "Male-centered academic situations," corresponding to the first and second places, represent explicit sexism, whereas "Exclusion and alienation inherent in the culture" and "Prejudice and generalization," corresponding to the third and fourth places, are close to implicit sexism. Interestingly, the standard deviation increased in the same order, indicating larger disagreement among the participants. That is, "Sexual objectification and lack of gender sensitivity" had the greatest impact on the perception of chilly climate, and its disagreement degree was small. Conversely, "Prejudice and generalization" had the least influence, and participants' responses varied greatly. It can be interpreted that, as explicit sexism is delivered more directly, women's perceptions and influence are not significantly different. However, as implicit sexism is conveyed more subtly, it is not easily recognized (Hall and Sandler, 1982; Swim et al., 2001), resulting in a larger perception variance. Goldman (2012) also states that some undergraduate women in STEM fields are not willing to connect their experience with gender, and if these women do not credit gender for maltreatment or sexism in the field, they experience it without noticing. Though prejudice and generalization scored the last in this study, its contents should not be overlooked as it scored over 4, which is the median value of a 7-point Likert scale. Instead, more attention may be needed as subtle sexism can affect women's mental health more seriously (Sue et al., 2007; Dumont et al., 2010).

4.1. Implications

First, the 52 statements drawn from this study provide useful information and insight into the subjective experience of female undergraduate engineering students. In sum, these students perceived a chilly climate from exclusion and alienation inherent in the engineering culture, sexual objectification and lack of gender sensitivity of their male peers, male-centered academic situations, and prejudice and generalization of women. These findings may shed light on the discriminatory environment and treatment that female engineering students have to endure and draw public attention to make improvements. In addition, they may bridge the perception gap that men and women in the engineering field have and help create climate change from within.

Second, this study is significant because it was the first to explore and expand the understanding of chilly climate using the concept mapping method. This study verified that the perception of chilly climate can be displayed on a two-dimensional diagram, with "task: academic"-"non-task: social" context and "explicit"-"implicit" sexism on both axes. Points on the map indicating the location of each statement and cluster make it easy to understand the degree of similarity of each content. The chilly climate identified in this study encompasses a broader range of sexism than gender microaggression and expands Hall and Sandler's definition of chilly climate.

Third, scores of influence rating indicate which specific content affects women more in perceiving a chilly climate. The findings may be utilized not only by interested parties in education, but also in psychological counseling. In the counseling scene, the scores can help set the weight and priority of the content needed to identify and understand the client's problem. They can also be used to develop a scale to facilitate identification of these problems or to develop a counseling program that supports and empowers female engineering students. Furthermore, these findings are expected to be used as evidence for social justice advocacy activities to improve the gender-discriminatory environment faced by female engineering students.

4.2. Limitations and future directions

First, as the findings of this study were based on only 13 participants, the results should be interpreted with caution. To cover as diverse experiences as possible, participants were recruited from seven different universities across the country, but this may still be insufficient to reflect the perspectives of a broader population of undergraduate engineering students. In addition, as the participants were recruited based on their voluntary application, and the study relied on self-reports, there may be a bias in reflecting reality.

Second, as this study was conducted with only Korean participants, it may be difficult to generalize the results to all cultures. In a more collectivist culture, the influence of the social environment is stronger (Kim and Choi, 2014). As Korea traditionally has a culture closer to collectivism than individualism, the frustration felt by participants in chilly social relations can be greater. Therefore, the research results should be further tested with larger populations covering more diverse cultures, majors, and age groups.

Third, as this study was conducted during the COVID-19 pandemic, many parts of the study were replaced with non-face-to-face methods. For example, obtaining consent, interviews, similarity classification, and rating processes were all conducted via e-mail, video conference, and other computer programs. These new attempts, different from the traditional concept mapping method, may have affected the participants' understanding of instructions or concentration. Therefore, efforts to overcome these limitations through repeated research and face-to-face research are required.

5. Conclusion

The present study examined South Korean female engineering undergraduate students' perceptions of chilly climate using a concept mapping method and provided a visual representation of the results. Fifty-two final statements from 13 participants were extracted, and their perceptions were grouped into four clusters in a two-dimensional model. The present study conceptualized the subjective experiences of female students in male-dominated campus environments and provided influence rating results so that opportune measures could be prioritized by interested parties in education and mental health care. The unfavorable climate for women can be ameliorated with the help of rigorous future research and social justice advocacy actions.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board of Seoul National University (IRB No. 2012/001-023). The patients/participants provided their written informed consent to participate in this study.

Author contributions

TK and DK: conceptualization and methodology. TK: data collection, formal analysis, and manuscript writing. DK: supervision. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2020S1A3A2A02103411).

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 potential conflicts of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahn, M.-S. (2017). Gender discriminatory environment and masculinity in university campus. *Women Stud.* 28, 113–134. doi: 10.1371/journal.pone.0190807

Allan, E. J., and Madden, M. (2006). Chilly classrooms for female undergraduate students: a question of method? *J. Higher Educ.* 77, 684–711. doi: 10.1353/jhe.2006.0028

Bedi, R. P. (2006). Concept mapping the client's perspective on counseling alliance formation. *J. Couns. Psychol.* 53, 26–35. doi: 10.1037/0022-0167.53.1.26

Blackburn, H. (2017). The status of women in STEM in higher education: a review of the literature 2007–2017. *Sci. Technol. Libr.* 36, 235–273. doi: 10.1080/01942 62X.2017.1371658

Borg, I., and Groenen, P.J. (2005). *Modern multidimensional scaling: theory and applications*. New York: Springer.

Cabay, M., Bernstein, B., Rivers, M., and Fabert, N. (2018). Chilly climates, balancing acts, and shifting pathways: what happens to women in STEM doctoral programs. *Soc. Sci.* 7:23. doi: 10.3390/socsci7020023

Cammaert, L. P. (1985). How widespread is sexual harassment on campus? *Int. J. Womens Stud.* 8, 388–397.

Cardador, M. T. (2017). Promoted up but also out? The unintended consequences of increasing women's representation in managerial roles in engineering. *Organ. Sci.* 28, 597–617. doi: 10.1287/orsc.2017.1132

Cardoso, P., and Moreira, J. M. (2009). Self-efficacy beliefs and the relation between career planning and perception of barriers. *Int. J. Educ. Vocat. Guid.* 9, 177–188. doi: 10.1007/s10775-009-9163-2

Chu, H. (2008). Women's deflection feeling in engineering schools. *Discourse* 11, 117–150. doi: 10.17789/discou.2008.11.3.005

Do, S. L. (2008). Differences of job related psychological variables between male and female engineering students. *Korean J. Educ. Psychol.* 22, 519–535.

Duberley, J., and Cohen, L. (2010). Gendering career capital: an investigation of scientific careers. *J. Vocat. Behav.* 76, 187–197. doi: 10.1016/j.jvb.2009.09.005

Dumont, M., Sarlet, M., and Dardenne, B. (2010). Be too kind to a woman, she'll feel incompetent: benevolent sexism shifts self-construal and autobiographical memories toward incompetence. *Sex Roles* 62, 545–553. doi: 10.1007/s11199-008-9582-4

Goldman, E. G. (2012). Lipstick and labcoats: undergraduate women's gender negotiation in STEM fields. *NASPA J. Women Higher Educ.* 5:115. doi: 10.1515/njawhe-2012-1098

Hair, J. F., and Black, W. C. (2000). "Cluster analysis" in *Reading and understanding more multivariate statistics*. eds. L. G. Grimm and P. R. Yarnold (Washington, DC: American Psychological Association), 147–205.

Hall, K. (2016). They believe that because they are women, it should be easier for them. *Subtle and overt sexism toward women in STEM from social media commentary. [dissertation]*. Richmond (VA): Virginia Commonwealth University

Hall, R. M., and Sandler, B. R. (1982). "The classroom climate: a chilly one for women?" in *Association of American Colleges. Project on the status and education of women*. ed. W.D.C.P.o.t.S.a.E.o.W (DC, Washington: Association of American Colleges)

Hall, R.M., and Sandler, B.R. (1984). *Out of the classroom: a chilly campus climate for women?* Washington, DC: Project on the Status and Education of Women, Association of American Colleges.

Han, J., Han, S., and Kim, Y. (2010). The need analysis for development of the customized educational programs for engineering female students. *J. Eng. Educ. Res.* 13, 130–142. doi: 10.18108/jeer.2010.13.4.130

Hatmaker, D. M. (2013). Engineering identity: gender and professional identity negotiation among women engineers. *Gend. Work. Organ.* 20, 382–396. doi: 10.1111/j.1468-0432.2012.00589.x

Hwang, S. (2020). A systematic review of female engineering students related studies. *J. Eng. Educ. Research.* 23, 31–42. doi: 10.18108/jeer.2020.23.2.31

Janz, T. A., and Pyke, S. W. (2000). A scale to assess student perceptions of academic climates. *Can. J. Higher Educ.* 30, 89–122. doi: 10.47678/cjhe.v30i1.183347

Jensen, L. E., and Deemer, E. D. (2019). Identity, campus climate, and burnout among undergraduate women in STEM fields. *Career Dev. Q.* 67, 96–109. doi: 10.1002/cdq.12174

Jeong, Y., Oh, M., and Kim, J. H. (2008). Examination of psychological correlates of woman engineering students. *J. Eng. Educ. Res.* 11, 34–45. doi: 10.18108/jeer.2008.11.4.34

Johns, M., Schmader, T., and Martens, A. (2005). Knowing is half the battle: teaching stereotype threat as a means of improving women's math performance. *Psychol. Sci.* 16, 175–179. doi: 10.1111/j.0956-7976.2005.00799.x

Kane, M., and Trochim, W.M. (2007). *Concept mapping for planning and evaluation*. California: Sage Publications

Kim, M. S., and Choi, B. G. (2014). The relation between contextual supports and sense of belonging among south Korean first year engineering students: mediating effects of academic self-efficacy and outcome expectations. *Korean J. Educ. Res.* 52, 111–138.

Kim, D. I., and Lim, C. H. (2011). Satisfaction and needs of female students in engineering with counseling men professors. *J. Eng. Educ. Research.* 14, 3–15. doi: 10.18108/jeer.2011.14.6.3

Kim, J. Y., and Meister, A. (2022). Microaggressions, interrupted: the experience and effects of gender microaggressions for women in STEM. *J. Bus. Ethics* 2022, 1–19. doi: 10.1007/s10551-022-05203-0

Kim, M. S., Yang, J. W., and Yon, K. J. (2016). Qualitative research on career choices and graduate program experiences of female graduate students in science and engineering. *Korean J. Couns. Psychother.* 28, 191–216. doi: 10.23844/kjcp.2016.02.28.1.191

Korea Foundation for Women in Science, Engineering and Technology. (2021). Analysis report on the statistics of development and utilization of women and men in STEM. Available at: https://wiset.or.kr/prog/pblcte/kor/sub02_03_01_03/rptpAll/view.do

Lee, M. J., Collins, J. D., Harwood, S. A., Mendenhall, R., and Huntt, M. B. (2020). 'If you aren't white, Asian or Indian, you aren't an engineer': racial microaggressions in STEM education. *Int. J. STEM Educ.* 7, 1–16. doi: 10.1186/s40594-020-00241-4

Makarem, Y., and Wang, J. (2020). Career experiences of women in science, technology, engineering, and mathematics fields: a systematic literature review. *Hum. Resour. Dev. Q.* 31, 91–111. doi: 10.1002/hrdq.21380

Marshall, A., Pack, A. D., Owusu, S. A., Hultman, R., Drake, D., Rutaganira, F. U. N., et al. (2021). Responding and navigating racialized microaggressions in STEM. *Pathog. Dis.* 79:ftab027. doi: 10.1093/femspd/ftab027

Miles, M. L., Brockman, A. J., and Naphan-Kingery, D. E. (2020). Invalidated identities: the disconfirming effects of racial microaggressions on black doctoral students in STEM. *J. Res. Sci. Teach.* 57, 1608–1631. doi: 10.1002/tea.21646

Min, M. S., and Lee, J. H. (2005). Analysis of educational and occupational experiences of women studying engineering. *Korean J. Sociol. Educ.* 15, 65–93.

Ministry of Education. (2019). Education statistics analysis data book: higher education statistics. Korean Education Development Institute. Available at: https://kess.kedi.re.kr/publ/view?survSeq=2019&publSeq=44&menuSeq=0&itemCode=02&language=.

Nadal, K. L., and Haynes, K. (2012). "The effects of sexism, gender microaggressions, and other forms of discrimination on Women's mental health and development" in *Women and Mental Disorders*. eds. P. K. Lundberg-Love, K. L. Nadal and M. A. Paludi (Westport, CT: Praeger/Santa Barbara, CA: ABC-CLIO), 87–101.

National Science Foundation, and National Center for Science and Engineering Statistics. (2021). Women, minorities, and persons with disabilities in science and engineering: 2021 [special report NSF 21–321]. Available at: https://ncses.nsf.gov/pubs/nsf21321/data-tables

O'Brien, L. T., Garcia, D. M., Adams, G., Villalobos, J. G., Hammer, E., and Gilbert, P. (2015). The threat of sexism in a STEM educational setting: the moderating impacts of ethnicity and legitimacy beliefs on test performance. *Soc. Psychol. Educ.* 18, 667–684. doi: 10.1007/s11218-015-9310-1

Park, M. J. (2019). Applying grounded theory to study the engineering identity construction of female engineering students. *J. Soc. Sci.* 26, 145–168. doi: 10.46415/jss.2019.03.26.1.145

Pascarella, E. T., Hagedorn, L. S., Whitt, E. J., Yeager, P. M., Edison, M. I., Terenzini, P. T., et al. (1997). Women's perceptions of a "chilly climate" and their cognitive outcomes during the first year of college. *J. Coll. Stud. Dev.* 38, 109–124.

Paulson, B. L., Truscott, D., and Stuart, J. (1999). Clients' perceptions of helpful experiences in counseling. *J. Couns. Psychol.* 46, 317–324. doi: 10.1037/0022-0167.46.3.317

Pololi, L. H., and Jones, S. J. (2010). Women faculty: an analysis of their experiences in academic medicine and their coping strategies. *Gend. Med.* 7, 438–450. doi: 10.1016/j.genm.2010.09.006

Renn, K.A. (2014). *Women's colleges and universities in a global context* Baltimore: Johns Hopkins University Press

Rice, K. G., Ray, M. E., Davis, D. E., Deblaere, C., and Ashby, J. S. (2015). Perfectionism and longitudinal patterns of stress for STEM majors: implications for academic performance. *J. Couns. Psychol.* 62, 718–731. doi: 10.1037/cou0000097

Richman, L. S., vanDellen, M., and Wood, W. (2011). How women cope: being a numerical minority in a male-dominated profession. *J. Soc. Issues* 67, 492–509. doi: 10.1111/j.1540-4560.2011.01711.x

Roper, R. L. (2019). Does gender bias still affect women in science? *Microbiol. Mol. Biol. Rev.* 83:e00018. doi: 10.1128/MMBR.00018-19

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

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

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M. B., Nadal, K. L., et al. (2007). Racial microaggressions in everyday life: implications for clinical practice. *Am. Psychol.* 62, 271–286. doi: 10.1037/0003-066X.62.4.271

Swim, J. K., Hyers, L. L., Cohen, L. L., and Ferguson, M. J. (2001). Everyday sexism: evidence for its incidence, nature, and psychological impact from three daily diary studies. *J. Soc. Issues* 57, 31–53. doi: 10.1111/0022-4537.00200

Thackeray, S. (2016). *Overcoming the toxic influence of subtle messaging: Utah women who persist in STEM [thesis]* (Boston (MA): Northeastern University)

Trochim, W. M. K. (1989). An introduction to concept mapping for planning and evaluation. *Eval. Program Plann.* 12, 1–16. doi: 10.1016/0149-7189(89)90016-5

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

Yang, Y., and Carroll, D. W. (2018). Gendered microaggressions in science, technology, engineering, and mathematics. *Leadersh. Res. Educ.* 4, 28–45.

Yim, O., and Ramdeen, K. T. (2015). Hierarchical cluster analysis: comparison of three linkage measures and application to psychological data. *Tutor. Quant. Methods Psychol.* 11, 8–21. doi: 10.20982/tqmp.11.1.p008

Young, D. M., Rudman, L. A., Buettner, H. M., and McLean, M. C. (2013). The influence of female role models on women's implicit science cognitions. *Psychol. Women Q.* 37, 283–292. doi: 10.1177/0361684313482109

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 17 August 2023
DOI 10.3389/feduc.2023.1194968

OPEN ACCESS

EDITED BY
Lisbet Rønningsbakk,
UiT The Arctic University of Norway, Norway

REVIEWED BY
Ian Thacker,
University of Texas at San Antonio, United States
Karen Blackmore,
University of Worcester, United Kingdom

*CORRESPONDENCE
Daniela Fernandez
dpf204@exeter.ac.uk

RECEIVED 27 March 2023
ACCEPTED 27 July 2023
PUBLISHED 17 August 2023

CITATION
Fernandez D, White S, Smith HCM, Connor PM and Ryan M (2023) Gender inequality in science, technology, engineering and mathematics: gendered time disparities in perceived and actual time spent in practical laboratory-based activities.
Front. Educ. 8:1194968.
doi: 10.3389/feduc.2023.1194968

Gender inequality in science, technology, engineering and mathematics: gendered time disparities in perceived and actual time spent in practical laboratory-based activities

Daniela Fernandez[1]*, Sarah White[1], Helen C. M. Smith[2], Peter M. Connor[2] and Michelle Ryan[3,4]

[1]Psychology Department, University of Exeter, Exeter, United Kingdom, [2]Engineering Department, University of Exeter, Exeter, United Kingdom, [3]Global Institute for Women's Leadership, Australian National University, Canberra, ACT, Australia, [4]Faculty of Economics and Business, Groningen University, Groningen, Netherlands

Lab-based activities provide essential skills for students within STEM disciplines, as lab activities provide students with research skills and science knowledge. Therefore, it is critical to note that female students have reported feeling less confident in conducting lab-based activities and report a lower sense of belonging in the lab. In two studies (N=544) we examined gender differences in the time that students spent, and perceived they spent, on various laboratory-based activities. We predicted that female (vs. male) students in science, technology, engineering and mathematics (STEM) would both perceive, and actually spend, less time in practical, science-specific activities, such as using equipment, compared to observing or note-taking. Study 1a (N=227) was an online, cross-sectional survey where university STEM students reported their perceptions of time spent during lab-based practical activities, and how satisfied they were with their time spent in these activities. Study 1b (N=318) was an observational study of university practical lab sessions in STEM disciplines. Our findings demonstrated that female (vs. male) students (1) spent more time recording and taking notes during lab sessions, (2) did not perceive, yet actually spent, less time in the lab using equipment, and (3) were equally satisfied with their time in the lab using equipment. Together, these results suggest that women occupy stereotypically gendered roles in the STEM lab, spending less time on activities that are key for their professional development. Furthermore, the fact that students from disciplines with more female participation were more satisfied with their time spent in lab activities can promote the insidious effects of assessing gender participation and equality in STEM through numbers only. The differences in time spent in lab activities-yet the lack of acknowledge of these differences-opens the discussion about how women might be receiving reduced utility from their programmes, and that universities may not be delivering on their obligations to ensure equal access to teaching resource and opportunities.

KEYWORDS

gender, STEM, higher education, laboratories, practicals

Introduction

Gender inequality science, technology, engineering and mathematics (STEM) disciplines is a well-documented problem and, despite educational institutions' efforts to improve women's participation in STEM disciplines, in terms of numbers (e.g., uptake in enrolment of STEM classes) and experiences (e.g., sense of belonging), gender inequalities remain. Women's participation in science, technology, engineering and mathematics disciplines (STEM) remains a challenge in higher education (Prieto-Rodriguez et al., 2020). Although women outnumber men in some of these areas (e.g., biosciences; Vincent-Lancrin, 2008; Cheryan et al., 2011a,b), they remain underrepresented in many STEM disciplines, such as computer sciences and engineering (Liben and Coyle, 2014). Indeed, in the United Kingdom, while women's enrolment in STEM disciplines is increasing, under-representation persists: (a) compared to male students, female students are still less likely to take qualifying STEM subjects in high school (Department for Education and Behavioural Insights Team, 2020); (b) only one in three STEM university majors and one in four STEM professionals are women (STEM Women, 2022); and (c) only 9% of STEM professors are women (Kirkup et al., 2010).

Research suggests that this continued under-representation of women in STEM is multiply determined. The lack of role models in STEM can lead female students to perceive that they do not "fit" into the field (Cheryan et al., 2011a,b). Moreover, if those examples of female success in STEM are portrayed as being distant from students (Leslie et al., 2015), students will not perceive the role models as attainable, thus affecting their motivation to persist in STEM (Gladstone and Cimpian, 2021). Research also demonstrates that even for those women in STEM, perceiving a lack of fit with the prototype can facilitate feelings of marginalisation from the broader STEM group (Kim et al., 2018).

In addition to these issues of role models and fitting in, women's under-representation in STEM can be, at least in part, explained by *gendered expectations* in STEM disciplines (Heilman, 2012). Stereotypes about women's abilities (or lack of them) in STEM subjects, have been shown to create stigma and affects women's STEM motivation (Pronin et al., 2004; Casad et al., 2017). Similarly, beliefs about the perceived competitive nature of STEM fields influences women's career choices (Buser et al., 2012), and as such, ideas that men are more talented and interested in sciences than women (Boston and Cimpian, 2018); while women are seen as more talented and interested in humanities/social disciplines (Trusz, 2020). Indeed, the stereotype of masculinity as "effortlessly successful" (Jones and Myhill, 2004; Jackson and Dempster, 2009; Jackson and Nyström, 2015) can lead to beliefs about women's lack of ability, and lack of fit, in STEM. In contrast, stereotypes of women as warmer, kinder and focused on communal goals are the opposite of representations of STEM disciplines as inherently competitive, independent and analytical (Carli et al., 2016; Boucher et al., 2017).

These gendered expectations are also part of STEM students' everyday experiences in their education. Indeed, as STEM associated degrees and modules are seen as an individual choice of students (Burkam et al., 1997), research has focused on understanding how classrooms and lecture hall dynamics might explain gender differences in academic experiences. For example, the perceived similarity and sense of belonging of female students in computer science increased when stereotypically masculine items were removed from the classrooms (Cheryan et al., 2011a,b). Moreover, women's experiences in STEM classrooms are likely to be shaped by explicit and subtle cues regarding their lack of ability to succeed in their degree (Pronin et al., 2004), bias in the evaluation of their performance (Andrus et al., 2018) and, overall, a "chilly" climate (Walton et al., 2015; Wilkins-Yel et al., 2022). These environmental characteristics can affect female students' participation in the classroom. Indeed, research using students' self-reported answers about their interactions and participation in the classroom showed that male students, compared to female students, were more likely to perceive higher levels of participation, in terms of their (a) participation engaging in discussions, (b) answering the instructor's questions and (c) taking a leader role in small group work (Eddy et al., 2015). For a review on gender disparities in classrooms see Eddy and Brownell (2016).

However, a less explored context where gender inequalities emerge in STEM education is the laboratory (lab). Lab-based activities are indeed a central aspect of most of STEM disciplines curricula (Velasco et al., 2016; Arnado et al., 2022). Lab-based activities provide essential skills for students within STEM disciplines, as lab activities provide students with research skills (Lopatto, 2004), and focus on inquiry-based and active learning (Wan et al., 2020), managing scientific equipment, and developing skills associated with team work (Batty and Reilly, 2022). Therefore, it is critical to note that female students have reported feeling less confident in conducting lab-based activities and report a lower sense of belonging in the lab (Batty and Reilly, 2022).

The disparities in how female and male students approach lab-based activities have been described not only through students' self-reports, but also through observation of lab groups. Ethnographic and qualitative research on pairs of physics students has demonstrated that there are disparities in task division in lab groups within physics classes between women and men. For instance, interviews and ethnographic observations showed that groups within lab sessions are likely to adopt a model of work that disadvantages women, whom are relegated to the Secretary and Hermione archetype, with men taking the task-related roles to use equipment (Doucette et al., 2020). Hence, gendered roles within mixed gender groups were also found to be more likely to occur, with women either undertaking the "Hermione" role, that is, taking on a disproportionate amount of the work compared to their male lab partners, or assuming the secretary role which involved female partners mainly taking notes and recording data whilst male partners interacted directly with the lab equipment (Doucette et al., 2020). Further research has focused on understanding different behaviours within the lab sessions, and how they might be different according to gender. For example, research has shown how in lab sessions where experimental work is emphasised, women are less likely to use equipment, compared to men (Quinn et al., 2018).

The use of observational data in STEM educational practices research has provided important insights about the persistence of gender inequalities and stereotyping roles in STEM lab-based activities (e.g., Lucht, 2015). However, the limited research we are able to find suggests observing labs in STEM is critical for two key reasons. First, a lesser interaction with lab equipment could contribute to women perceiving themselves to be incongruent with the role of a scientist (Doucette et al., 2020). As one of the first exposures to the practical aspects of their discipline, such activities shape students' identities as scientists (Gonsalves et al., 2016) and as participants in a wider scientific community (Quinn et al., 2018). Indeed, as individuals see themselves being part of the STEM community, they identify and feel that they belong to this group (Kim et al., 2018), which contributes to individuals' motivation and wellbeing. Second, lab practical activities are important for students to see STEM knowledge as less abstract and more connected with their everyday lives, leading to higher engagement, motivation, and interest in these disciplines (Holmes et al., 2022). Similarly, previous research with secondary students demonstrated that higher hands-on making attitude-this is, the preference towards objects-was associated with higher curiosity and, in turn, with higher STEM career interest (Cui et al., 2022). Hence, the fact that women's participation is restricted to certain roles in lab practices might constrain their experiences and, therefore, motivation and engagement towards STEM disciplines.

Therefore, these initial findings are particularly concerning as laboratory-based practical activities are a key aspect of STEM studies. Despite the fact that there is an increased interest on addressing gender inequalities in STEM (Holmes et al., 2022), research looking at lab practices is-from our knowledge-limited. Moreover, the existing evidence has shown that students from different genders approach lab activities differently. However, this previous research has reported these activities focusing on (a) counting behaviours, (b) only observations of lab activities, without including other data, such as students' perceptions about their roles in lab activities, and (c) qualitative techniques that allow us to understand differences in how lab activities are performed by students, but with less clarity about how much time students spend in different lab activities.

Considering the importance of lab activities in STEM disciplines, we argue that it is key to look at how students approach lab activities and how these approaches might be different according to their gender. Furthermore, we also argue that it is important to focus not only on observations, but also in how students perceive they navigate lab practical sessions, in terms of the activities they conduct and the time spent on those activities. Indeed, an important aspect to understand students' learning experiences is how they monitor their actual task performance, as well as their assessment of the actual performance. This process, also known as calibration (Alexander, 2013), provides evidence to understand the importance of forming sound judgement about one's abilities, which might impact on individuals' meta-cognitive skills and strategic behaviour (Alexander, 2013). Hence, research about lab practicals and gender distributed activities can be of benefit when it takes into account students' perceptions and actual time spent in activities, as both elements inform us about students' judgments and potential future strategic behaviour when they participate in lab activities in STEM contexts.

In this study, we aim to contribute to this previous research, by now integrating two dimensions of analysis regarding lab practical activities: students perceived and actual time spent on activities. Moreover, we aim to analyse potential gender differences in both categories.

The current research

The present research builds on this work by (a) comparing lab equipment time use by female and male students across multiple STEM disciplines in a United Kingdom university, with larger mixed gender groups, through video recordings of lab activities and (b) asking students directly about their perceived time spent on specific tasks in lab groups, and their satisfaction with this time. Thus, enabling a direct comparison between university students' perceptions of their involvement in lab group activities, to their actual amount of time spent on these activities through time stamped data, within a particular university.

Following this, our research aims to explore how students perceive their peers in lab-based practical activities, as well as how students experience their participation in terms of time spent doing specific activities. Within these experiences, lab-based activities provide a first approach to practical STEM work and are fundamental to the process of becoming a scientist. However, as it was discussed, female students are likely to keep facing inequalities in terms of how these tasks are distributed, with women undertaking more "administrative" and passive work in the lab (e.g., taking notes or observing), and men undertaking more active and stereotypically "scientific" work (e.g., using equipment). The present research aims to provide further evidence on this issue, now taking into account gender differences in the time that students spend in different lab-based activities, considering students self-reports and time measured at the lab. Indeed, despite the benefits of observing gender bias in educational settings, this methodology has faced criticism due to the potential researcher bias (Blickenstaff, 2005). Hence, our study measured the time for the activities observed.

We report the results of one study including two phases with undergraduate students from STEM disciplines in a United Kingdom university. We included two dimensions described in the research on gender equality in STEM education: (a) self-reports on the perception of time spent and (b) the actual time spent on different tasks in STEM disciplines (through the analysis of recorded lab sessions). We asked a group of students from different STEM disciplines about the time they spent in different lab-based activities, such as using equipment, recording and taking notes, and observing. We also asked them about their satisfaction with their time spent on these activities (Study 1a). Afterwards, we conducted a separate study where we recorded different lab sessions where students from different STEM disciplines participated, and measured the time that students spent in the lab-based activities described (Study 1b). We predicted that (a) female students, compared to male students, would perceive spending less time using equipment, more time recording and observing data (H1); (b) female students, compared to male students, would be less satisfied with the time that they spent on these activities (H2); and (c) within practical lab-based settings, female students, compared to male students, would spend less time using equipment, and more time recording and observing data (H3).

Study 1a

In the first phase of our study, we examined female and male students' perceptions of their lab-based practical experiences, with an emphasis on the perceived time they spent on specific lab activities

TABLE 1 Descriptive statistics and bivariate correlations (Study 1a).

	Women	Men	Bivariate correlations				
	M (SD)	*M (SD)*	2	3	4	5	6
1. Relative time spent using equipment	3.05 (0.44)	3.17 (0.48)	0.18***	−0.11	−0.03	−0.35***	−0.09
2. Satisfaction with time spent using equipment	3.40 (0.85)	3.20 (0.95)	—	−0.05	0.28***	−0.11	0.36***
3. Relative time spent recording/taking notes	3.16 (0.54)	3.02 (0.56)	—	—	0.27***	−0.04	0.09
4. Satisfaction time spent recoding/taking notes	3.16 (0.84)	3.02 (0.83)	—	—	—	−0.03	0.40***
5. Relative time spent observing	2.95 (0.51)	2.82 (0.56)	—	—	—	—	0.12
6. Satisfaction with time spent observing	3.13 (0.88)	2.90 (0.91)	—	—	—	—	—

***$p < 0.001$.

and their satisfaction with that time. Undergraduate students were recruited by staff in STEM-facing colleges.

Participants

We recruited 370 STEM undergraduate students from three STEM-facing Colleges within a United Kingdom university who volunteered to participate in the study. We also recruited participants from the College of Medicine and Health, but we decided to focus on disciplines traditionally associated with STEM. Analysis results including the College of Medicine and Health can be found in Supplementary materials. We excluded participants who (a) had not participated in lab-based practical activities, (b) had not fully completed the survey, and (c) did not identify as women or men, leaving useable data for 227 participants (56.8% women, 43.2% men). Students were in their first (54.2%) or second (47.6%) years of study (third students were excluded from participating to avoid a conflict with a nation-wide survey of third year students). *A priori* G*Power analysis (v.3.1, Faul et al., 2007) revealed that a sample of 200 was needed to reach an 80% power to detect a small effect size ($f = 0.20$). Students participated in Spring 2020.

Procedure

This study was part of wider research about students' perceptions about their lab-practical based experiences. During the recruitment process, we contacted participants via email inviting them to take part in the survey. Participants first received an email providing brief details of the survey, following a link to access the project information sheet explaining the study and the consent form. Consenting participants were then directed to the survey. Participants answered questions on both group and individual lab-based work. Following the survey, participants were debriefed in full and provided with the opportunity to be entered into a prize draw for a £75 gift voucher.

Measures

After completing demographic questions (gender, year of study, college, and discipline of study), students were asked whether they had undertaken any lab-based practical work during their course at the university, and whether this practical work was mainly carried out in groups or individually. Students that (a) had not undertaken any lab-based practical work during their course and (b) mainly carried out this practical work individually, were excluded from the study.

To measure students' perceived relative time (a) using equipment, (b) recording or taking notes, and (c) observing, we used a single-item Likert scale type from 1 ("much less than other people") to 5 ("much more than other people"), with one item for activity with the following question: "relative to other people you work with, how much time do you spend (a) using equipment, (b) recoding/taking notes and (c) observing."

We also asked to students to assess their satisfaction with the amount of time that they were currently spending (a) using equipment, (b) recording or taking notes, and (c) observing ("how satisfied are you with the amount of time spent in the data recording/note-taking role?"). We used a single-item Likert scale type from 1 ("not at all satisfied") to 5 ("extremely satisfied"). Finally, as a covariate, we included female participation by discipline, following Athena Swan datasets and administrative information. We developed a ranking for 11 disciplines (1 meaning low numbers of women participating and 11 meaning higher numbers of women participating see Supplementary material). For descriptive and correlations see Table 1.

Results

Gender differences in lab experiences

To test H1, we conducted an ANOVA test (gender: female vs. male) female participation by discipline as the covariate. Female participation by discipline was indeed positively associated with students' satisfaction in terms of their time spent using equipment, and recording and taking notes. However, students from disciplines with lower female participation, perceived to spend more time observing and were less satisfied with their time observing[1] (Table 2). Results showed that students in disciplines with higher female participation were more satisfied with their time spent on these activities. There were gender differences in how students perceived their time spent on different practical lab-based work. However, these

1 We conducted exploratory analysis to know whether students' gender had an interaction with the levels of female participation on their perceived time and time satisfaction. However, results were not significant.

TABLE 2 Effect of participation by discipline on time perceived and satisfaction with time in lab based practical activities.

	Type III sum of squares	$F(1, 224)$	Sig.	$\eta^2 p$
Relative time spent using equipment	0.45	2.18	0.14	0.01
Satisfaction with time spent using equipment	6.62	8.51	0.00	0.04
Relative time spent recording/taking notes	3.46	12.15	<0.00	0.05
Satisfaction with time spent recoding/taking notes	1.48	2.14	0.15	0.01
Relative time spent observing	1.68	6.10	0.01	0.03
Satisfaction with time spent observing	4.82	6.17	0.01	0.01

differences were focused on perceptions of time recording and taking notes, and not in terms of time using equipment and observing. Thus, female students-compared to male students-perceived to spend more time recording and taking notes, $F(1, 224) = 8.86$, $p = 0.00$, $\eta^2 p = 0.04$. There were no reported gender differences for perceived time spent using equipment, $F(1, 224) = 1.96$, $p = 0.16$, $\eta^2 p = 0.01$; and observing, $F(1, 224) = 0.954$, $p = 0.33$, $\eta^2 p = 0.01$.

Furthermore, both women and men were equally satisfied with their time spent on lab activities, in terms of satisfaction with time spent using equipment, $F(1, 224) = 0.39$, $p = 0.53$, $\eta^2 p = 0.00$, satisfaction with time spent taking and recording notes, $F(1, 224) = 2.57$, $p = 0.11$, $\eta^2 p = 0.01$, or satisfaction with time spent observing, $F(1, 224) = 0.02$, $p = 0.27$, $\eta^2 p = 0.01$. In other words, women reported spending more time on recording notes than men did (Supplementary Figure S1), but women and men were equally satisfied with the amount of time they spent on this task (Supplementary Figure S2).

Discussion

In Study 1a, we found partial support for our hypotheses, regarding differences in how male and female students perceived the time spent on specific activities in practical lab-based settings. As we predicted, female students perceived they spent more time than other students recording and taking notes during lab-based activities. However, no differences were reported regarding students' perceived time using equipment and observing, nor their satisfaction with the time spent on these activities. That is to say, despite perceiving different amounts of time recording, female and male students reported similar levels of satisfaction with the time they reported spending conducting these activities. This may be related to the positive effect of levels of female participation within the discipline on students' satisfaction, as female participation was positively associated with levels of satisfaction. As female students tended to participate in predominantly female disciplines, this might have boosted their satisfaction with their time spent.

In Study 1a we examined gender differences in how students perceived their time in different practical lab-based activities. Although our exploratory results partially supported H1, showing differences in how students perceived their time in different practical lab-based activities, particularly in the ones considered as less "scientific" (e.g., taking notes), results are based on students' perceptions. Following this, our next study explored whether there were gender differences in the *actual* time that students spent in different practical lab-based activities. Hence, in Study 1b, we examined gender differences in actual practical lab participation, via recorded practical lab sessions.

Study 1b

In Study 1b, we aimed to explore H3 in a realistic setting: practical lab-based activities. Hence, instead of focusing on students' perceptions about their time in different lab-based activities-as it was analysed in Study 1a-we decided to look at the actual time that students spent on these activities, and analysed whether gender differences could be identified. Although previous research has used observation in educational settings to explore learning practices in STEM (e.g., Velasco et al., 2016; Stains et al., 2018; Wan et al., 2020), from our knowledge, this method has not been used to analyse gender differences in participation in lab-based settings. Therefore, in this study, we explored a lesser known aspect of gender inequality in STEM, that is, practical lab-based settings. For this aim, we recorded different sessions of practical lab-based activities and measured the time (in seconds) that students spent using equipment, recording/taking notes and observing. We then tested potential gender differences in the time measures. We hypothesised that, during lab-based practical activities, female students would spend more time recording, taking notes, and observing, and less time using equipment than male students (H3).

Participants

We recruited 335 university students enrolled in STEM disciplines, at the same University as Study 1a (including Engineering and Chemistry). Students participated in one of their programme lab classes. Of those, we excluded 8 participants as their gender was not clear in the recording and could not be coded, leaving a total of 327 participants (20.5% women, 79.5% men). A sensitivity G*Power analysis (v.3.1, Faul et al., 2007) revealed that this sample gave us an 80% power to detect a medium effect size ($d = 0.37$).

Procedure

At the start of each lab session, we requested permission to film each group of students in the class. Students, lab academics and PhD lab supervisors were advised that we were filming with the aim of improving the quality of lab teaching. We provided a consent to film form for each student, a number of whom declined consent and they were not filmed. We took care that these students would not show up in the background of other videos, and confirmed for those that did consent, that the video would not be seen by the tutors marking any assessed element of the labs. For details regarding the

10.3389/feduc.2023.1194968

number of sessions recorded and length of each session see Supplementary material.

We filmed lab sessions between 7 November 2019 and 9 March 2020 (pre-COVID). There were four cameras at each location. Where possible, the majority of sessions were filmed by a researcher with some distance from the student classes being filmed. The researchers taking consent and setting up cameras were also briefed that the work was to improve lab quality (rather than specifically to look at gender). We targeted mixed gender groups where possible in order to get a wider sample, as single gender groups were in the large majority, even where there was a rough gender balance in overall student numbers. The students either self-assigned their working groups in the lab or were assigned by us to ensure a gender split, yet not mentioning the study aims to the group. For disciplines with a smaller number of women than men in the labs, we worked with the lecturing staff to set up the lab groups to ensure that the groups were mixed where possible, and to avoid all-male and all-female groups.

With filming complete, three undergraduate interns were given access to the recordings to code them. The interns completed a timing template of all participants' involvement in the activities (using the headings: equipment, observing, recording data, group instruction, group write up, other actively engaged, and other-disengaged see Supplementary material). The timings were recorded in hours, minutes and seconds. After each video recording had been coded, the researchers transformed the time to seconds, and checked for data accuracy with regards to the initial coding. A check was run between 10% of the raw coded data and the converted SPSS file to check for conversion accuracy, which found 84% accuracy.

Results

Preliminary analyses showed that the data was not normally distributed (Shapiro–Wilk Test of Normality for all variables, $p < 0.001$). Hence, we conducted an independent samples Mann–Whitney U test (gender: female vs. male) on each of the lab activities. As we partially predicted, women spent less time using equipment in lab activities compared to male students, $U = 7021.5$, $p = 0.01$. There were no differences between men and women in the time that students spent recording data, $U = 7529$, $p = 0.08$, or observing in lab sessions, $U = 7820.5$, $p = 0.20$. For means according to gender, see Table 3.

Discussion

As we expected, when the time in lab based practical activities was measured, gender differences emerged. Following H3, female students spent less time using equipment during lab based practical activities than male students.

General discussion

Across one study including two phases, a cross-sectional survey of perceived time spent and an observational study of actual time spent, we analysed how the time spent in lab-based STEM practicals might be gendered. In line with our predictions, female students actually spent less time using equipment, compared to male students. Previous research has shown that using equipment is a key aspect of students' experiences at lab, as it is associated with students' engagement with their discipline (Keskin-Geçer and Zengin, 2015). Furthermore, specific activities within the lab-such as using equipment-have shown to be critical in understanding students' development of their identity as scientists (Doucette et al., 2020), which is associated with their sense of belonging to their discipline (Chen et al., 2020). For instance, engaging in practical activities in the lab (e.g., using equipment) contributes to students' perceptions of competence (perceived abilities to participate in lab based practical activities), performance (perceived practical work at the lab) and recognition (being recognised by others as part of the science group). All are key elements of Science and STEM identity (Carlone and Johnson, 2007). Hence, the fact that women actually spent less time using equipment might be detrimental not only for their academic engagement and motivation, but also for their career motivation and sense of belonging to the science community.

Unexpectedly, female students did not perceive spending less time using equipment, and were equally satisfied with their time in different lab activities, compared to male students. Self-reports did not support gender differences in these areas. Hence, although the students that participated in Study 1a and 1b were not the same (see Limitations and future research section), it is concerning that students perceived to spend the same amount of time using equipment than the rest of their peers, and were satisfied with this time. These results might be explained due to the environments where female students conduct lab activities: if female students are more likely to participate in disciplines where they are the majority, then they will perceive to spend similar time as their peers on activities (as their peers are also women), or perceive that the higher levels of women are a sign of more equality in their labs. However, previous research has shown that numbers are not enough, and even in STEM disciplines where women outnumber men (e.g., veterinary, life sciences), inequalities persist (Begeny et al., 2020; Bloodhart et al., 2020). Indeed, women participate in STEM more than previously, but they tend to participate in "female" disciplines congruent with gender stereotypical roles (Garcia-Retamero and López-Zafra, 2006). Our study (including both 1a and 1b) results can provide further support to this idea, demonstrating that although female students in STEM might perceive to spend equal time in lab activities compared to male students, they actually do not. This dissonance could be problematic, as a perception of equality—despite the differences with the actual time in this activity—may lead to a false sense of equality in the lab, and in turn, reproduce traditional

TABLE 3 Time spent in practical lab-based settings by gender (in seconds).

	Women M (SD)	Men M (SD)	Total M (SD)
1. Time using equipment	215.45 (273.52)	382.80 (483.92)	348.51 (453.62)
2. Time recording data/taking notes	284.93 (395.56)	221.36 (370.35)	234.39 (375.91)
3. Time observing	410.16 (492.34)	326.07 (457.84)	343.30 (465.58)

gender roles in STEM. Hence, while these results clearly demonstrate that gender inequality in practical STEM activities do persist, the magnitude of this inequality may be masked by a perceived sense of equality that flows from the higher numbers of women enrolled in the disciplines.

Furthermore, our study results contribute to previous research looking at inequalities in STEM settings in laboratories (Batty and Reilly, 2022); classrooms (Wieselmann et al., 2019), and online learning (Nurramadhani et al., 2021) with a novel approach to explore gender inequalities in STEM, looking at lab practical activities and the time that students perceive and spend on them. Our results provide evidence of the importance in analysing STEM educational settings, and considering their particularities in terms of learning contribution and group dynamics.

Theoretical and practical implications

Our findings contribute to the understanding of gender differences in the practical delivery of STEM education. In particular, within the context of practical laboratory activities and the persistence of gender inequality in the distribution of lab-based tasks. The novel approach of assessing data from multiple perspectives (e.g., self-reports in Study 1a, observation in Study 1b) provided a more rounded understanding of the extent of gender inequality in STEM, as well as how students perceived these inequalities. This is important given the incongruencies we found between perceptions and reality. The disparity between perceptions and reality regarding access to resources has also been demonstrated elsewhere, namely in the field of mentoring, potentially to the detriment of women's career advancement. Previous experimental research by Welsh and Diehn (2018) on the mentoring relationship found that women perceived that they were receiving more mentoring than men, even when provided with the same descriptions of the mentoring relationship. This potentially leaves women at risk of missing opportunities for introductions and career progress, as they do not fully experience the benefits of mentoring, such as developing a network and learning informal knowledge to navigate one's career. Moreover, our research can contribute further to the calibration literature (Alexander, 2013), providing initial evidence supporting the theory that the potential mismatch between students' actual task performance and their judgement about the same performance can also be expanded to other settings and measures. In this case, lab practical sessions and time. Considering the implications discussed early about the importance of lab activities for students' learning, academic and identity processes, calibration might inform how lab activities and gender differences also have an impact on students' meta-cognitive skills and strategic behaviour towards scientific participation.

Our research explores a distinctive setting (lab practical activities) where perceived/actual time differences can also lead to detrimental consequences for women equally access to educational opportunities and development of skills in their field. Indeed, the disparities in the results, in terms of perception of time spent and actual time spent using equipment showed that to analyse the effectiveness of gender equality interventions in STEM, different settings need to be considered (such as equipment use in lab-based practicals), as well as different sources of analysis to evaluate how organisations are working to improve gender equality in STEM.

Practical lab-based work is important not only in terms of students learning experiences, but also to help them develop their identities as scientists (Doucette et al., 2020). Our findings contribute to understanding further factors that may shape differences in perception between genders and what it means to identify with being a scientist (e.g., Starr, 2018; Chen et al., 2020), in this case, the time that students spend in different lab activities. Indeed, measuring time in lab activities, from our knowledge, has not been explored in gender equality research in STEM and, following our study results, it might be an important factor to consider when students identification with their discipline (e.g., "being a scientist") is analysed.

Our study showed that gender proportions in discipline participation was associated with students' time satisfaction with lab activities. Although this observation can have positive outcomes, for example that more equal settings are beneficial for all students, it can also mean that the increase in the enrolment of women in STEM disciplines might lead students to perceive a false sense of gender equality in educational settings. Our study provides initial evidence that numbers are not enough, and that multiple sources of evidence (e.g., surveys, observations) need to be included when gender equality is discussed.

These results also have practical implications. Most interventions aimed at improving gender equality in STEM higher education settings have focused their efforts on motivating women to apply and enrol in STEM disciplines (e.g., National Academies of Science, Engineering, and Medicine, 2020). Although this is a critical first step, organisations also need to promote diligence to ensure equal learning opportunities continue within the setting after women have enrolled in these disciplines.

Additionally, our study has implications for how gender equality interventions are conceived and implemented. To be fully effective, these interventions must be conceived considering not only students self-report, but also group settings. Furthermore, interventions must be implemented in different STEM educational settings beyond classrooms, where students can also learn different skills and develop a sense of their identity as scientists (Doucette et al., 2020).

Limitations and future research

Both phases of our study have a number of limitations that need to be acknowledged. Firstly, Study1b showed a non-normal distribution, and despite having a sample higher than 300 participants, we decided to conduct a non parametric test. Hence, these results need to be interpreted cautiously and more follow up studies need to be conducted to establish time differences in lab activities. Moreover, the samples from both phases were different, that is, the students that participated in the self-report study were not the same ones participating in the lab observation study. Hence, although students were part of the same college and department, we cannot follow up on whether the same students that perceived spending equal time using equipment, for example, were the same students that actually spent less/more time using equipment. Future studies could compare not only students' time perception and satisfaction between groups (e.g., women and men), but also within the groups, looking at differences in how students perceive and spend time in the lab. Similarly, Study 1a and 1b focused on gender differences, without taking into account discipline differences (comparing disciplines recognised with higher female participation to disciplines with lower female participation). Future studies need to acknowledge this aspect, creating studies that compare gender and discipline differences, as well as the interaction of both groups (e.g., how women in certain disciplines perceive their participation versus their actual participation). Moreover, our samples in both phases were unequal in terms of gender and discipline distribution (Study 1a included a majority of women, and Study 1b a majority of men). Future studies should better balance the proportion of participants.

Thirdly, our study focused on gender without considering other contextual factors. For example, in Study 1a we measured "perception" of time without including how level of perceptions might interact with other socio-cognitive variables, such as gender stereotypes (see McKinnon and O'Connell, 2020). Indeed, observational data alone do not allow us to determine whether gender differences are indeed explained only by gender (Eddy and Brownell, 2016). Furthermore, how students approach educational activities in STEM disciplines has been associated with previous experiences in education (Bian et al., 2017), and exposure to gender stereotypes in the media (Steinke, 2017), as well as other intersectional identities (National Academies of Sciences, Engineering, and Medicine, 2020), such as race (Ireland et al., 2018). For instance, future experimental research could manipulate the salience of gender stereotypes such as the "secretary" (Doucette et al., 2020) and examine whether they are still serving to influence both female and male student's expectations of role distribution, and thus their active participation in practical lab sessions.

Fourth, Study 1a included the use of single-item questions. Although previous research has showed that the use of single-item measures can be useful, valid, and reliable (Fisher et al., 2016; Allen et al., 2022), it is important to acknowledge the potential limitations of single-item measures in our study. The single item measures used were not previously tested nor validated in prior studies. Future research needs to test (a) the reliability and validity of the single item measures used in these studies, (b) previously used single item measures, and (c) multiple item measures in similar studies.

Finally, although our study included a total sample of over 500 participants, both phases demonstrated a small effect. Future studies should replicate this study, and include a larger sample to ensure the generalisability of our findings. Another limitation that could affect the replicability of our findings is the challenges of setting up cameras to record in the lab, which can prove challenging in terms of human and material resources. Future research needs to consider the implementation of high quality and the quantity of equipment available to conduct recording research.

Conclusion

Our study contributes to the limited literature focusing on gender inequalities in lab-based activities within STEM disciplines in undergraduate programmes. With this study, we propose a novel approach to investigate gender inequalities, that is, the use of observation in educational settings and the measure of the time that students spend in different activities. Our research shows that focusing on students' self-reports is important, but not enough, to evaluate gender equality, and multiple approaches and methodologies are needed to analyse the state of gender inequalities. Indeed, paradoxically, in STEM disciplines numbers are not enough, and we need to look at not only the levels of participation of women in STEM, but also the quality of their participation. Otherwise, universities will contribute to a false sense of inequality and leaving women outside critical educational experiences for their discipline and identities as scientists.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by College of Social Sciences and International Studies Ethics Committee. University of Exeter. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

DF: data curation, formal analysis, project administration, validation, visualisation, writing—original draft, and writing—review and editing. SW: formal analysis, visualisation, and writing—original draft. HS and PC: conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, and writing—review and editing. MR: conceptualisation, data curation, funding acquisition, investigation, methodology, resources, software, supervision, validation, and writing—review and editing. All authors contributed to the article and approved the submitted version.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2023.1194968/full#supplementary-material

References

Alexander, P. A. (2013). Calibration: what is it and why it matters? An introduction to the special issue on calibrating calibration. *Learn. Instr.* 24, 1–3. doi: 10.1016/j.learninstruc.2012.10.003

Allen, M. S., Iliescu, D., and Greiff, S. (2022). Single item measures in psychological science: a call to action. *Eur. J. Psychol. Assess.* 38, 1–5. doi: 10.1027/1015-5759/a000699

Andrus, S., Jacobs, C., and Kuriloff, P. (2018). Miles to go: the continuing quest for gender equity in the classroom. *Phi Delta Kappan* 100, 46–50. doi: 10.1177/0031721718803570

Arnado, A. A., Pene, A. J. P., Fuentes, C. J. F., and Astilla, K. M. (2022). Fostering sustainable STEM education: attitudes and self-efficacy beliefs of STEM teachers in conducting laboratory activities. *Int. J. Stud. Educ. Sci.* 3, 54–74. doi: 10.46328/ijses.33

Batty, L., and Reilly, K. (2022). Understanding barriers to participation within undergraduate STEM laboratories: towards development of an inclusive curriculum. *J. Biol. Educ.*, 1–23. doi: 10.1080/00219266.2021.2012227

Begeny, C. T., Ryan, M. K., Moss-Racusin, C. A., and Ravetz, G. (2020). In some professions, women have become well represented, yet gender bias persists—perpetuated by those who think it is not happening. *Sci. Adv.* 6:eaba7814. doi: 10.1126/sciadv.aba7814

Bian, L., Leslie, S., 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

Blickenstaff, J. C. (2005). Women and science careers: leaky pipeline or gender filter? *Gend. Educ.* 17, 369–386. doi: 10.1080/09540250500145072

Bloodhart, B., Balgopal, M. M., Casper, A. M. A., Sample McMeeking, L. B., and Fischer, E. V. (2020). Outperforming yet undervalued: undergraduate women in STEM. *PLoS One* 15:e0234685. doi: 10.1371/journal.pone.0234685

Boston, J. S., and Cimpian, A. (2018). How do we encourage gifted girls to pursue and succeed in science and engineering? *Gift. Child Today* 41, 196–207. doi: 10.1177/1076217518786955

Boucher, K. L., Fuesting, M. A., Diekman, A. B., and Murphy, M. C. (2017). Can I work with and help others in this field? How communal goals influence interest and participation in STEM fields. *Front. Psychol.* 8:901. doi: 10.3389/fpsyg.2017.00901

Burkam, D. T., Lee, V. E., and Smerdon, B. A. (1997). Gender and science learning early in high school: subject matter and laboratory experiences. *Am. Educ. Res. J.* 34, 297–331. doi: 10.3102/00028312034002297

Buser, T., Niederle, M., and Oosterbeek, H. (2012). *Gender, competitiveness and career choices* National Bureau of Economic Research, Inc. No. 18576, NBER working papers, Available at: https://EconPapers.repec.org/RePEc:nbr:nberwo:18576.

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

Carlone, H. B., and Johnson, A. (2007). Understanding the science experiences of successful women of color: science identity as an analytic lens. *J. Res. Sci. Teach.* 44, 1187–1218. doi: 10.1002/tea.20237

Casad, B. J., Hale, P., and Wachs, F. L. (2017). Stereotype threat among girls: differences by gender identity and math education context. *Psychol. Women Q.* 41, 513–529. doi: 10.1177/0361684317711412

Chen, S., Binning, K. R., Manke, K. J., Brady, S. T., McGreevy, E. M., Betancur, L., et al. (2020). Am I a science person? A strong science identity bolsters minority students' sense of belonging and performance in college. *Pers. Soc. Psychol. Bull.* 47, 593–606. doi: 10.1177/0146167220936480

Cheryan, S., Meltzoff, A. N., and Kim, S. (2011a). Classrooms matter: the design of virtual classrooms influences gender disparities in computer science classes. *Comput. Educ.* 57, 1825–1835. doi: 10.1016/j.compedu.2011.02.004

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B. J., and Kim, S. (2011b). 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

Cui, Y., Hong, J. C., Tsai, C. R., and Ye, J. H. (2022). How does hands-on making attitude predict epistemic curiosity and science, technology, engineering, and mathematics career interests? Evidence from an international exhibition of young inventors. *Front. Psychol.* 20:13. doi: 10.3389/fpsyg.2022.859179

Department for Education and Behavioural Insights Team. (2020). *Applying behavioural insights to increase female students' uptake of STEM subjects at A Level*. Government Social Research. Available at: https://www.gov.uk/government/publications/applying-behavioural-insights-to-increase-female-students-uptake-of-stem-subjects-at-a-level

Doucette, D., Clark, R. D., and Singh, C. (2020). Hermione and the secretary: how gendered task division in introductory physics labs can disrupt equitable learning. *Eur. J. Phys.* 41:035702. doi: 10.1088/1361-6404/ab7831

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. doi: 10.1103/PhysRevPhysEducRes.12.020106

Eddy, S. L., Converse, M., and Wenderoth, M. P. (2015). PORTAAL: a classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE Life Sci. Educ.* 14:ar23. doi: 10.1187/cbe.14-06-0095

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fisher, G. G., Matthews, R. A., and Gibbons, A. M. (2016). Developing and investigating the use of single-item measures in organizational research. *J. Occup. Health Psychol.* 21, 3–23. doi: 10.1037/a0039139

Garcia-Retamero, R., and López-Zafra, E. (2006). Prejudice against women in male-congenial environments: perceptions of gender role congruity in leadership. *Sex Roles* 55, 51–61. doi: 10.1007/s11199-006-9068-1

Gladstone, J. R., and Cimpian, A. (2021). Which role models are effective for which students? A systematic review and four recommendations for maximizing the effectiveness of role models in STEM. *Int. J. STEM Educ.* 8:59. doi: 10.1186/s40594-021-00315-x

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. doi: 10.1103/PhysRevPhysEducRes.12.020120

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Res. Organ. Behav.* 32, 113–135. doi: 10.1016/j.riob.2012.11.003

Holmes, K., Berger, N., Mackenzie, E., Attard, C., Johnson, P., Fitzmaurice, O., et al. (2022). Editorial: the impact of place-based contextualised curriculum on student engagement and motivation in STEM education. *Front. Educ.* 6:826656. doi: 10.3389/feduc.2021.826656

Ireland, D. T., Freeman, K. E., Winston-Proctor, C. E., DeLaine, K. D., McDonald Lowe, S., and Woodson, K. M. (2018). (Un)hidden figures: a synthesis of research examining the intersectional experiences of black women and girls in STEM education. *Rev. Res. Educ.* 42, 226–254. doi: 10.3102/0091732X18759072

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. Gend. Stud.* 18, 341–356. doi: 10.1080/09589230903260019

Jackson, C., and Nyström, A.-S. (2015). 'Smart students get perfect scores in tests without studying much': why is an effortless achiever identity attractive, and for whom is it possible? *Res. Pap. Educ.* 30, 393–410. doi: 10.1080/02671522.2014.970226

Jones, S., and Myhill, D. (2004). 'Troublesome boys' and 'compliant girls': gender identity and perceptions of achievement and underachievement. *Br. J. Sociol. Educ.* 25, 547–561. doi: 10.1080/0142569042000252044

Keskin-Geçer, A., and Zengin, R. (2015). Science teachers' attitudes towards laboratory Practises and problems encountered. *Int. J. Educ. Res.* 3 Available at: https://ijern.com/journal/2015/November-2015/12.pdf

Kim, A. Y., Sinatra, G. M., and Seyranian, V. (2018). Developing a STEM identity among young women: a social identity perspective. *Rev. Educ. Res.* 88, 589–625. doi: 10.3102/0034654318779957

Kirkup, G., Zalevski, A., Maruyama, T., and Batool, I. (2010) *Women and men in science, engineering and technology: the UK statistics guide 2010* Bradford, UK Resource Centre for Women in Science, Engineering and Technology.

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

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

Lopatto, D. (2004). Survey of undergraduate research experiences (SURE): first findings. *Cell Biol. Educ.* 3, 270–277. doi: 10.1187/cbe.04-07-0045

Lucht, P. (2015). De-gendering STEM - lessons learned from an ethnographic study of a physics laboratory. *International Journal of Gender, Science and Technology* 8, 67–81.

McKinnon, M., and O'Connell, C. (2020). Perceptions of stereotypes applied to women who publicly communicate their STEM work. *Human. Soc. Sci. Commun.* 7:160. doi: 10.1057/s41599-020-00654-0

National Academies of Sciences, Engineering, and Medicine (2020). *Promising practices for addressing the underrepresentation of women in science, engineering, and medicine: opening doors*. Washington, DC: The National Academies Press.

Nurramadhani, A., Lathifah, S. S., and Yamin. (2021). Gender differences in science learning: how is students' questioning quality through STEM based e-module? *J. Phys. Conf. Ser.* 1806:012134. doi: 10.1088/1742-6596/1806/1/012134

Prieto-Rodriguez, E., Sincock, K., and Blackmore, K. (2020). STEM initiatives matter: results from a systematic review of secondary school interventions for girls. *Int. J. Sci. Educ.* 42, 1144–1161. doi: 10.1080/09500693.2020.1749909

Pronin, E., Steele, C. M., and Ross, L. (2004). Identity bifurcation in response to stereotype threat: women and mathematics. *J. Exp. Soc. Psychol.* 40, 152–168. doi: 10.1016/S0022-1031(03)00088-X

Quinn, K. N., McGill, K. L., Kelley, M. M., Smith, E. M., and Holmes, N. G. (2018). Who does what now? How physics lab instruction impacts student behaviors. 2018 Physics Education Research Conference Proceedings. doi: 10.1119/perc.2018.pr.Quinn

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., et al. (2018). Anatomy of STEM teaching in north American universities. *Science* 359, 1468–1470. doi: 10.1126/science.aap8892

Starr, C. R. (2018). I m not a science nerd! *Psychol. Women Q.* 42, 489–503. doi: 10.1177/0361684318793848

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

STEM Women (2022). *Women in STEM Statistics*. Retrieved from https://www.stemwomen.com/women-in-stem-percentages-of-women-in-stem-statistics.

Trusz, S. (2020). Why do females choose to study humanities or social sciences, while males prefer technology or science? Some intrapersonal and interpersonal predictors. *Soc. Psychol. Educ.* 23, 615–639. doi: 10.1007/s11218-020-09551-5

Velasco, J. B., Knedeisen, A., Xue, D., Vickrey, T. L., Abebe, M., and Stains, M. (2016). Characterizing instructional practices in the laboratory: the laboratory observation protocol for undergraduate STEM. *J. Chem. Educ.* 93, 1191–1203. doi: 10.1021/acs.jchemed.6b00062

Vincent-Lancrin, S. (2008), "The reversal of gender inequalities in higher education: an on-going trend"higher education to 2030. *Higher education to 2030 volume 1: demography*. OECD Publishing, Paris.

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

Wan, T., Geraets, A. A., Doty, C. M., Saitta, E. K. H., and Chini, J. J. (2020). Characterizing science graduate teaching assistants' instructional practices in reformed laboratories and tutorials. *Int. J. STEM Educ.* 7:30. doi: 10.1186/s40594-020-00229-0

Welsh, E. T., and Diehn, E. W. (2018). Mentoring and gender: perception is not reality. *Career Dev. Int.* 23, 346–359. doi: 10.1108/cdi-11-2017-0198

Wieselmann, J. R., Dare, E. A., Dare, E. A., and Roehrig, G. H. (2019). "I just do what the boys tell me": exploring small group student interactions in an integrated STEM unit. *J. Res. Sci. Teach.* 57, 112–144. doi: 10.1002/tea.21587

Wilkins-Yel, K. G., Arnold, A., Bekki, J., Natarajan, M., Bernstein, B., and Randall, A. K. (2022). "I can't push off my own mental health": chilly STEM climates, mental health, and STEM persistence among Black, Latina, and White graduate women. *Sex Roles* 86, 208–232. doi: 10.1007/s11199-021-01262-1

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 06 October 2023
DOI 10.3389/feduc.2023.1137563

Check for updates

OPEN ACCESS

EDITED BY
Karen Blackmore,
University of Worcester, United Kingdom

REVIEWED BY
Ove Edvard Hatlevik,
Oslo Metropolitan University, Norway
Dan Whittaker,
University of Worcester, United Kingdom

*CORRESPONDENCE
Tove Leming
✉ tove.leming@uit.no

RECEIVED 04 January 2023
ACCEPTED 21 September 2023
PUBLISHED 06 October 2023

CITATION
Leming T and Johanson LB (2023) "And then I check to see if it looks legit" – digital critical competence in teacher education.
Front. Educ. 8:1137563.
doi: 10.3389/feduc.2023.1137563

"And then I check to see if it looks legit" – digital critical competence in teacher education

Tove Leming* and Lisbeth Bergum Johanson

Department of Education, Faculty of Humanities, Social Sciences and Education, UiT The Arctic University of Norway, Tromsø, Norway

This study investigates pre-service teachers' understanding and use of their own digital critical competence. In the Norwegian teacher training programmes, pre-service teachers at all levels are required to develop critical reflection skills and learn basic digital skills. They have to be able to communicate digitally and at the same time be able to reflect on how developments in technology entail a growing need for critical assessment of digital media. Through a qualitative approach, we interviewed 17 social studies pre-service teachers at UiT The Arctic University of Norway and asked how they understand and use digital critical competence. Our study shows that the students' understanding of digital critical competence as a concept in the early stage of their education is mainly linked to source criticism and can largely be classified as a procedural understanding. A procedural approach means acting without any consideration of the underlying intentions; it requires little thought, and cognitive operations are ignored. More experienced pre-service teachers can link digital critical competence more clearly to the teaching profession and the school context and can reflect on didactic perspectives. They have a more norm-critical approach and question how information and knowledge are established. We find that they have developed a more critical and reflexive approach. The implication for teacher education is that digital critical competence should have a space in all subjects, not just social studies. This will help improve the quality of education and equip the pre-service teachers for everyday life as critical and reflexive teachers.

KEYWORDS

critical thinking, teacher education, digital competence, pre-service student teacher, critical-reflexive

Introduction

Social media are vital to every part of our life, and the Internet is a source of quick and easy access to information. Everyone can share and spread their thoughts on political, cultural, economic, historical, and social matters from different perspectives and in whatever "truth" they believe in. Castellvi et al. (2020, p. 1) argue that the digital age has changed how we interact through the media, how we learn, how we communicate, and how we access information. Grut (2021, p. 9) claims that the sheer volume of digital information makes it very challenging to ascertain whether the content is reliable or not. Castellvi et al. (2020, p. 1) write that "the information it offers often reproduces hegemonic narratives and can be lacking in rigor, or even biased or false." Several studies, according to Breakstone et al. (2022, p. 963–964), shows that university students struggle to evaluate information that are online. This also aligns with a previous study of McGrew et al. (2018, p. 165) where they find that students from middle school, high school and college had difficulties evaluating sources on internet. Critical thinking is therefore vital for assessing stories and information found online and is an important key competence for both today and the future (Hulin, 2018, p. 86). Critical thinking is an important factor in digital competence (Nascimbeni and Vosloo, 2019, p. 11). Digital competence can broadly be defined as "the confident, critical and creative use of ICT to achieve goals related to work, employability, learning, leisure, inclusion and/or participation in society" [defined by Kampylis et al. (2015, p. 39); Mattar et al. (2022, p. 12)]. To develop digital competences, organisations such as UNESCO, the European Commission and International Telecommunication Union, commercial actors such as Microsoft and Google (Nascimbeni and Vosloo, 2019, p. 5), educationists, and schools all over the world are now participating in educating citizens to become digitally critically competent.

In the Norwegian education context, digital competence is of great importance. In a white paper concerning the school of the future (Official Norwegian Reports NOU, 2015: 8, p. 26), it is stated that the use of technology is a significant part of our lives and that digital competence is a prerequisite to participating in education and society. As an answer to the future demands and challenges of a digital society, the government of Norway has adopted a national curriculum that focuses on 21st-century skills (Bakken and Andersson-Bakken, 2021, p. 729). These skills, among others, include critical thinking, technology skills, and digital literacy and are incorporated in our understanding and use of the concept of digital competence.

The latest national curriculum reform called "Fagfornyelsen"[1] [literally: subject renewal] included a strengthening of critical thinking as a topic and highlighted critical thinking as a fundamental aspect of the underlying core values for education (Norwegian Directorate for Education and Training, 2017, p. 5–6).[2] In the curriculum of pre-service social studies teachers, it is emphasised that the subject must help educate pupils to become "critically thinking citizens" who possess digital competence (Norwegian Directorate for Education and Training, 2020, pp. 5, 6 and 12). In other words, they must become digitally critically competent. In a study of two cohorts of pre-service teachers' professional digital competence, it was found that after the teacher education reform of 2017 the students had improved their professional digital competence (Andresen et al., 2022). Despite this study and these explicit and clearly defined requirements for digital critical competence, research in the national context (Røkenes and Krumsvik, 2016; Instefjord and Munthe, 2017; Instefjord, 2018; Langset et al., 2018; Mikkelsen and Rist, 2018; Gudmundsdottir and Hatlevik, 2020) shows that there is still a considerable mismatch between the requirements imposed in teacher education programmes and the digital competencies teacher graduates have when they enter the profession as schoolteachers. In addition, recent research (Weyergang and Frønes, 2020) has revealed that there is broad variation in Norwegian primary school pupils' ability for critical thinking. Furthermore, Ferguson and Krange (2020) found that teachers have not received sufficient training in how to develop pupils' critical thinking. This may have consequences for how the teacher education contributes to the development of the pre-service teachers' critical thinking and their digital competence. Thus, there is a need for more studies on critical digital competence in education and in teacher education in particular.

At UiT The Arctic University of Norway (hereinafter UiT), critical thinking and digital competence are explicit goals in pre-service social studies education. It is particularly central in one of the master's degrees courses, namely social studies didactics and critical thinking (Programme Description, 2022). However, we do not have much information about the level of digital critical competence among social studies pre-service teachers, neither when they enter the programme nor when they finalize their education. By conducting interviews of 17 pre-service social study teachers in the second and fourth year of their education programme, the current study investigated the following research question: how do social studies pre-service teachers understand and use their own digital critical competence?

To answer this research question, we formulated some sub-questions for elaboration. We asked the pre-service teachers how they understand the concept of critical thinking and how they define the concept of digital critical competence. In addition, we asked the pre-service teachers to describe the process they use when they assess digital sources in a critical way. Exploring this research question might indicate if the pre-service teacher's perception of their own digital critical competence is in line with their practice.

We will start by exploring the theoretical framework of the concept of digital competence in an educational context. Thereafter we will present the methodological approach of the study. We chose to combine the results and discussions, based on the interviewees' interpretation of the concepts and the description of practice together with the progression and development of competence. We thereafter summarize our findings and highlight the didactical implications for teacher education programmes in general.

Theoretical framework: critical thinking and digital critical competence

Because critical thinking is a part of digital competence, it is important to elaborate on our understanding and use of the concept critical thinking. There is no single definition of critical thinking, and it can be interpreted in myriad ways. Ryen et al. (2019), for example, define critical thinking as the ability to use rationality in an independent and investigative way to identify the premises that underlie various claims. Ferrer and Wetlesen (2019, p. 11) offer a similar definition, saying that critical thinking is "reflective thinking and an active and evaluative approach to assumptions and accepted truths." They stress that critical thinking is more than source work and methodology alone. Critical thinking is a creative process whereby a

1 The curriculum upgrade was introduced in the Norwegian school system from 2020 as a process for "developing and introducing new curricula in the National Curriculum for Knowledge Promotion" (Norwegian Directorate for Education and Training, 2020).

2 The Norwegian Directorate for Education and Training has the overarching responsibility for education at the kindergarten, primary, and lower secondary and upper secondary levels. The Directorate's role is to ensure equitable offerings, administer and interpret legislation, develop curricula and framework plans, be responsible for mapping, examinations, and national tests, obtain and collate knowledge, and contribute to the development of competencies and skills (regjeringa.no).

person reflects on different possibilities and perspectives. Alexander (2014, p. 470) and Ferguson and Krange (2020, p. 196) argues that critical thinking involves thinking "deeply and flexibly about important matters." Lim (2015) and Ryen et al. (2019, pp. 3–4) focuses on the fact that critical thinking should not only be a purely technical exercise, but that people ought to be trained in justifying their arguments and that awareness about this process must be raised. Critical thinkers can see connections in addition to being able to apply a diversity of perspectives and, not least, to clarify the importance of the power perspective. The power perspective is an important focus of Røthing (2020, p. 27), who believes that critical thinking involves critical reflection on power relations and structures in society. Critical thinking is necessary in order to make informed choices and is thus a prerequisite for democratic participation. It should therefore be possible to translate critical thinking into action in order to foster active citizens (Ferrer and Wetlesen, 2019). Thus, in our study critical thinking entails more than purely a technical approach.

There are several different concepts that coexist, like digital skills, 21st-century skills, information and computer literacy, and digital competence, but according to Nascimbeni and Vosloo (2019, p. 10) digital literacy can be seen as an umbrella concept, and they propose that digital literacy is the most appropriate concept to use. This is "because it clearly entails skills, uses and outcomes." Additionally, they believe digital literacy entails a more holistic approach that also includes critical thinking.

However, in our context we find digital competence to be a more appropriate and relevant concept. According to Amdam et al. (2022), the concept of competence in Norway and in the other Nordic countries often has a broader meaning than in English-speaking countries. It includes skills, literacies, and bildung (Erstad et al., 2021). Furthermore, the concept of digital competence is normally applied in policy documents and curricula in Norway in addition to its use in research (Røkenes and Krumsvik, 2016; Erstad et al., 2021; Andresen et al., 2022; Krumsvik, 2022). Thus, digital competence "can be broadly defined as the confident, critical, and creative use of ICT to achieve goals related to work, employability, learning, leisure, inclusion and/or participation in society. Digital competence is a transversal key competence which enables the acquisition of other key competencies. It is related to many of the so-called '21st Century skills', which should be acquired by all citizens, to ensure their active participation in society and the economy" (Joint Research Centre et al., 2012; Kelentrić et al., 2017, p. 12). This is in line with how the EU's DigiComp project defines digital competence. In our understanding and interpretation digital competence entails critical thinking, and we thus use the term digital critical competence.

Furthermore, in our study we find it helpful to apply the concepts of 'procedural', 'critical', and 'reflexive' that Hulin (2018, p. 85) applied in a study to connect critical thinking to a digital context. The aim of Hulin's study was how to teach students critical thinking by combining a procedural, critical, and reflexive approach. A procedural approach means acting without any consideration of the underlying intentions; it requires little thought, and the cognitive operations are ignored. A critical and reflexive approach, in contrast, entails assessing and understanding the background for the action; for example, by identifying whether power structures have set the premises for how we act. Applying these perspectives to social studies in teacher education, it is obvious that digital critical competence entails more than just a scientific method and more than just a technical and instrumental understanding. It entails an extended analytical and scientific approach to a phenomenon that is more closely aligned to the descriptions given above. This is largely consistent with Hulin's (2018) distinction between a procedural approach and a critical and reflexive approach, and this is the framework of our discussion.

Methods

Type of study

The aim of this study was to establish a wider picture on social studies pre-service teachers' perspectives on digital critical competence. Our research question was how the social studies pre-service teachers understand and use their own digital critical competence.

We chose a qualitative approach to get more knowledge about the context and the pre-service teachers' overall perception of the topic. To obtain and interpret pre-service teachers' knowledge and practice, we applied a phenomenological approach, where the pre-service teachers presented their experiences and perspectives. Our research question and sub questions derived originally from a recent quantitative survey on digital competence by Johanson et al. (2022). One finding from this previous study, where we focused on the pre-service teachers' digital competence with respect to digital interaction and communication when they first entered the programme, was that the pre-service teachers themselves believed that they had adequate skills in terms of finding and assessing digital information. Quantitative surveys may have some limitations with respect to understanding the context in which the questions are asked, and there might be a risk that the pre-service teachers' voices are not sufficiently reflected in these kinds of surveys (Creswell, 2006). In the present study, we sought a deeper understanding of this context using more qualitative approaches. Our aim was to get more knowledge about how they understand "adequate skills" and how they find and assess digital information.

Context and participants

The context of the study was pre-service teachers in UiT's 5-year primary and lower secondary teacher education programme. There was a total of 60 pre-service teachers in this programme, and 17 of them participated in the study. All of them had elected social studies as one of various subjects in the programme. The researchers and authors are both professors in social studies, one based at Campus Tromsø and the other at Campus Alta.

We recruited six pre-service teachers from the second year of the primary and lower secondary teacher education programme from both Campus Alta and Campus Tromsø who had participated in the previous year's survey. The interviews were conducted on Zoom and lasted between 20 and 45 min. We thereafter conducted one focus group interview with a group of 11 pre-service teachers. This was conducted in person and lasted 1 h. This group consisted of fourth-year social studies pre-service teachers in the primary and lower secondary teacher education programme from Campus Tromsø. This interview was conducted using the same interview guide.

Data sources

We used a strategically selected sample of interviewees in the sense that the pre-service teachers all had chosen social studies as part of their teacher education. The reason for this was that we were particularly interested in ascertaining what prior digital competence social studies pre-service teachers have when they embark on their education. This choice means that we cannot generalise the results to apply to all pre-service teachers. We developed an interview guide with a total of seven questions where we asked them about their daily use of digital platforms in general, and if they could identify any challenges in finding relevant information using internet. We also inquired how they understood the concepts *critical thinking* and *digital critical competence*. Further we asked what they had learned earlier about digital critical competence. To get a picture on how they practiced their knowledge, we invited them to describe in concrete detail how they find information online and how they assess the information they find. We also asked them to describe in detail how they considered the information they found; how they for example sorted out not valid information. All the participants were informed about the project both in writing and orally and consented to the use of the data in an anonymised form. Both researchers took part in the interviews, in the processing of the responses, and in the subsequent analysis of the material.

Furthermore, using both individual interviews and a focus group interview may have some limitation to our study. The focus group have the advantages of building on each other's arguments, and when stimulating the interaction between participants we may generate different data and knowledge. We were aware of this and made sure that all the students were involved when answering and discussing the questions. Also focus interviews may give more nuanced answers and show differences in opinions and experiences. In addition, the second year students participated by using zoom (because of the corona-restrictions) which may have played a role in how we read their body language or how well we communicated with the students. This seemed not to be a problem, as the students had used digital communications for quite some time and were familiar to this way of communication. We find that the answers from both groups, even though applying different approaches, contributes to the research question and provides and identifies preservice teachers' perspectives. However, due to this methodological limitation, a follow up study further exploring our research question is desirable.

Analyses

We chose a stepwise deductive-inductive model to process the data material (Tjora, 2018, 2021). In this method, induction and deduction are used alternatingly in a separate and stepwise process (Tjora, 2018). Consequently, we started out by reading through and interpreting the material – first separately, then together – to get an idea of what kind of key words and concepts characterised the pre-service teachers' responses. In an inductive stepwise process, we then sorted these into categories with codes. Some of the codes we worked with were source criticism, including reliability and credibility, independence, multiple perspectives, a lot of information and time-consuming, interpretation and threat to democracy, digital education, citizenship, and professional relevance. In our second review and discussion of the codes and different interpretations of these, we found that some of the concepts the students used, could be associated with a more analytic and reflexive understanding.

The use of advanced, in contrast to everyday language, indicated that there is a gap in competence amongst the informants. Two main categories emerged: "basic understanding and experience of digital critical competence" and "advanced understanding and experience of digital critical competence." In the subsequent process, which comprised a more deductive approach, we linked these categories to theories and concepts that could help explain the phenomenon and to cast light on the context of the information (Hammersley and Atkinson, 1996). In this phase, we analysed the material using theoretical models, primarily Hulin's model on procedural and critical-reflexive approach.

As researchers and social studies teachers, we are interested in this type of competence, and our different backgrounds in history and interdisciplinary social sciences, respectively, mean that we may have different experiences of and different knowledge about critical digital competence. We believe that this has contributed positively to the study in that it has broadened our approach to the topic. Doing a critical examination of our knowledge and different perspectives could increase validity of the study, but we choose here to focus on the students' perspectives.

We do have some ethical concerns regarding the selection and processing of data because we both have professional relations to the pre-service teachers we interviewed. We are teachers in their subjects, and this asymmetrical power relationship may have coloured the data we received. Are they giving the answer they think the teacher wants to hear, or are they saying what they themselves really think? At the beginning of the interviews, we spent some time explaining that this was not something that would be evaluated or have an impact on them in terms of their grades or assessment. It is important to bear this perspective in mind in all phases of a study, and especially when considering the reflexive element of the researcher role (Hammersley and Atkinson, 1996).

How do the pre-service teachers understand and apply the concepts of critical thinking and digital critical competence?

In this first part, we will present the findings from the interviews of how pre-service teachers, in their second year, understand the concept of critical thinking. These perceptions contribute to giving us a picture of how they understand digital critical competence. We thereafter analyse and discuss how the different pre-service teacher groups understand digital critical competence.

Critical thinking – many understandings?

When we asked the pre-service teachers in their second year of study how they understand critical thinking, we note that all the respondents stated that it is important to question what is being told or to question the sources, i.e., what we understand as source criticism. For example, one of them said: "The first thing that comes into my mind is source criticism." Most of the pre-service teachers claimed that they check the sources to find who created the information provided, and they want to ascertain whether the person or people who have provided the information can be trusted or not. The pre-service teachers also stated that it is important to understand why something is said or written. For example, one of the interviewees said that critical thinking is about being critical towards the information, why it is being conveyed, and who is conveying it. Another mentioned that it is important to know who you can trust and that the sources must therefore be checked several times. Several of them referred to reliability as a key concept. They also stated that it is important to know there are different voices and opinions:

> "It is a matter of questioning what you see and perceive, and how things work. So, in a more general level (…) being able to see things from multiple angles and form an opinion about the information available to you (…)".

Many of the pre-service teachers stressed the importance of being able to see that there are multiple perspectives and the ability to think independently. One of them pointed out that after starting on the degree programme they realized that most books and articles used in the programme were based on something that had previously been written that the content was an interpretation, and that they should again interpret what had been written. The pre-service teachers stated this can be challenging, "What was the real meaning behind the original message?" The more experienced pre-service teachers mentioned source criticism as an important part of critical thinking but found that the concept had a wider meaning. For example, for them critical thinking could include a norm-critical approach.

Discussion: how do the pre-service teachers understand critical thinking?

As described in the introduction and in the theory section, in this article we regard critical thinking as more than mere competence in source criticism or what we might call a technical or procedural approach (Hulin, 2018). Critical thinking involves deconstructing the narratives in books, online, and elsewhere, and it must be possible to assess and compare claims against each other, and we need to be able to distinguish between how narratives and history are used and misused. Critical thinking entails seeing things from different perspectives, asking critical questions on distributions of power, and being critical towards how structures and norms in society are established and maintained.

What emerged from the interviews with the pre-service teachers who were still at a relatively early stage of their studies was that most of them thought that critical thinking is largely a matter of source criticism. They were particularly concerned with whether the sources they read, or the author are credible and whether the sources are reliable. This might not be so surprising because they all were introduced to source criticism in the early stage of their teacher education. Most of the pre-service teachers reported that they were familiar with source criticism from secondary school. This understanding was limited compared to how we understand critical thinking or how literature such as Ferrer and Wetlesen (2019) or Røthing (2020) describe it. However, even though most of the pre-service teachers perceived critical thinking as source criticism, with its more technical or formulaic approach to the sources, some, and especially the more experienced pre-service teachers, recognized that the sources and information they receive represent interpretations and different perspectives on reality. This can be seen as an advanced understanding of critical thinking that is closer to a more critical and reflexive approach as Hulin (2018) and Ferrer and Wetlesen (2019) define it.

How do pre-service teachers understand digital critical competence?

Findings from second-year pre-service teachers

When asked what they think digital critical competence is, most of the pre-service teachers answered that it means having a critical attitude towards digital media. One participant said:

"Personally, I've never heard that term before, but I would imagine it means practicing source criticism and using digital resources in a critical way. Being able to think independently about whether what you find online is credible." Another commented: "and perhaps not using information you find online entirely indiscriminately. If you cannot confirm that it was written by an expert, then maybe it is better not to use this opinion uncritically." This is in line with how they perceive critical thinking in general, i.e., that it is a technical exercise and basic understanding. The pre-service teachers found the sheer volume of information challenging, and it could be difficult to "get to the bottom" of the information.

> "There is simply so much that is written, and you get so many hits for every single search word. And then when different websites say different things, what should we trust, what is correct? For example, if two reputable websites say different things, how are you supposed to know which is right, using the competencies you have? It makes it very challenging".

One of the more experienced pre-service teachers made an interesting statement, linking this to challenges to democracy:

> "We get so much information today that you don't really have time to be critical…. From a societal perspective, this is incredibly worrying, because we are being bombarded with so much information that is not correct. We do not have time to check any corrections that are added afterwards…. These corrections tend to get overlooked because we've already got the information we were looking for. We trust that it's right… and this, ultimately, poses a threat to democracy. It's worrying."

Findings from the experienced pre-service teachers

The fourth-year pre-service teachers linked the concept of digital critical competence to digital education, which encompasses aspects such as qualified judgement, impact assessment, and credibility in addition to elements of source criticism. They were concerned with "how we as teachers handle sensitive information about students." One of the examples they cited was awareness regarding the use of e-mail for sensitive information in relation to school–home relations. Some focused on raising awareness amongst pupils on how certain actors on digital platforms can have an agenda for influencing, for example, political issues. Another area has a commercial nature, influencing choices and habits in their favour. They were especially concerned with how so-called "echo chambers" emerged:

> "If you, for example, are in an echo-chamber online, you will find support for the information you already have obtained. It can be difficult to judge whether you are inside an echo-chamber or not, and then it will be even more difficult to find if the information you have is realistic or not".

Discussion: how do pre-service teachers understand digital critical competence?

When it comes to understanding the concept of digital critical competence, the less experienced pre-service teachers found it far more complex than the pre-service teachers expressed in the quantitative survey we conducted earlier (Johanson et al., 2022). In the present study we found that the pre-service teachers expressed difficulties in identifying which sources are credible. In addition, source-critical checking of online sources is time-consuming and resource-intensive. Nevertheless, it seems that they have the tools to do this. Digital critical competence is particularly linked to the vast amount of information available online. In this respect, this kind of an approach corresponds to what we would categorize as a basic understanding of digital critical competence and what Hulin (2018) calls a procedural or technical approach. This involves a formulaic approach to sources and information. However, we regard the pre-service teachers' ability to recognize that the sheer volume of information can pose a "threat to democracy" as a form of advanced understanding of digital critical competence. This can be interpreted as a critical and reflexive approach. Hulin (2018) points out that it is important to assess, understand, and identify the structure of power. This understanding is in line with Kelentrić et al. (2017, p. 12) and with official policy documents, including the Norwegian curricula for various levels of education.

The more experienced pre-service teachers highlighted the importance of a critical approach to digital learning resources. As expected, these pre-service teachers had other ways of expressing themselves through more advanced use of academic terminology coupled with both theoretical knowledge and personal experience from practice. They related digital critical competence to democracy and citizenship. As one of them put it: "So, digital critical competence is a bit of a step up from digital literacy. It is a step up in terms of being a good citizen."

When questioning whether the pre-service teachers defined or understood digital critical competence differently from critical thinking, we found that they were somewhat alike as, for example, source criticism. However, evaluating information online is much more complex and challenging because of the vast amount of information on the Internet. Additionally, some worried that the vast amount of information on the Internet is a "threat to democracy." The more experienced pre-service teachers linked digital critical thinking not only to source criticism, but also to ethics, democracy, and citizenship and to the choice and evaluation of digital tools or platforms. They expressed that digital critical competence should be a part of digital education in general.

Assessing digital information – a difficult exercise?

Having established how the pre-service teachers understands both critical and digital critical competence, we will discuss how they proceed when they assess digital information.

Findings of second-year pre-service teachers

When we asked the less experienced pre-service teachers to describe how they differentiate between reliable and less reliable sources on the Internet and the process they use to assess online sources, the answers varied. Most of them seemed to understand what a reliable source on the Internet was and they seemed to be equipped with quite high level of skills to find and assess digital information. Several of the pre-service teachers cited examples of what they perceived to be adequate and reliable sources. These included government and public websites such as the Norwegian Ministry of Education and Research and the Norwegian Directorate for Education and Training,[3] the national broadcasting company NRK,[4] the Norwegian Digital Learning Arena (NDLA),[5] and the Great Norwegian Encyclopedia (SNL).[6] The commercial television channel TV2 was also regarded as a good source. In addition, the website Forskning.no[7] was referred to as an adequate source, although the pre-service teacher who mentioned this source pointed out that the website is selective in its interpretation of research when they write an academic article.

> "Journalists who refer only to research papers, and articles that are coloured more by the journalist's opinion of the matter than what the researchers meant. As a result, it becomes more of a populist article. But at least I can look at the references and find the original articles. I've done that on several occasions."

Several of the pre-service teachers gave examples of what they perceived as less reliable sources – usually sensationalist articles or online news outlets that are widely shared on Facebook. Blogs and Wikipedia were also cited as less reliable sources, with statements such as: "I try to use Wikipedia as little as possible" or "we have learned from early age that Wikipedia is not a reliable source." Another said:

> "I remember being taught in middle school, back home … 'always be slightly critical towards what you read on Wikipedia. Scroll down and look at the references Wikipedia cites and go directly to those sources'. We were told that the Great Norwegian

3 Ministry of Education and Research and the Norwegian Directorate for Education and Training.

4 Norwegian Broadcasting Corporation.

5 National Digital Learning Arena, which offers open learning resources.

6 Store Norske Leksikon/the Great Norwegian Encyclopedia (SNL): free and publicly available encyclopaedia written by professionals.

7 Forskning.no (2002–present) is an online newspaper about Norwegian and international research.

> Encyclopedia (SNL) is a better source than Wikipedia for information on specialist subjects. So that's when I had my first encounter with critical digital thinking and all that."

According to one of the pre-service teachers, these kinds of websites spread a lot of fake news on topics such as dieting and immigration, for example. They pointed out that to check the credibility of the sources, you must "examine what the authors claim very closely by checking references to the original source." Another said: "Who has written it, what they have written, and who has published it are perhaps the three main things to check." One of the interviewees said that importance should be attached to the first impression you get when you go to a website: "The first impression, i.e., how it is structured, in a purely aesthetic sense – and then I think I can distinguish between an adequate source and an inadequate source."

Another pre-service teacher gave an example of searching for a specific topic:

> "I was searching for something related to Nazism and stumbled across a website that looked very poorly designed. It appeared to be a right-wing extremist blog, and the content was sheer nonsense, but it looked like the kind of website where information is shared. This was probably an extreme example. While the websites with the best design, such as the Great Norwegian Encyclopedia (SNL) and institutional websites, each tend to have their own distinctive style, I notice how they are organized, since that's what makes the first impression. It's a bonus if the author's name is given at the beginning of a text you're looking at. It helps strengthen the reader's confidence in the website."

An interesting answer that one of the participants gave, and which shows a more advanced understanding of digital critical competence, concerned the use of a noncredible sources, such as sensationalist articles, in a news context. The pre-service teacher stated that this is problematic because this way of writing "pulverizes democracy." According to this pre-service teacher, people who disseminate this type of news are not aware of their role in society. Other pre-service teachers found it difficult to distinguish a reliable source from an inadequate or false source, simply because there is so much information and so many people writing about the same topic, thus making it difficult to sort through it all. Despite this, the pre-service teachers stressed that they are critical of sources, that they check references to the original source, and that they check the academic competence of the author of the text. One of them said that they have a checklist that they run through and that this has now become so ingrained that they do it automatically. They check the address, among other things, to see if there is anyone to contact, "and then I check to see if it looks legit."

Findings of the experienced pre-service teachers

We discussed the differences between reliable and an inadequate or false sources with the more experienced pre-service teachers as well, and how they go about checking this. They felt that their own learning process linked to this was important, primarily to be able to teach pupils "how we can tell that something is not true." Cross-checking information and checking multiple sources were highlighted as important. This is an essential skill because "if you search to find out whether the Earth is flat, you will always find someone who says the Earth is flat" or, as another put it, "You will always find someone online to support your claim."

They went on to describe the method as checking the origin of websites, checking the text to ascertain whether it is an opinion piece or part of a debate, finding out who the authors of the text are, and checking the bibliography and what references have been used. Where reference is made to previous research, it should be clear which research is being referred to. On websites, checks include assessing the layout, determining whether it is a reliable source (for example, the BBC), or if it is a blog, and they mentioned that influencers can exert a strong influence through blogs or social media posts. It is a matter of trying to judge which interests are behind the various sources, such as whether there are any underlying financial or commercial interests. The interviewees described that it is important to be aware that many websites created by non-experts can look very professional and that it sometimes can be difficult to tell the difference. As some of the pre-service teachers summed it up:

> "We talked about what kind of website had posted the information, the purpose behind and goal of this website, and whether there are people or organizations behind the website who want to present their views on a particular topic. You must perform quite a thorough investigation to find out who wrote it."

In this respect, the more experienced pre-service teachers were quite like the second-year pre-service teachers in how they checked whether information or sources are reliable or not. In addition, the group of pre-service teachers reflected on the "echo chamber" effect. According to them, people tend to frequent websites that support the information they have already found. In this context, they pointed out how various news articles, although they may use reliable sources, only present parts of the information to promote their own point of view:

> "…articles published in VG [a Norwegian national daily newspaper] promote a specific view. So, they might have given the date, named the author, and provided information about them and all that, and there may well be some sources at the bottom that show where they have got their information from. Those sources might be a PhD thesis, but it is still not necessarily so that the article you are reading is credible. The person concerned may have only cited a few cherry-picked paragraphs from a large thesis that support their point of view and left out points that criticize the views the author wants to promote."

Discussion: is assessing digital information a difficult exercise?

All of the interviewees were able to give examples of what they consider a reliable source and an inadequate source. However, there was quite wide variation in how they go about assessing whether a source is adequate or not. One interviewee said that they adopt a very thorough approach, while others seem to have a slightly vaguer method. This may be related to the background that the pre-service teachers reported they had. In the individual interviews, we got an impression of the kind of prior knowledge the pre-service teachers had when they embarked on the teacher education programme. Naturally, this prior knowledge varied according to several factors, including age and interest in digital media in general. One of the interviewees had previously worked in the ICT sector, while some had a vocational background. However, all interviewees mentioned lower secondary school and upper secondary school as their most important arena for training related to digital media, with source criticism as the main area they had learned about. Several cited Wikipedia as a concrete example of a source they had learned to be critical towards at school, and they reported that their teachers often mentioned the Great Norwegian Encyclopedia SNL as an example of a more credible source. At the same time, a couple of the interviewees described how they had used Wikipedia as their main source of information in schoolwork without this being corrected. This clear distinction between Wikipedia as an unreliable source and the SNL as a reliable source can serve as an example of a rather unnuanced understanding of what critical digital thinking entails. Critical thinking can involve examining what kind of values are represented in the SNL, thereby raising awareness of the fact that there are multiple ways of describing a phenomenon. The SNL can serve as an example of one of them, and Wikipedia as an example of another. For instance, Brox (2012, 2016) shows in her studies that using Wikipedia in a learning context can contribute to increased critical competence. Through projects where pupils and students themselves contributed to articles on Wikipedia, they became aware of how knowledge is constructed and maintained (Brox, 2012, 2016).

We found that the more experienced pre-service teachers had acquired a more professional vocabulary when using their digital critical competence. They could contextualize knowledge through practice and experience, thus linking it to the school system, teaching profession, parent–home relations, pupils, and learning processes. In other words, they had a more advanced understanding of what critical digital competence entails, more akin to Hulin's (2018) critical and reflexive understanding. In addition, in their development towards becoming a teacher they had gained some professional and academic pegs on which they could hang these concepts.

Even though the second year pre-service teachers' competence could be connected to a more technical understanding of digital critical competence, we found examples of pre-service teachers who had a more critical and reflexive approach, comparable to Hulin's (2018). One example was the pre-service teacher who presented a power-critical view of the role of the media.

Towards a better teacher education?

How can we as providers of teacher education use the experiences we have gained from these surveys to help pre-service teachers develop digital critical competence? When asking pre-service teachers what they miss in their education or would like to learn more about, several of the second-year pre-service teachers responded that they would like more instruction in assessing online sources. They want to learn how to evaluate different types of sources and to develop this digital skill earlier in their studies. In other words, they want more of the kind of formulaic procedural approach to critical digital competence linked to source criticism that is familiar to them from school. This kind of competence is important, especially in view of the sheer volume of the digital information flow. However, it is important that the pre-service teachers work on acquiring more reflexive critical thinking skills related to digital critical competence. The responses from the fourth-year pre-service teachers indicated that this is something that is being worked on during their education. These more experienced pre-service teachers, but also some of the less experienced pre-service teachers, understood that digital critical competence entails much more than simply running through a source criticism checklist – that it also requires a certain level of critical reflection. The implication for teacher education is that digital critical competence should have a space in all subjects, not just social studies. This will help improve the quality of education and equip the pre-service teachers for everyday life as critical and reflexive teachers.

Conclusion

The point of departure for this study was the question of how the social studies pre-service teachers in the teacher education programme understand and use their own digital critical competence. Our findings shows that the pre-service teachers' understanding of critical thinking as a concept in the early stage of the education is mainly linked to source criticism. Pre-service teachers' conceptual understanding of critical thinking evolves as they progress through the teacher education programme. When it comes to the concept of digital critical competence, most of the second-year pre-service teachers have a technical understanding and are concerned with various challenges, for example, the "enormous volume of information." Within the frames of Hulin (2018), this can be classified as a procedural understanding of the concept. A procedural approach means acting without any consideration of the underlying intentions; it requires little thought, and cognitive operations are ignored. More experienced pre-service teachers can link digital critical competence more clearly to the teaching profession and the school context and can also reflect on didactic perspectives. They have a more norm-critical approach and question how information and knowledge are established. We find that they have developed a more critical and reflexive approach (Hulin, 2018). A critical and reflexive approach, by contrast, entails assessing and understanding the background for the action, for example, by identifying which power structures have set the premises for action. It is important to develop critical thinking within digital praxis.

It is not unexpected that pre-service teachers develop their competencies at different levels of their teacher education programme. However, our agenda has been to see and reflect on how the pre-service teachers themselves think about and articulate both a procedural and extended competence within the field of critical digital competence. This study can contribute towards a broader and more nuanced understanding of how these social studies pre-service teachers perceive, reflect upon, and develop their own digital critical competencies.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by SIKT-Norwegian Agency for Shared Services in Education and Research. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

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

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alexander, P. A. (2014). Thinking critically and analytically about critical-analytic thinking: an introduction *Educ. Psychol. Rev.* 26, 469–476. doi: 10.1007/s10648-014-9283-1

Amdam, S. H., Kobberstad, L. R., and Tikkanen, I. (2022). Professional digital competence in strategy and management: a case study of three teacher education programs in Norway. *Nordic J. Digit. Lit.* 17, 16–30. doi: 10.18261/njdl.17.1.2

Andresen, J. K., Tømte, C., and Bergan, I. and, Kovac, V. B. (2022). Professional digital competence in initial teacher education: an examination of differences in two cohorts of pre-service teachers *Nordic J. Digit. Lit.* 17, 61–74. doi: 10.18261/njdl.17.1.5

Bakken, J., and Andersson-Bakken, E. (2021). The textbook as a genre. *J. Curric. Stud.* 53, 729–748. doi: 10.1080/00220272.2021.1929499

Breakstone, J., Smith, M., Ziv, N., and Wineburg, S. (2022). Civic preparation for the digital age: how college students evaluate online sources about social and political issues. *J. High. Educ.* 93, 963–988. doi: 10.1080/00221546.2022.2082783

Brox, H. (2012). The elephant in the room: a place for Wikipedia in higher education? *Nordlit.* 16, 143–155. doi: 10.7557/13.2377

Brox, H. (2016). Troublesome tools: how can Wikipedia editing enhance student teachers' digital skills? *Acta Didactica Norge.* 10, 329–346. doi: 10.5617/adno.2493

Castellvi, J., Diez-Bedmar, M.-C., and Santisteban, A. (2020). Pre-service teachers' critical digital literacy skills and attitudes to address social problems. *Soc. Sci.* 9, 1–11. doi: 10.3390/socsci9080134

Creswell, J. W. (2006). "Understanding mixed methods research" in *Designing and Conducting Mixed Methods Research*. eds. J. W. Creswell and V. L. Plano Clark (Thousand Oaks, CA: SAGE Publications)

Erstad, O., Kjälander, S., and Järvelä, S. (2021). Facing the challenges of 'digital competence' a Nordics agenda for curriculum development for the 21st century. *Nordic J. Digit. Lit.* 16, 77–87. doi: 10.18261/issn.1891-943x-2021-02-04

Ferguson, L. E., and Krange, I. (2020). Hvordan fremme kritisk tenkning i grunnskolen? *Norsk Pedagogisk Tidsskrift.* 2, 194–205. doi: 10.18261/issn.1504-2987-2020-02-09

Ferrer, M., and Wetlesen, A. (eds.) (2019). *Kritisk Tenkning i Samfunnsfag*. Oslo: Universitetsforlaget.

Grut, S. (2021). *Digital Kildekritikk*. Oslo: Gyldendal Norsk Forlag AS.

Gudmundsdottir, G. B., and Hatlevik, O. E. (2020). "I just Google it" – developing professional digital competence and preparing student teachers to exercise responsible ICT use. *Nordic J. Comp. Int. Educ.* 4, 39–35. doi: 10.7577/njcie.375

Hammersley, M., and Atkinson, P. (1996). *Feltmetodikk. Grunnlaget for Feltarbeid og Feltforskning*. Oslo: Ad Notam Gyldendal.

Hulin, T. (2018). How critical thinking should be at the heart of digital interactions? *J. Commun. Stud.* 11, 85–105.

Instefjord, E. J. (2018). Professional Digital Competence in Teacher Education: A Mixed Methods Study of the Emphasis on and Integration of Professional Digital Competence in Teacher Education [Doctoral Dissertation, University of Stavanger]. Available at: https://uis.brage.unit.no/uisxmlui/bitstream/handle/11250/2501440/Elen_J_Instefjord.pdf?sequence=1&isAllowed=y (Accessed March 01, 2023).

Instefjord, E. J., and Munthe, E. (2017). Education digitally competent teachers: a study of integration of professional digital competence in teacher education. *Teach. Teach. Educ.* 67, 37–45. doi: 10.1016/j.tate.2017.05.016

Johanson, L. B., Leming, T., Johannessen, B.-H., and Solhaug, T. (2022). Competence in digital interaction and communication – a study of first-year preservice teachers' competence in digital interaction and communication at the start of their teacher education. *Teach. Educ.* 58, 270–288. doi: 10.1080/08878730.2022.2122095

Joint Research Centre, Institute for Prospective Technological Studies, and Ferrari, A. (2012). *Digital competence in practice: an analysis of frameworks*. Publications Office. Available at: https://data.europa.eu/doi/10.2791/82116

Kampylis, P., Punie, Y., and Devine, J. (2015). *Promoting effective digital-age learning: a European framework for digitally-competent educational organisations, EUR 27599 EN*. Luxembourg: Publications Office of the European Union.

Kelentrić, M., Helland, K., and Arstorp, A.-T. (2017). Rammeverk for Lærerens Profesjonsfaglige Digital Kompetanse. Senter for IKT i Utdanningen. Available at: https://www.udir.no/contentassets/081d3aef2e4747b096387aba163691e4/pfdk-rammeverk-2018.pdf, (Accessed March 01, 2023)

Krumsvik, R. (2022). Digital competence across the education and health sector. *Nordic J. Digit. Lit.* 17, 149–154. doi: 10.18261/njdl.17.3.1

Langset, I. D., Jacobsen, D. Y., and Haugsbakken, H. (2018). Digital professional development: towards a collaborative learning approach for taking higher education into the digitalized age. *Nordic J. Digit. Lit.* 13, 24–39. doi: 10.18261/ISSN.1891-943X-2018-01-03

Lim, L. (2015). Critical thinking, social education and the curriculum: foregrounding a social and relational epistemology. *Curriculum J.* 26, 4–23. doi: 10.1080/09585176.2014.975733

Mattar, J., Santos, C. C., and Cuque, L. M. (2022). Analysis and comparison of international digital competence frameworks for education. *Educ. Sci.* 12:932. doi: 10.3390/educsci12120932

McGrew, S., Breakstone, J., Ortega, T., Smith, M., and Wineburg, S. (2018). Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory Res. Soc. Educ.* 46, 165–193. doi: 10.1080/00933104.2017.1416320

Mikkelsen, R., and Rist, M. (2018). "Lærerstudenters digitale praksismøte [when preservice teachers encounter digital praxis]" in *Restart: Å Være Digital i Skole og Utdanning*. eds. L. B. Johanson and S. S. Karlsen (Oslo: Universitetsforlaget), 163–177.

Nascimbeni, F., and Vosloo, S. (2019). Digital Literacy for Children: Exploring Definitions and Frameworks. UNICEF Office of Global Insight and Policy. Available at: https://www.unicef.org/globalinsight/media/1271/file/%20UNICEF-Global-Insight-digital-literacy-scoping-paper-2020.pdf, (Accessed March 01, 2023)

Norwegian Directorate for Education and Training (2017). Core Curriculum – Values and Principles for Primary and Secondary Education. Available at: https://file:///C:/Users/ljo053/Downloads/Overordnet%20del%20p%C3%A5%20engelsk.pdf (Accessed March 01, 2023)

Norwegian Directorate for Education and Training (2020). Curriculum for Social Studies (SAF01-04). Available at: https://www.udir.no/lk20/saf01-04?lang=eng (Accessed March 01, 2023)

Official Norwegian Reports NOU (2015). The School of the Future – Renewal of Subjects and Competences. Available at: https://www.regjeringen.no/en/dokumenter/nou-2015-8/id2417001/ (Accessed March 01, 2023)

Programme Description (2022). Studieplan. Master i Grunnskolelærerutdanning for 5.10. UiT Norges Arktiske Universitet. Available at: https://uit.no/Content/729822/cache=20222609093702/studieplan%205-10.pdf (Accessed March 01, 2023)

Røkenes, F. M., and Krumsvik, R. J. (2016). Prepared to teach ESL with ICT? A study of digital competence in Norwegian teacher education. *Comput. Educ.* 97, 1–20. doi: 10.1016/j.compedu.2016.02.014

Røthing, Å. (2020). *Mangfoldskompetanse og Kritisk Tenkning*. Oslo: Cappelen Damm Akademisk.

Ryen, E., Jøsok, E., Jegstad, K. M., and Sandvik, M. (2019). Kritisk Tenkning. Bedre Skole. 31. Available at: https://utdanningsforskning.no/artikler/2019/kritisk-tenkning/ (Accessed March 01, 2023)

Tjora, A. (2018). *Viten Skapt – Kvalitativ Analyse og Teoriutvikling*. Oslo: Cappelen Damm Akademisk.

Tjora, A. (2021). *Kvalitative Forskningsmetoder I Praksis*. Oslo: Gyldendal Akademisk.

Weyergang, C., and Frønes, T. S. (2020). "Å lese kritisk: Elevers vurderinger av teksters troverdighet og pålitelighet" in *Like Muligheter Til God Leseforståelse? 20 år Med Lesing i PISA*. eds. I. C. Weyergang and T. S. Frønes (Oslo: Universitetsforlaget), 166–195.

Frontiers in Education

Explores education and its importance for individuals and society

A multidisciplinary journal that explores research-based approaches to education for human development. It focuses on the global challenges and opportunities education faces, ultimately aiming to improve educational outcomes.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact