NEW CHALLENGES IN THE RESEARCH OF ACADEMIC ACHIEVEMENT: MEASURES, METHODS, AND RESULTS

EDITED BY: Juan Luis Castejon, José Carlos Núñez, Raquel Gilar-Corbi and Iván Manuel Jorrín Abellán

PUBLISHED IN: Frontiers in Psychology and Frontiers in Education

frontiers Research Topics

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88966-507-5
DOI 10.3389/978-2-88966-507-5

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

NEW CHALLENGES IN THE RESEARCH OF ACADEMIC ACHIEVEMENT: MEASURES, METHODS, AND RESULTS

Topic Editors:
Juan Luis Castejon, University of Alicante, Spain
José Carlos Núñez, Universidad de Oviedo Mieres, Spain
Raquel Gilar-Corbi, University of Alicante, Spain
Iván Manuel Jorrín Abellán, Kennesaw State University, United States

Citation: Castejon, J. L., Núñez, J. C., Gilar-Corbi, R., Abellán, I. M. J., eds. (2021). New Challenges in the Research of Academic Achievement: Measures, Methods, and Results. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-507-5

Table of Contents

05 ***Optimization: In-Depth Examination and Proposition***
Huy Phuong Phan, Bing Hiong Ngu and Alexander Seeshing Yeung

21 ***The Educational Situation Quality Model: A New Tool to Explain and Improve Academic Achievement and Course Satisfaction***
Fernando Doménech-Betoret, Amparo Gómez-Artiga and Laura Abellán-Roselló

30 ***The Influence of Self-Regulation Behaviors on University Students' Intentions of Persistance***
Ana Bernardo, María Esteban, Antonio Cervero, Rebeca Cerezo and Francisco Javier Herrero

38 ***Children's Academic, Artistic, and Athletic Competencies: Successes are in the Eye of the Beholder***
Sarah J. Racz, Diane L. Putnick, Gianluca Esposito and Marc H. Bornstein

53 ***School Readiness Profiles and Growth in Academic Achievement***
Qianqian Pan, Kim T. Trang, Hailey R. Love and Jonathan Templin

70 ***Relationship Between Digital Creativity, Parenting Style, and Adolescent Performance***
María del Carmen Pérez-Fuentes, María del Mar Molero Jurado, Nieves Fátima Oropesa Ruiz, María del Mar Simón Márquez and José Jesús Gázquez Linares

80 ***Academic Performance in Adolescent Students: The Role of Parenting Styles and Socio-Demographic Factors – A Cross Sectional Study From Peshawar, Pakistan***
Sarwat Masud, Syed Hamza Mufarrih, Nada Qaisar Qureshi, Fahad Khan, Saad Khan and Muhammad Naseem Khan

92 ***Identification of Factors Associated With School Effectiveness With Data Mining Techniques: Testing a New Approach***
Fernando Martínez-Abad

105 ***Evaluating Impact on Motivation and Academic Performance of a Game-Based Learning Experience Using Kahoot***
Andrés Fuster-Guilló, María Luisa Pertegal-Felices, Antonio Jimeno-Morenilla, Jorge Azorín-López, María Luisa Rico-Soliveres and Felipe Restrepo-Calle

113 ***Relationship Between SES and Academic Achievement of Junior High School Students in China: The Mediating Effect of Self-Concept***
Shifeng Li, Qiongying Xu and Ruixue Xia

120 ***Can Academic Achievement in Primary School Students Be Improved Through Teacher Training on Emotional Intelligence as a Key Academic Competency?***
Teresa Pozo-Rico and Ivan Sandoval

128 ***Longitudinal Stability of Reading Difficulties: Examining the Effects of Measurement Error, Cut-Offs, and Buffer Zones in Identification***
Maria Psyridou, Asko Tolvanen, Marja-Kristiina Lerkkanen, Anna-Maija Poikkeus and Minna Torppa

142 ***Academic Achievement, Self-Concept, Personality and Emotional Intelligence in Primary Education. Analysis by Gender and Cultural Group***
Lucía Herrera, Mohamed Al-Lal and Laila Mohamed

155 ***Gender Differences in Mathematics Motivation: Differential Effects on Performance in Primary Education***
Susana Rodríguez, Bibiana Regueiro, Isabel Piñeiro, Iris Estévez and Antonio Valle

163 ***University Entrance Examinations in Spain: Using the Construct Comparability Approach to Analyze Standards Quality***
Alejandro Veas, Leandro Navas, Teresa Pozo-Rico and Pablo Miñano

171 ***Schools That Promote the Improvement of Academic Performance and the Success of All Students***
Pilar Arnaiz-Sánchez, Remedios de Haro, Salvador Alcaraz and Ana Belén Mirete Ruiz

179 ***Students' Evaluation of Teaching and Their Academic Achievement in a Higher Education Institution of Ecuador***
Tarquino Sánchez, Raquel Gilar-Corbi, Juan-Luis Castejón, Jack Vidal and Jaime León

189 ***How Schools Affect Student Well-Being: A Cross-Cultural Approach in 35 OECD Countries***
Elena Govorova, Isabel Benítez and José Muñiz

203 ***Predicting Mathematics Achievement in Secondary Education: The Role of Cognitive, Motivational, and Emotional Variables***
Amanda Abín, José Carlos Núñez, Celestino Rodríguez, Marisol Cueli, Trinidad García and Pedro Rosário

213 ***Design, Implementation and Evaluation of an Emotional Education Program: Effects on Academic Performance***
María-José Mira-Galvañ and Raquel Gilar-Corbi

222 ***Future Time Perspective and the Achievement of Optimal Best: Reflections, Conceptualizations, and Future Directions for Development***
Huy P. Phan, Bing H. Ngu and Kelvin McQueen

235 ***Neural Network Model for Predicting Student Failure in the Academic Leveling Course of Escuela Politécnica Nacional***
Iván Sandoval-Palis, David Naranjo, Raquel Gilar-Corbi and Teresa Pozo-Rico

frontiers in Psychology

CONCEPTUAL ANALYSIS
published: 19 June 2019
doi: 10.3389/fpsyg.2019.01398

Optimization: In-Depth Examination and Proposition

Huy Phuong Phan[1], Bing Hiong Ngu[1] and Alexander Seeshing Yeung[2]*

[1]School of Education, Faculty of Humanities, Arts, Social Sciences and Education, University of New England, Armidale, NSW, Australia, [2]Institute for Positive Psychology and Education, Australian Catholic University, Sydney, NSW, Australia

One notable concept that is of interest is a *person's state of optimal functioning*. Achieving optimal functioning (e.g., subjective well-being at school), aside from personal autonomy, requires some form of "optimization." *Optimization*, we argue, is more than just an "enhancement," a "predictive effect," and/or a "causal flow" between an independent variable (IV) and a dependent variable (DV). We note from existing literature that optimization has often been referred to without a clear, definitive explanation of what this term actually entails. At the same time, we acknowledge that unlike other areas of development (e.g., engagement), no theoretical article is available to explain the concept of optimization. This article considers a number of theoretical tenets for advancement: (1) the tenet of *three major criteria* that could assist in the explanation, assessment, and measurement of optimization, (2) the tenet of the *development of a methodological conceptualization* that could measure and assess optimization, and (3) the tenet of the "quantification" of optimization, and in particular, a proposed *index of optimization* and a corresponding scientific notation of "γ," which we coin as an "optimizing effect." Overall, we contend that this examination is insightful and holistic, seeking clarity into an important topical theme in psychology.

Keywords: optimal functioning, optimization, index of optimization, energization, subjective well-being, positive psychology, cognitive load theory, optimizing effect

OPEN ACCESS

Edited by:
Juan Luis Castejon, University of Alicante, Spain

Reviewed by:
Shi Yongchuan, Wenzhou University, China
Peter Merrotsy, University of Western Australia, Australia

****Correspondence:***
Huy Phuong Phan
hphan2@une.edu.au

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

Received: *17 March 2019*
Accepted: *29 May 2019*
Published: *19 June 2019*

Citation:
Phan HP, Ngu BH and Yeung AS (2019) Optimization: In-Depth Examination and Proposition. Front. Psychol. 10:1398. doi: 10.3389/fpsyg.2019.01398

INTRODUCTION

One notable line of research in psychology that has recently received considerable interest is the *operational nature of optimal functioning*. Optimal functioning, which may be in physical, cognitive, emotional, and/or social terms, emphasizes the importance of a person's *inner strength*, *state of resilience*, *virtue*, and the *maximization in capability* (Source: Applied Psychology: Health and Well-being). Optimal functioning reflects the paradigm of *positive psychology* (Seligman and Csikszentmihalyi, 2000; Seligman, 2010), and may in the context of academia, involve the experience of mastery, and/or the achievement of an exceptional academic result. Optimal functioning in a nonacademic arena, likewise, may indicate a football player's exceptional achievement to score 50 goals in one season, for example. This theoretical concept of optimal functioning is in direct contrast to personal experiences of stagnation and pessimism, highlighting weakness, sub-optimal performance, and minimal potential. The concept of optimal functioning therefore takes a positive perspective. However, what optimal functioning constitutes and how optimization of human functioning operates are not clearly defined and understood.

The aim of this article is to conduct an in-depth examination of the theories related to the concept of optimization and to propose future directions for research advancement.

UNDERSTANDING OPTIMIZATION

Relating to the concept of optimal functioning is a question that we, as researchers, have made concerted attempts to address: *how does a person reach an optimal state of functioning?* This important question has led to our numerous empirical and conceptual undertakings, which specifically focus on the complexity of the operational mechanism of optimal functioning. What causes an exceptional state of functioning? What actually occurs as a state of functioning improves from one level to that of another level? How does the cause of optimal functioning associate with a level of optimal functioning? These three major questions have, to date, formed the central premise of existing research inquiries and our own contributions. Understanding this complexity of optimal functioning (e.g., how a person reaches a state of optimal cognitive functioning) is innovative, especially in terms of educational and social practices for implementation. From the context of successful schooling, for example, we could capitalize on this line of research development and design appropriate educational programs and/or pedagogical strategies, which may closely align with the optimization of students' learning experiences.

The study of the processes of optimal functioning, from our point of view, is emerging and has received moderate attention. We recognize there are some prominent theoretical tenets that have, likewise, considered the improvement of cognitive functioning. For example, Vygotsky's (1978, 1981) *sociocultural theory of cognitive development* stipulates the potent impact of the contextual environment to shape a person's cognitive development. Psychological tools and cultural artifacts, such as mathematical symbols and notations may mediate a student's progress in his/her understanding of problem solving. In particular, Vygotsky (1978) makes reference to an important term, coined as the "zone of proximal development," which depicts the difference between what a person can do without help and what he/she can do with help (e.g., scaffolding). Piaget's (1963) *theory of personal constructivism*, somewhat different from Vygotsky's (1978) theory, contends that cognitive growth arises from a person's experience resolution of disequilibrium *via* means of adaptation. In school contexts, according to Piaget's (1963) theory, effective learning occurs when a child experiences a mental state of cognitive conflict. Learning outcomes that do not stimulate intellectual challenges or "flow" (Csíkszentmihályi, 1990; Seligman and Csikszentmihalyi, 2000) are more likely, in this analysis, to limit enriched cognitive experiences.

In sum then, our brief introduction contends that an optimal state of functioning indicates personal growth, improvement, and exceptional performance. Achieving this optimal state of functioning requires some form of scaffolding from the external world. Notwithstanding existing theoretical contributions (Piaget, 1963; Vygotsky, 1978), one element that has gone amiss is an in-depth examination of the actual operation involved in the achievement of optimal functioning. This operation, from our point of view and proposition, is known as the process of "optimization" (Phan et al., 2019a,b). The term optimization, extensively used in the academic literature (e.g., Freund and Baltes, 1998; Fraillon, 2004; Ziegelmann and Lippke, 2007) is inconsistently explained, and has not been adequately addressed. The true nature of optimization, we argue, is relatively unknown at present in terms of analysis and understanding. What actually occurs during the process of optimization? How does the process of optimization explain a person's optimal best practice? Can the process of optimization be "quantified" and be represented by a scientific notation? These questions indicate the totality of our understanding of optimal functioning.

OPTIMAL FUNCTIONING: AN INTRODUCTION

Optimal functioning is a perceived positive theoretical concept that emphasizes the importance of improved competence, personal best or exceptionality, and a strong sense of motivation and resilience. Optimal functioning situated within the context of academia is also analogously termed as *optimal best practice* (Phan et al., 2016, 2018a) and *personal best* (Martin, 2006, 2011). An analysis of the literature indicates that, likewise, educators and researchers have often referred to the notion of an "optimal condition" for effective learning and enriched schooling experiences. An educator, for example, may consider strategies and/or programs that could stimulate and foster a positive social climate for learning, which in turn could instill a perceived sense of school belonging for students (Goodenow, 1993; Goodenow and Grady, 1993).

Optimal functioning is a central feat of human agency and may apply to different complex contexts in life (Straszewski and Siegel, 2018; Wiese et al., 2018). Optimal functioning, in this case, may consist of different facets – for example, optimal physical functioning, optimal cognitive functioning, optimal emotional functioning, etc. In the areas of health and subjective well-being, researchers have, for example, explored the concept of optimal subjective well-being (Fraillon, 2004; ACU and Erebus International, 2008). This research inquiry, indeed, has led to the propositions of a number of definitions and views about the nature of optimal functioning. The literature review published by the Australian Catholic University (ACU) in 2004 specifically elucidated the essence of optimal functioning, which the researchers expressed their understanding – "maximizing one's potential" (Dunn, 1961; Ryff, 1995), "pursuit of excellence in physical, mental, emotional, and spiritual realm" (Ardell, 1982), "an active process of fulfillment" (Hettler, 1984), "living and working effectively" (Corbin, 1997), "living fully in the natural community" (Witmer and Sweeney, 1998), "resilience and successful community participation" (Weisner, 1998), "holistic, positive emotions" (Stewart-Brown, 2000), "positive emotions, life satisfaction, and absence of negative emotions" (Diener and Biswas-Diener, 2002), "positive feelings and positive psychosocial functioning" (Keyes, 2002), "resilience, satisfaction,

and maximizing one's potential" (Bornstein et al., 2003), "positive feelings and life satisfaction" (Headey and Wooden, 2004), "positive state and satisfaction of needs" (Prilleltensky and Prilleltensky, 2006), and "resilience and maximizing one's potential" (WHO, 2007).

From this theoretical overview (ACU and Erebus International, 2008), a person's achievement of optimal functioning indicates numerous attributes that are positive – for example, self-fulfillment and inner satisfaction, exceptional accomplishment, and enrichment and personal growth. Depending on the nature of the context, a person may experience different types of attributes when he/she achieves optimal functioning. From an educational perspective, optimal functioning in an academic subject may reflect different learning experiences: a student's ability to continuously perform and achieve outstanding results in Year-8 mathematics and receiving an "A" grade at the end of the school term (Phan et al., 2017), or a student's seeking of mastery to know the different pedagogical approaches that could enable in-depth understanding of a topical theme (i.e., ability to solve challenging transfer percentage problems) (Ngu et al., 2018). At the same time, aside from mastery and performance-based accomplishments, optimal functioning may indicate a student's heightened state of motivation (e.g., intrinsic) to persist with his/her studies (Church et al., 2001; Elliot and Murayama, 2008). From a noneducational point of view, likewise, optimal functioning on a daily basis may indicate a person's positive outlook about life, and his/her strong state of personal resolute and resilience to combat health-related matters. Low optimal functioning, in this case, may result in feelings of pessimism and helplessness, and a belief that existing health issues are not worth combatting.

Overall then, from the aforementioned description, we contend that optimal functioning is an important element of a person's development. Optimal functioning, in its simplistic summation, is concerned with an individual state of "change" that a person experiences for the better. Job satisfaction, combatting health, a positive outlook of life, personal best in sports performance, and successful schooling are some examples of a person's positive experience of optimal functioning. Of relevance and significance in this discussion, which we next discuss, is an in-depth analysis and understanding of how an optimal level of functioning is accomplished. For example, within the context of academia, we want to consider in-class pedagogical strategies, school-based educational programs, and/or the use of intellectual capitals to enhance and optimize students' cognitive functioning. This feat concerning the *nature of achievement* of optimal functioning has not been adequately addressed. We do not have clear evidence at present, both conceptually and empirically, to explain how a state of optimal best is ascertained. What are the underlying processes, which may govern our drive to achieve a state of exceptionality?

Our proposition of a detailed conceptualization of optimization, which may explain the intricate processes of achievement of optimal functioning, draws from existing theorizations (e.g., Csíkszentmihályi, 1990; Fraillon, 2004; Phan et al., 2017) and empirical research findings (Martin, 2011; Liem et al., 2012; Phan et al., 2018a,b,c). Optimization, as we conceptualize, is not an outcome or a relationship, but rather depicts an underlying process that in turn "optimizes" an entity in question (e.g., a person's academic learning experience in a subject matter). In addition, we have also recently considered a related theoretical matter, namely, the conceptualization and development of appropriate methodologies that could enable the assessment and validation of optimization. This research-based discourse is innovative as emphasis is placed on researchers' theoretical contributions to the study of a conceptualized inquiry.

THE OPERATIONAL NATURE OF OPTIMIZATION

What is optimization? In the preceding sections, we mentioned that optimization is an intricate process that closely aligns with the achievement of optimal functioning. An examination of the literature indicates that researchers have extensively used the term optimization in their researches (e.g., Freund and Baltes, 1998; Fraillon, 2004; Ziegelmann and Lippke, 2007). We contend there is ambiguity as to what optimization actually entails as a process. From a generic, simple point of view, optimization may be perceived as a "vehicle" that operates to maximize a person's state of functioning from T_1 to T_2. In recent years, researchers in the areas of *subjective well-being* (Fredrickson, 2000; Keyes et al., 2002) and *healthcare and aging for senior citizens* (Freund and Baltes, 1998; Ziegelmann and Lippke, 2007) have made extensive reference to the concept of optimization. For example, in relation to healthcare for senior citizens, a number of researchers have theorized that optimization serves as a process of engagement in goal-directed actions and means to pursue and maintain personally relevant goals (e.g., a goal of adopting and maintaining a physically active lifestyle). In relation to the study of positive psychology, likewise, Noble and McGrath (2008) proposed a *Positive Educational Practices* (PEPs) *Framework* that focuses on five specific foundations of well-being, namely: (1) social and emotional competency, (2) positive emotions, (3) positive relationships, (4) engagement through strengths, and (5) a sense of meaning and purpose. This framework, according to the authors, provides guidance to educators, school administrators, and researchers in the optimization of positive educational initiatives. The PEPs Framework, in this case, facilitates and encourages students to find a sense of meaning at school, and a purpose in life. In a similar vein, Seligman's (e.g., Seligman and Csikszentmihalyi, 2000; Seligman, 2010, 2011) work on the *PERMA Framework* has also acknowledged the importance of happiness, resilience, and personal growth. One central aspect of human endeavor encompasses an inner desire and striving for one to lead and live a meaningful and enriching life.

Other researchers, similarly, have explored other comparable concepts that we believe reflect the relatedness to the process of optimization. Diener (e.g., Diener et al., 2009, 2010) and other colleagues (e.g., Keyes, 2002; Huppert and So, 2013) have explored the concept of *flourishing*, which is defined as a person's experience that life is going well. In a similar vein, a research focus on the proactivity and enrichment of life has

led to the propositions of theoretical constructs such as *thriving* (Su et al., 2014; Wiese et al., 2018), defined as "a state of positive functioning at its fullest range" (Su et al., 2014), and *personal striving* (Phan and Ngu, 2015; Phan et al., 2018a,b,c), defined as "a person's effort attempt to seek out realistic and/or ambitious endeavor for accomplishment" (Phan et al., 2018a,b,c). Flourishing, thriving, and personal striving are in accord with the paradigm of positive psychology (Seligman and Csikszentmihalyi, 2000; Seligman, 2010), and place emphasis on a person's seeking to achieve optimal endeavors.

Understanding the true mechanism of optimization, theoretically and/or empirically, is relatively unknown at this stage. This consideration, in particular, depicts the finer detail of the "steps" involved in the achievement of optimal functioning. A few researchers have, in this instance, provided comparable explanations of the operational nature of optimization. Fraillon's (2004) discussion paper on the subject of student well-being, for example, described an interesting tenet – namely a person's *actual best functioning* (ABF) and his/her subsequent *notional best functioning* (NBF). Optimization, for the author, is defined as the difference between ABF and NBF. From Fraillon's (2004) brief account, Phan et al. (2017) presented an elaborated conceptualization of the relationship between two levels of functioning – *realistic best* (RB) and *optimal best* (OB). Importantly, the authors' conceptualization proposes an important element coined as the "zone of optimization," which is defined as the difference or range between RB and OB. The zone of optimization varies in the magnitude of the difference or range between the two levels of functioning. Moreover, as a point for consideration, the zone of optimization seeks to explain the "amount" of optimization that would be needed to help optimize the achievement of OB from RB.

Fraillon's (2004) initial, but brief description of optimization and Phan et al.'s (2017) subsequent analysis both have provided theoretical grounding for further development. In our own recent research inquiries pertaining to the nature and scope of *mindfulness* (Phan et al., 2019a,b), we offered an expanded perspective and explanation of optimization. Our conceptualization, as shown in **Figure 1**, is more detailed and technical. In terms of different levels of functioning (e.g., RB: Phan et al., 2017), we argue that *time precedence* is an important element for incorporation – in other words, different levels of functioning cannot take place simultaneously.

From **Figure 1**, we propose that optimal functioning is the result of a progression from an existing level of functioning, denoted as L_1, to a level that is more exceptional, denoted as L_2. Mathematically, in this analysis, a person's progression from L_1 to L_2 is denoted as $\Delta_{(L_2-L_1)}$. We argue for the inclusion of time difference because, as Fraillon (2004) and Phan et al. (2017) concur, L_1 is indicative of a what person is capable of at present, whereas L_2 is concerned with his/her maximum outcome. Being able to achieve L_2 from L_1 does not occur instantaneously, but rather requires an adequate timeframe for completion. Hence, from our conceptualization, we equate L_1 to situate at T_1 and L_2 to situate at T_2 – hence, overall, the achievement of optimal functioning may be defined as $\Delta_{(L_2T_2-L_1T_1)}$.

Methodologically, from a quantitative point of view (Tabachnick and Fidell, 2007), we may consider the assessment, measurement, and validation of $\Delta_{(L_2T_2-L_1T_1)}$. Social science's research has used complex quantitative methodological designs to investigate associative and predictive effects of psychological and educational variables. Nonexperimentally, in this instance, we could consider the introduction of a variable A, which is then proposed to help "optimize" the improvement in score of L_1 to L_2 (**Figure 2**). Moreover, we expect to find that $\Delta_{(L_2T_2-L_1T_1)}$ would be positive in value. This proposition, in this case, stipulates an association between Variable *A* and $\Delta_{(L_2T_2-L_1T_1)}$.

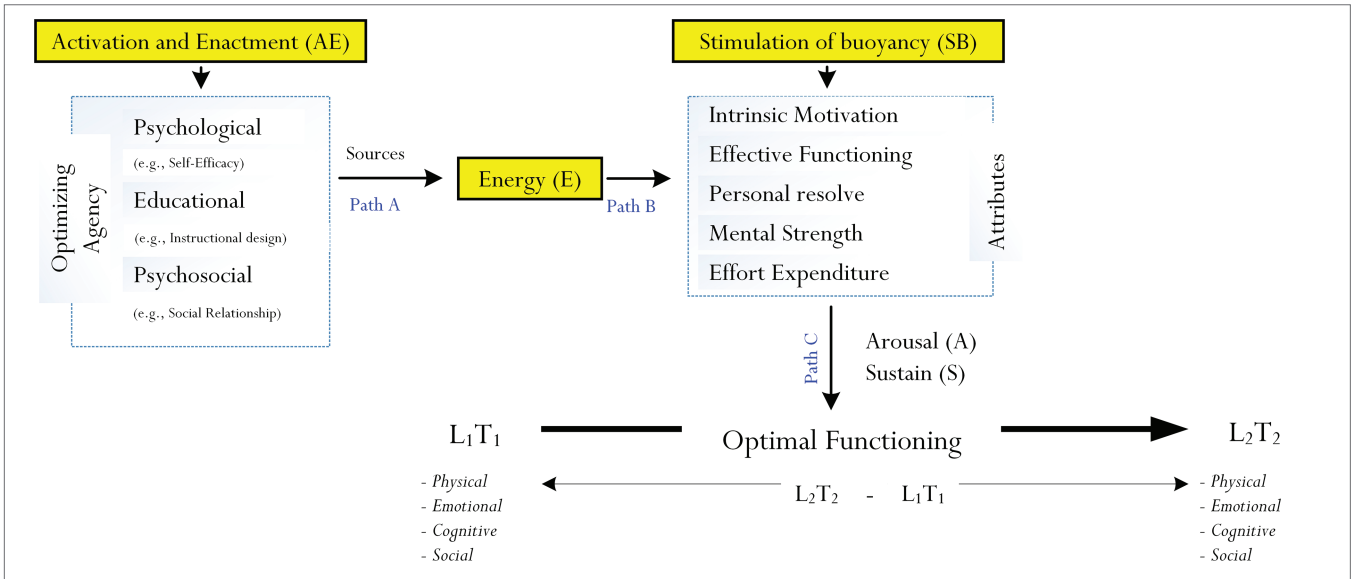


FIGURE 1 | Proposition of the operational nature of the process of optimization.

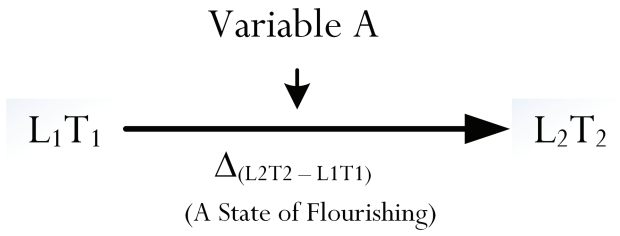


FIGURE 2 | Simple methodological design of optimal functioning.

Baron and Kenny's (1986) seminal publication has resulted in extensive research development into the importance of mediating effects of variables (e.g., Grice et al., 2015; Kline, 2015; Tate, 2015; Trafimow, 2015). In this analysis, referring to our explanation, a central variable *A* could operate to mediate the effect of L_1 at T_1 onto L_2 at T_2 which can be tested in a series of path models: (1) *Test 1*: estimates a model in which only L_1 predicts L_2, (2) *Test 2*: estimates a model in which only Variable *A* predicts L_2, (3) *Test 3*: estimates a model in which only L_1 predicts Variable *A*, and (4) *Test 4*: assesses the reduction in the path from L_1 to L_2 with the introduction of Variable *A* as a mediator. Importantly though, in order to determine a true mediating effect and the potency of a mediator, we need to have evidence of *causal effects*, which in this case requires an experimental treatment or treatments, and the precedence of time difference.

Referring to **Figure 2**, and in tandem with Baron and Kenny's (1986) criteria, it is poignant for us to consider the use of an intervention between T_1 and T_2. Referring to our previous discussion, Variable *A* could be considered as an "optimizing variable" between a determinant, L_1, and an outcome, L_2 – that is: L_1 at T_1 (determinant) → Variable *A* (optimizing variable) → L_2 at T_2 (outcome). In sum, from this introduction of a simple methodological design of optimal functioning, we propose three major criteria:

1. L_1 as an informational source, which then serves as a *point of reference* for the achievement of L_2.
2. The requirement of *timeframe* in order for a person to develop and experience an "increase" in optimal functioning (e.g., emotional functioning) – that is, the existence of multiple time points, which correspond with different levels of functioning – for example, L_1 at T_1, L_2 at T_2, etc.
3. The introduction of an *intervention*, which could operate as an "optimizing agent" in order to enhance and optimize L_1 to L_2.

Aside from a methodological account, we also need to consider the nature of Variable *A*. How does Variable *A* operate in order to facilitate an improvement of L_2 from L_1? The operational nature of Variable *A*, we contend, is intricate for its variation, which closely associates with the *complexity* of the $\Delta_{(L_2T_2-L_1T_1)}$. This proposition is similar to Phan et al.'s (2017) zone of optimization, whereby this "zone" differs and connotes a magnitude in strength for the process of optimization. What does this actually mean? For example, in relation to optimal health functioning, a person may require a substantial time period to combat an illness. An optimizing agent (i.e., Variable *A*) to improve the person's health, in this case, may consist of an effective therapy, varying in intensity in accordance with the difference between L_1 and L_2. In a similar vein, a secondary school student wishing to achieve an optimal level of best practice in the topical theme of essay composition, based on his/her previous experiences, may require a lesser amount of time in terms of optimization. An optimizing agent to address $\Delta_{(L_2T_2-L_1T_1)}$ for writing composition may consist of an academic skills program that also vary in intensity.

An In-Depth Analysis of the Operational Nature of Optimization

Variable *A*, as depicted in **Figure 2**, is proposed to operate as an optimizing agent, which then optimizes and enhances the achievement of L_1 to L_2. This proposition has been extensively detailed in Phan et al.'s (2017) theorization of optimization. From a methodological point of view, we could treat Variable *A* as a mediator between L_1 and L_2, and that there is a direct predictive path from L_1 to L_2. However, aside from its assessment and measurement, we contend that the totality and operational nature of Variable *A* is much more complex than it being viewed and treated as a mediator. Phan et al. (2017), in this case, proposed an underlying process encompassing this complexity, which comprises of two major sub-processes:

1. *Sub-process 1* concerns the "enactment" of different types of *psychological* (e.g., the positive impact of hope: Snyder, 2004), *educational* (e.g., an appropriate pedagogical practice: Ngu et al., 2014), and *psychosocial* (e.g., the complexity of the home environment: Daulta, 2008) *agencies*, which then initiate sub-process 2.
2. *Sub-process 2* involves the activation of the attributes of *persistence*, *effort expenditure*, and *effective functioning*, which then operate to optimize a state of functioning.

Both Fraillon's (2004) consideration and Phan et al.'s (2017) theoretical model of optimization suggests that the process of optimization is more than just a directional association between sub-process 1 and sub-process 2. The "totality" of optimization, we propose, encompasses the stimulation and enrichment of experience of *vitality and buoyancy*. In this analysis, the enactment of optimization is likely to result in an enriched state of energy, strength, and liveliness, which would then enable a person to engage in proactive functioning. This complexity, we contend, is more accurately indicative of what actually occurs within the process of optimization. As shown in **Figure 1**, there are three pathways: Path A, Path B, and Path C. These paths tend to operate in a sequential manner, following these steps:

1. *Step 1*: This step, in line with Phan et al.'s (2017) theorization, is concerned with the activation and enactment (i.e., denoted as "AE") of different psychological (e.g., the impact of hope: Snyder et al., 2000), educational (e.g., an appropriate instructional design: Ngu and Yeung, 2013), and/or psychosocial (e.g., the impact of teacher-student relationship: Roorda et al., 2011) agencies that then serve as sources of a person's state of "energy" (we denote this as "E"). We argue

that the activation and enactment of a particular agent (e.g., psychological agency) does not necessarily influence cognitive or motivational processes directly. Rather, the execution (i.e., activation and enactment) of an optimizing agent (e.g., psychological agent) serves to produce an experience of high "energy." *Energization* is therefore an underlying sub-process of optimization, which in this case entails the experience and indication of vitality and buoyancy, assisting a person to stay focused on task.

Vitality, from our conceptualization, forms a central element of the process of optimization. Personal experience of vitality (e.g., "I feel very energized at the moment") is positive and enriching, predisposing a person to strife for the achievement of an optimal state of functioning. Importantly, of course, vitality is concerned with the observation and reporting of "stamina and liveliness" in cognition, behavior, and/or emotion, contrasting to a state of pessimism and procrastination, which correspondingly associate with a low level of energy. The selection, activation, and enactment of a specific optimizing agent (e.g., the use of hope as a psychological agent), in this case, depend upon the type of optimal functioning that a person is striving to achieve (e.g., optimal cognitive functioning in the area of mathematics).

2. *Step 2*: Personal experience of energization from Step 1 is postulated to stimulate the buoyancy of five comparable psychological attributes: *intrinsic motivation* (i.e., defined as a person's intrinsic motive to persist a course of action – for example, learning Calculus), *personal resolve* (i.e., defined as a person's internal state of decisiveness and resolute to strive for optimal functioning), *effective functioning* (i.e., defined as a person's purposive state of organization, structured thoughts, and behavioral patterns and a deliberate intent to succeed), *mental strength* (i.e., defined as a person's mindset of having the capacity to deal with obstacles, stressors, and pressure – for example, a tennis player is able to bounce back after losing two out of three games in competition tournament), and *effort expenditure* (i.e., a person's conscious attempt to invest effort in order to achieve a particular outcome).

   Further to Fraillon's (2004) brief description and Phan et al.'s (2017) subsequent conceptualization of optimization, we offer an expanded analysis where the sub-process of energization positively influences the operational nature of different types of psychological attributes (e.g., the stimulation of buoyancy of effort) that we perceive as being positive in nature. For example, one psychological attribute that we propose as being potent is a person's internal mental strength to persevere, whereas another notable and related attribute is that of effort expenditure. Our recent correlational research, likewise, has attested to the direct and mediating effects of personal resolve and effective functioning (e.g., Phan et al., 2018a,b,c, 2019a).

3. *Step 3*: The stimulation of buoyancy of intrinsic motivation, personal resolve, effective functioning, mental strength, and effort expenditure *via* positive energy is postulated to *arouse* a person's state of functioning at T_1 at and sustain it to T_2 (e.g., optimal cognitive functioning) (i.e., denoted as "AS"). For example, within the context of secondary schooling, the stimulation of buoyancy of intrinsic motivation may arouse a student's interest in understanding Calculus, which could then help sustain a state of cognitive functioning. The student's aroused state of cognitive functioning of mathematics learning, sustaining in progress from T_1 to T_2, may also involve the stimulation of buoyancy of effort expenditure, personal resolve, etc.

   A person's aroused and sustained state of functioning within a particular context (e.g., academic learning in a subject matter) reflects the effectiveness of the stimulation of buoyancy of different types of psychological attributes. The effective stimulation is facilitated by an enriching state of energy, which arises from the activation and enactment of a relevant educational, psychological, and/or psychosocial agent. A low level of energy, in contrast, is likely to produce the inaction of different types of psychological attributes, resulting in sub-optimal functioning.

In summary, the pivotal components of optimization consist of the activation and enactment of psychological, educational, and psychosocial agencies, which then serve as sources of energy in order to stimulate the buoyancy of the five mentioned comparable attributes. Intrinsic motivation, personal resolve, effective functioning, mental strength, and effort expenditure in turn would individually, and/or in tandem, arouse and sustain a person's progress in functioning from T_1 to T_2. In its simplistic term then, we can summarize the operational nature of optimization as follows: $AE + E + SB$.

Our methodological conceptualization of optimization, which we theoretically derive from previous inquiries (Fraillon, 2004; Phan et al., 2017), partially reflects Vygotsky's (1978, 1981) *sociocultural theory of cognitive development* as well as other theories. For example, aligning to Vygotsky's (1978, 1981) sociocultural theory of cognitive development, our conceptualization highlights three major facets: (1) extensive contributions from an external agent, especially in terms of the provision of opportunities of different types of agency for achievement of optimal functioning (e.g., a child's exposure to different instructional designs/pedagogical practices from a teacher: Ngu et al., 2014), (2) the "internalization" of a particular agent and its "transformation" into a form of positive energy, and (3) the progress in a person's state of functioning (e.g., cognitive functioning), consequently, as a result of external scaffolding. Other researchers, in contrast, have been less clear in their explanatory accounts and descriptions of optimal functioning and optimization-related entities.

METHODOLOGICAL DEVELOPMENT OF OPTIMIZATION

Our theoretical development of optimization has also led us to consider an important inquiry – namely, the development of what we coin as "methodological conceptualization," which

places an emphasis on the measurement, assessment, and evaluation of optimal functioning, and more importantly, the process of optimization. This methodological inquiry has theoretical, methodological, and empirical implications for consideration. From the perspective of quantitative methodology in the social sciences, there is acknowledgment that researchers may use both experimental and nonexperimental research designs to study associative patterns between variables (Bordens and Abbott, 2008; Gravetter and Forzano, 2009; Babbie, 2014). An important question for discussion then, is how do we measure, assess, and quantify the process of optimization?

Assessing Optimal Functioning

From the preceding sections, the concept of optimal functioning reflects a number of analogous attributes, such as "personal best," "maximization in capability," "fullest potential," and "exceptionality." Our previous discussion has emphasized a reference point (e.g., T_1) for benchmarking and comparison – this reference point may be denoted as L_1T_1, where L_1 = initial level of functioning (e.g., cognitive functioning), T_1 = time 1. For example, in the area of mathematics learning, we could consider a student's *current* cognitive competence to solve *linear equations* with one unknown, x (e.g., solve for x: $5x - 11 = 4$), as L_1. This initial level of cognitive functioning (i.e., L_1), known as actual functioning in Fraillon's (2004) terms, or realistic achievement best in Phan et al.'s (2017), is postulated to act as a focal point for benchmarking. The student's optimal level, denoted as L_2 and benchmarked against L_1, may consist of a competence to solve *quadratic equations* with one unknown, x [e.g., solve for x: $(x - 5)^2 = 20$]. L_2 (i.e., learning quadratic equations), compared to L_1 (i.e., learning linear equations), is more advanced in terms of quality and cognitive complexity.

The achievement of optimal functioning from a current state of functioning, reflecting personal growth (i.e., "increase in a state of functioning"), may be defined as follows: $\Delta L_{21} = L_2\ T_2 - L_1\ T_1$. From a quantitative point of view, we need to equate L_1 and L_2 with actual numerical values in order to determine what ΔL_{21} is. The "equating" of L_1, L_2, etc., with specific numerical values is subjective – that is, a student may equate L_1 (i.e., knowing how to solve for x: $5x - 11 = 4$) with an arbitrary value of 12 (e.g., out of 20), say, and L_2 (i.e., knowing how to solve x: $(x - 5)^2 = 20$) with an arbitrary value of 15 (e.g., out of 20), etc. Why do we want to quantify L_1, L_2, etc.? We contend that quantifying L_1, L_2, etc. with numerical values (e.g., 9, 10, 11, …, etc.) makes it relatively easy for researchers to rationalize the meaning of ΔL_{21}, ΔL_{32}, etc. In other words, quantitatively, an optimal level of functioning is more meaningful when it is denoted by a definitive numerical value.

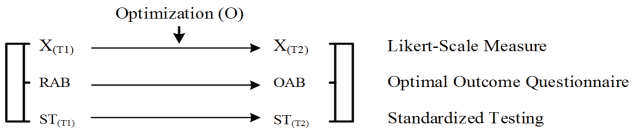
Quantifying different levels of functioning with numerical values (e.g., "Provide an arbitrary score that you believe best describes your current level of emotional functioning"), of course, may pose a few problems for researchers, such as *inconsistency*, *subjective bias*, and *miscalibration*. A student's inexperience in personal reflection, for example, may result in unintentional biased alignment of L_1 (e.g., 3 out of 20), L_2 (e.g., 12 out of 20), etc. when, in fact, this is not the case. Researchers focusing on students' self-efficacy for academic learning (Bandura, 1986, 1997), likewise, have reported on the problem of underestimation and overestimation of judgments of perceived competence (Pajares and Kranzler, 1995; Pajares, 1996a,b). This problem of miscalibration of competence beliefs (e.g., underestimation), we contend, may arise from a student's lack of focus, lack of concentration, and misunderstanding of instruction.

Aside from instructing a person to equate his/her level of an internal state of functioning with a corresponding numerical value, it is also possible to use Likert-scale measures and/or open-ended surveys. In the broad area of *subjective well-being*, for example, a number of researchers have developed different Likert-scale measures, such as the Comprehensive Inventory of Thriving (CIT) Scale (Su et al., 2014), the Flourishing Scale (Diener et al., 2010), and the Academic Striving Subscale (Phan et al., 2018a,b,c). The use of Likert-scale measures, administered to subjects on multiple occasions, is straightforward and may provide fruitful information about their current state of functioning and the potential of achieving optimal functioning. For example, consider a participant's response to the Flourishing Scale (Diener et al., 2010) on two occasions, denoted as: Response-FT_1 to Response-FT_2. A positive change in scores from Response-FT_1 and Response-FT_2 ($\Delta_{(\text{Response-FT2 – Response-FT1})}$ = +ve), in this analysis, would indicate an improvement in personal flourishing from T_1 to T_2. A negative difference ($\Delta_{(\text{Response-FT2 – Response-FT1})}$ = −ve), in contrast, would suggest a decline in a person's state of flourishing. It is possible, too, for us to explore and identify linear and/or nonlinear trajectories of a person's subject well-being. The use of latent growth modeling (LGM) procedures, in particular, may also enable researchers to test for effects of extraneous influences on growth trajectories (Bollen and Curran, 2006; Hancock and Lawrence, 2006).

More recently, in an attempt to study the process of optimization (Fraillon, 2004; Phan et al., 2017), Phan et al. (2016) developed a Likert-scale questionnaire to measure and assess current level and optimal level of subjective academic well-being. The *Realistic Achievement Best Subscale* (e.g., "I am content with what I have accomplished so far for this subject"), according to the authors, explores a person's actual functioning, whereas the *Optimal Achievement Best Subscale* (e.g., "I can achieve much more in this subject than I have indicated through my work so far") reflects the person's notional best functioning. The Optimal Outcome Questionnaire, as Phan et al. (2016) proposed, may serve as a diagnostic tool to assess students' "profiles" of cognitive competence in their academic learning (Phan et al., 2018a,b,c). Furthermore, in their detailed theorization of optimization, Phan et al. (2016) postulate the forming of two subscale scores [i.e., the Realistic Achievement Best (RAB) Subscale and the Optimal Achievement Best (OAB) Subscale scores] that would assist in the assessment, measurement, and evaluation of the process of optimization. What is unclear though, from this consideration, is how we could use the RAB and OAB scores to measure and assess the operational nature of optimization.

From an educational perspective then, measuring and assessing a current level of cognitive functioning and an optimal level of cognitive functioning may involve the use of comparable quantitative methodologies, such as Likert-scale measures and cognitive competence tests (e.g., quiz). A robust

methodological approach, in this case, may consist of an integration of three comparable measures: Likert-scale measures, the Optimal Outcome Questionnaire (Phan et al., 2016), and standardized testing (Phan et al., 2019a,b). This methodological conceptualization is depicted as follows:



Note: the Likert-scale measure X = Comprehensive Inventory of Thriving (CIT) Scale (Su et al., 2014), the Flourishing Scale (Diener et al., 2010), the Academic Striving Subscale (Phan et al., 2018a,b,c), etc.

The above methodological conceptualization is insightful as it enables us to cross-validate the three comparable measures. The Optimal Outcome Questionnaire is administered to participants at a particular time point (Phan et al., 2016, 2017), and measures and assesses a person's current level of functioning (i.e., the RAB Subscale) and his/her optimal level of functioning (i.e., the OAB Subscale). Of concern, from our viewpoint, is whether the OAB score actually indicates a person's optimal best, or whether it is simply an indication of "miscalibrated" and potential optimal best (Phan et al., 2018a,b,c). On this basis, it would be appropriate to cross-validate the Optimal Outcome Questionnaire with another comparable Likert-scale measure (e.g., coined as "X"), administered to a participant on multiple occasions. Longitudinal research designs emphasize the importance of *time precedence*, stipulating the administration of the same Likert-scale measure on multiple occasions (Rogosa, 1979; MacCallum and Austin, 2000) – for example, a current time point, T_1, and a future time point, T_2. Hence, considering the Optimal Outcome Questionnaire and Likert-scale measure X, we propose a first iteration (*Iteration 1*) where there are two associations: (1) $r_{XT1\text{-}RAB}$, which depicts the association between a Likert-scale measure X administered at T_1 and the RAB Subscale, and (2) $r_{XT2\text{-}OAB}$, which depicts the association between the same Likert-scale measure X administered at T_2 and the OAB Subscale. In this analysis then, in terms of consistency and accuracy, we would expect similar rating scores for the RAB Subscale and the Likert-scale measure X at T_1, and for the OAB Subscale and the same Likert-scale measure X at T_2, respectively.

A scrutiny of the possibility of miscalibration is important (e.g., Pajares and Miller, 1994; Pajares and Kranzler, 1995; Pajares, 1996a,b) because miscalibration may result in either inflated (e.g., overconfidence of optimal level of cognitive functioning) or deflated (e.g., underconfidence of current level of cognitive functioning) responses. In a similar vein, the use of an identical Likert-scale measure on multiple occasions also poses problems such as identification of familiarity of items, and evidence of autocorrelated errors between items – for example, Item 1 at T_1 and Item 1 at T_2, Item 2 at T_1 and Item 2 at T_2, etc. (Bandalos et al., 1995; Marsh and Yeung, 1997; Guay et al., 1999). Addressing these potential problems, we propose a second iteration (*Iteration 2*), namely: (1) $r_{XT1\text{-}STT1}$, which depicts the association between the Likert-scale measure X administered at T_1 and a standardized performance test, denoted as STT1, and (2) $r_{XT2\text{-}STT2}$, which depicts the association between the same Likert-scale measure X administered at T_2 and a different standardized performance test, denoted as STT2. Again, in terms of consistency and accuracy, we would expect a similar rating score for the Likert-scale measure X at T_1 and the performance score of the standardized test at T_1, and for the same Likert-scale measure X at T_2 and the performance score of the standardized test at T_2.

Finally, in order to validate the nature of the Optimal Outcome Questionnaire (Phan et al., 2016) and taking into account the possible shortcomings of Likert-scale measures, we consider a third iteration (*Iteration 3*), which emphasizes the potential associations between the Optimal Outcome Questionnaire and standardized testing at T_1 and T_2: (1) $r_{RAB\text{-}STT1}$, which depicts the association between the RAB Subscale and a standardized test administered at T_1, and (2) $r_{OAB\text{-}STT2}$, which depicts the association between the OAB Subscale and a standardized test administered at T_2. Similar to the two previous iterations, in terms of consistency and accuracy, we would expect a similar rating score for the RAB Subscale and the performance score of the standardized test at T_1, and the OAB Subscale and the performance score of the standardized test at T_2.

We contend that the three iterations outlined, in their totality, make substantive contributions to the study of measurement and assessment of optimal functioning. From our rationalization, $r_{XT1\text{-}RAB}$, $r_{RAB\text{-}STT1}$, and $r_{XT1\text{-}STT1}$ would provide theoretical understanding of a current level of functioning [i.e., $X_{(T1)} \approx RAB \approx ST_{(T1)}$], whereas $r_{XT2\text{-}OAB}$, $r_{OAB\text{-}STT2}$, and $r_{XT2\text{-}STT2}$ would provide theoretical understanding of an optimal level of functioning [i.e., $X_{(T2)} \approx OAB \approx ST_{(T2)}$]. The use of any of the three measures alone is somewhat limited, whereas a combination of two or all three measures is more stringent in terms of elucidating the complex nature of optimal functioning. One notable inquiry that has, to date, remained elusive is our theoretical inference and interpretation of Δ [i.e., $\Delta_{(XT2\,-\,XT1)}$, $\Delta_{(OAB\text{-}RAB)}$, and $\Delta_{(STT2\text{-}STT1)}$]. For example, given a participant's response to a Likert-scale measure X at T_1 and T_2, the Optimal Outcome Questionnaire at T_1, and a cognitive competence test at T_1 and T_2, can we use this information to explain the process of optimization?

The Quantification of Optimization: A Proposed Index of Optimization?

The preceding discussion pertaining to the assessment and measurement of optimal functioning is insightful for the purpose of our proposition: *the potential "quantitative" measure of the process of optimization*. Referring to our previous mentioning of the three comparable iterations of optimal functioning, we have $X_{(T1)}$, RAB, and $ST_{(T1)}$ as indicators of a current level of functioning, and $X_{(T2)}$, OAB, and $ST_{(T2)}$ as indicators of an optimal level of functioning. Of interest, in this analysis, is whether and/or to what extent the derivative of Δ notation (e.g., $\Delta_{(XT2\,-\,XT1)}$) could align with, and/or fit in with out proposed conceptualization (i.e., activation and enactment of an agent → the sub-process of energization → stimulation of buoyancy of psychological attributes; **Figure 1**). Empirical validation of optimization (O_B), as an underlying process, *does not* equate

to a "predictive effect," an "enhancement," and/or a "causal flow." In other words, at this stage, methodologically and statistically, very little is known about the "quantitative representation" of the process of optimization. A predictive effect, denoted as a beta value (i.e., β), may simply inform us of a positive association between an educational, psychological, or psychosocial agent and an adaptive outcome. For example, in a recent longitudinal, nonexperimental study, Phan et al. (2018a,b,c) found that effective functioning exerted a positive effect on school experience ($\beta = 0.62$, $p < 0.001$) and academic achievement ($\beta = 0.30$, $p < 0.001$). Likewise, in an earlier research, McCartney et al. (2007) reported the positive effect of high quality child care, as an index of a psychosocial agent, on school readiness ($\beta = 0.21$, $p < 0.01$).

A complex issue then relates to the "transformation" of an r value (i.e., an association) or a β value (i.e., a predictive effect) into an "optimizing effect." What is an "optimizing" effect, and how do we define and/or calculate this optimizing effect? We postulate that an optimizing effect, denoted as "γ," is derived from three "pathways," as shown in **Figure 1**: (1) Path A describes the result of the activation and enactment of psychological, educational, and psychosocial agents, which then results in the process of energization (i.e., AE → E), (2) Path B describes the result of energization, which consequently leads to the stimulation of buoyancy of different psychological attributes (i.e., E → SB), and (3) Path C describes the arousal of an internal state of functioning and its sustained positioning from T_1 to T_2 (i.e., SB → AS).

Having identified these specific paths, we need to conceptualize the "intensity" of optimization by assigning a numerical value to each effect (e.g., 0 for minimal optimizing effect to 1 for maximal optimizing effect). The quantification of γ, in this instance, would reflect the totality of effects (i.e., the combined effects of Path A, Path B, and Path C). In other words, as a point of summary: γ = Path A + Path B + Path C. An important question for us to consider then, is why would γ vary in its magnitude? Referring back to our conceptualization of optimization, one notable aspect is the difference between L_1T_1 and L_2T_2. The $\Delta_{(L_2T_2-L_1T_1)}$, we argue, is likely to vary in accordance with a person's current level of functioning (L_1T_1) *and* his/her subsequent level of optimal functioning (L_2T_2). For example, consider mathematics learning for the topic of *Algebraic expressions* with two different scenarios:

Scenario 1.

L_1T_1 = knowing how to solve equations with one unknown, x: $x + 8 = 10$, evaluate x?

L_2T_2 = knowing how to solve quadratic equations with one unknown, x: $(x - 10)^2 = 20$, evaluate x?

Scenario 2.

L_1T_1 = knowing how to solve equations with one unknown, x: $x + 8 = 10$, evaluate x?

L_2T_2 = knowing how to solve simultaneous equations with two unknowns, x and y: $(2x + y) = 9$ and $(5x - 10y) = 20$, solve for x and y.

An analysis of the two mentioned scenarios indicates that L_2 cognitive functioning is more complex for Scenario 2 (i.e., simultaneous equations that have two unknowns) than that for Scenario 1 (i.e., equations that have one unknown), suggesting that $\Delta_{(L_2T_2-L_1T_1)}$ (i.e., the *range*) is "larger" in scale or amount for the former. Achieving L_2 (i.e., an optimal level) from L_1 for Scenario 2 requires "more" effort in terms of optimization. This example, interestingly, emphasizes the potential interrelations between the magnitude (i.e., intensity or strength) of the process of optimization and the range or difference between L_1T_1 and L_2T_2. On this basis, the magnitude of the quantification of γ (i.e., reflecting the totality of the process of optimization) is postulated to associate with the "complexity" of L_2, and how this optimal level of functioning differs from L_1. In formulating a quantitative derivative of this consideration, we recently proposed a theoretical concept, which we coined as the "index of optimization" (i.e., denoted as IO) (Phan et al., 2019a,b). The IO is defined as: $\Delta_{(L_2T_2-L_1T_1)} \times \gamma$, where γ = Path A + Path B + Path C.

The index of optimization is the *combination* (i.e., multiplication) of the difference between L_1T_1 and L_2T_2 and the magnitude of the optimizing effect of an educational, psychological, or psychosocial agent. How does the IO help us in our understanding of optimal functioning and optimization? A quantified numerical value of IO, which we propose to range from 0 (e.g., minimal IO) to 1, may elucidate the complexity of $\Delta_{(L_2T_2-L_1T_1)}$, and the amount of resources that would be needed for optimization to achieve L_2 (e.g., ability to solve simultaneous equations with two unknowns)? Importantly, the quantification of IO (close to 1) may also reveal a person's energy level. A high value of IO, for example, would indicate a person is completely energized, and that the stimulation of buoyancy of different psychological attributes is more likely. A low value of IO (close to 0), in contrast, would indicate a low level of vitality and liveliness.

How do we standardize the measurement and assessment of the IO? Aside from the complexity of $\Delta_{(L_2T_2-L_1T_1)}$, it is important to highlight that the "combination" in effects of Path A (i.e., AE → E), Path B (i.e., E → SB), and Path C (i.e., SB → AS) in the process of optimization is not easily measured and/or computed. Consider the personal experience of energization, which arises from the activation and enactment of educational, psychological, and psychosocial agents. Measurement and assessment of the sub-process of energization, along with the delving into the subsequent arousal and sustaining of an internal state of functioning is a difficult feat to ascertain. It would be of interest for future research to focus on the development of appropriate methodological designs and measurements that could validate and standardize the proposed IO. For example, the level of optimization to assist a person's optimal level of emotional functioning (e.g., a positive state of happiness) would differ from that of the level of optimization to facilitate optimal physical functioning (e.g., being able to score 50 goals in one football season). A γ value of "0.4" for the achievement of optimal cognitive functioning would not, in our view, equate to the same γ value of 0.4 for the achievement of optimal physical functioning. In other words, from this

comparison, we contend that the index of optimization would vary in accordance with a particular type of functioning (e.g., cognitive functioning *versus* physical functioning).

DIFFERENTIAL INFLUENCES OF HUMAN AGENCIES: AN EXAMPLE OF OPTIMAL COGNITIVE FUNCTIONING

One notable component of our conceptualization of optimization that is worthy for discussion is the activation and enactment of different agencies to serve as sources of energy. We argue that the differential influences of psychological, educational, and psychosocial agencies are subject to the contextual situation at hand, as well as the timely opportunity that may arise. For example, the optimization of physical functioning (e.g., a football player's scoring of goals) may benefit more from psychological (e.g., the use of self-efficacy beliefs to convince the football player's resolve) and/or psychosocial (e.g., the provision of an adequate environment for training) agencies. However, educational agencies (e.g., the teaching of an effective instructional design) could be more appropriate in the optimization of cognitive functioning (e.g., a student's academic performance in mathematics). In a similar vein, we argue that on a daily basis, the provision of opportunities for optimization purposes may vary in accordance with the contextual situation. What this means is that at any point in time, there are variations in the exposure of psychological, educational, and psychosocial agencies.

Personal energy, we postulate, differentially influence intrinsic motivation, personal resolve, effective functioning, mental strength, and effort expenditure. The stimulation of buoyancy of the five personal attributes that serve to arouse and sustain a person's progress is likely to vary in accordance with the contextual subject matter. For example, in the context of academic learning, a student may show personal resolve as he or she seeks achievement of optimal best (e.g., achieving mastery of a particular concept). Likewise, an academic subject matter that is of interest and has authentic relevance may energize a student's intrinsic motivation. In a nonacademic sense, in contrast, an athletic may exhibit a high level of mental strength as he makes attempts to achieve optimal best in long-distance running. The impact of a psychosocial agency (e.g., the provision of emotional and social support) may, in contrast, serve to energize the person's effort expenditure as she seeks to adjust to a new social environment.

Hence, from our conceptualization, the process of optimization is dynamic in terms of the availability of different agencies. The dynamic of the process of optimization is postulated to intricately link with the contextual matter or situation, at hand – for example, a senior citizen's seeking to achieve optimal health after surgery, or a student's fulfillment of mastery competence in Calculus. The contextual matter or situation, from our point of view, then corresponds with a related agency for the personal experience of energization. This consideration places an emphasis on different "pathways" of optimization: (1) psychological agency (e.g., the impact of personal self-efficacy: Bandura, 1997) → energization → stimulation of buoyancy of intrinsic motivation, or (2) educational agency (e.g., an appropriate instructional design: Ngu and Yeung, 2013) → energization → stimulation of buoyancy of effort expenditure, or (3) psychosocial agency (e.g., the impact of teacher-student relationship: Roorda et al., 2011) → energization → stimulation of buoyancy of mental strength.

For this final section of the article, we discuss the comparable influences of psychological, educational, and psychosocial agencies on the optimization of cognitive functioning. From previous research development, we consider the importance of *personal self-efficacy* (Rosenberg, 1965; Bandura, 1997; Trautwein et al., 2006), effective *instructional designs* (Ngu et al., 2014; Star et al., 2015), and *social relationships at school* (Cornelius-White, 2007; Roorda et al., 2011) as psychological, educational, and psychosocial agencies, respectively. Optimal cognitive functioning, within the contexts of schooling, may consist of a student's academic performance in a subject area, his or her willingness to show mastery competence in a topical theme, or successful school adjustment.

An Example of Psychological Agency: The Impact of Personal Self-Efficacy

Personal self-efficacy (Bandura, 1997), which forms part of *the self-beliefs system*, is a notable construct that serves as a strong predictor of educational and noneducational outcomes. Personal self-efficacy, according to Bandura (1997), refers to "beliefs in one's capabilities to organize and execute the course of action required to produce given attainments" (p. 3). This definition contends that self-efficacy is not concerned with a person's actual capability, but rather his/her self-judgment of perceived competence (e.g., regardless of my current ability, do I believe that I have the capability to complete this mathematics task?) Self-efficacy is a potent predictor of different types of adaptive outcomes (e.g., academic performance), as it mobilizes a person's state of persistence and effort expenditure, governs his or her choices in life, and regulates appropriate emotional responses. In accordance with Bandura's (1997) theory, a high level of academic self-efficacy is likely to assist a student to choose an appropriate course of action (e.g., choosing a mathematics-related career pathway: Betz and Hackett, 1983, 1986).

We contend that personal self-efficacy is analogously related to the paradigm of positive psychology (Seligman and Csikszentmihalyi, 2000; Seligman, 2010). As existing research has shown, a heightened state of self-efficacy is associated with improvement in corresponding outcomes (Schunk, 1995; Pajares, 1996a,b; Bandura, 1997). A weakened state of self-efficacy, in contrast, is more likely to result in engagement of maladaptive functioning (e.g., orientation toward performance-avoidance goals: Liem et al., 2008). From the perspective of schooling, in terms of optimization of enjoyable learning experiences, we could use academic self-efficacy as a source of energy to stimulate the buoyancy of intrinsic motivation, personal resolve, effective functioning, mental strength, and/or effort expenditure. To our knowledge, to date, no research has yet considered the conceptualization of academic self-efficacy as an operator of a person's energy that manifests in his or her stamina and liveliness.

How does academic self-efficacy instill a level of stamina and liveliness in the teaching and learning processes? Our conceptualization, in this case, considers the "potency" of academic self-efficacy to not only predict different types of future educational outcomes (e.g., Fast et al., 2010; Martin et al., 2010; Yailagh et al., 2013), but to also yield a corresponding level of "energy" (i.e., self-efficacy → level of energy). In this analysis, from the characteristics and nature of self-efficacy (Bandura, 1997), we propose that a high level of perceived competence would instill confidence, "feel-good" experiences, and a state of deliberate focus, all of which then transform into a source of energy, acting as an intermediary outcome to stimulate the buoyancy of intrinsic motivation, personal resolve, effective functioning, mental strength, and effort expenditure (i.e., energy → intrinsic motivation, etc.). This postulation of optimization gives a noteworthy positioning of academic self-efficacy (Bandura, 1997) as a source of energy for further prediction.

The proposition regarding a person's experience of energy, which results from a heightened level of self-efficacy, is an interesting tenet and requires further consideration and development. The main emphasis, in this case, is the saliency of an "interjection" of energy between self-efficacy and a corresponding criterial outcome (i.e., self-efficacy → energy → outcome). Previous correlational studies, in contrast, have attested to the interjecting role of other educational and/or psychological variables. For example, in one of their studies, Pajares and Johnson (1996) used path analysis techniques to highlight the "in-between" role of apprehension between self-efficacy and academic performance. Statistically, taking into account Baron and Kenny's (1986) writing, it is also appropriate for us to infer that energy could serve as a mediator between self-efficacy and different types of educational outcomes. In the context of optimization, we contend that energy, as an in-between variable, would mediate the effect of academic self-efficacy on intrinsic motivation, personal resolve, effective functioning, mental strength, and/or effort expenditure. An important focus of inquiry, in this case, considers the specific pathways that originate from self-efficacy to intrinsic motivation, personal resolve, effective functioning, mental strength, and effort expenditure, *via* a level of energy. We purport that the stimulation of buoyancy of the five mentioned attributes and their subsequent effects to arouse and sustain progress would vary in accordance with a student's experience, and the contextual nature of the subject matter. In other words, from this theoretical account, personal experience of energy may selectively influence some but not all of the five attributes. For example, a topical theme that is of interest is more likely to yield a student's experience of energy that gears toward intrinsic motivation and effort expenditure, whereas another student's previous experience of repeated successes in a subject matter could energize a high level of personal resolve, effective functioning, and mental strength.

An Example of an Educational Agency: An Instructional Design

Cognitive load theory (Sweller et al., 2011; Sweller, 2012), for example, has assisted the design and implementation of different instructional designs for effective mathematics learning (e.g., Ngu et al., 2016; Ngu and Phan, 2016). Situating within our explanatory account of optimization, we argue that an instructional design may optimize a student's mathematics learning experience (e.g., better comprehension and understanding of instructional materials). We consider cognitive load theory as a basis to determine to what extent an instructional design could act as a source of energy during the process of optimization. By this account, a question then is how an instructional design could cultivate positive emotions, which in turn energize a student and stimulates the buoyancy of intrinsic motivation, effective functioning, personal resolve, mental strength, and effort expenditure.

Cognitive Load Theory and Element Interactivity

Cognitive load theory (Sweller et al., 2011; Sweller, 2012) highlights the interaction between the acquisition of schemas and a person's human cognitive architecture. Basically, it focuses on the management of the limited working memory load to process complex cognitive tasks in order to facilitate acquire acquisition. It also seeks to capitalize on the unlimited capacity of the long-term memory that stores a huge number of schemas. Processing schemas retrieved from the long-term memory reduces working memory load.

Sweller (2010) argued that element interactivity is a common factor across the three types of cognitive loads (i.e., extraneous cognitive load, intrinsic cognitive load and germane cognitive load). Element interactivity refers to the interaction between elements within a learning task, which must be processed simultaneously in working memory to allow understanding to occur. An element refers to any item that requires learning (e.g., a number, a symbol, a concept, a procedure, etc.) (Chen et al., 2017). Investing cognitive resources to process interacting elements that hampers learning constitutes extraneous cognitive load, which can be reduced by altering the *design of the instruction*. Investing cognitive resources to process element interactivity that arises from the inherent complexity of material constitutes intrinsic cognitive load. There is an inverse relation between the amount of intrinsic cognitive load and learners' expertise in a domain. The intrinsic cognitive load of the material is fixed with a given level of the learner's expertise in the domain. Investing cognitive resources to process element interactivity of the material that contributes toward learning constitutes germane cognitive load. The germane cognitive load depends on the intrinsic cognitive load because the level of element interactivity that determines germane cognitive load is associated with the intrinsic cognitive load of the material.

Instructional Design, Cognitive Load, and Emotion

Research has indicated that negative emotions (e.g., anxiety) increase cognitive load imposition and decrease working memory capacity for processing information, resulting in reduced learning (Fraser et al., 2014). However, less is known about the relation between cognitive load imposition, positive emotions, and learning outcomes (Fraser et al., 2012). It is possible an effective instruction that imposes low cognitive load would cultivate positive emotions, which in turn increase a student's energy level. Based on cognitive load theory, we propose the benefit

of acquiring a higher level schema by building on a lower level of schema in learning linear equations.

Element Interactivity and Instructional Design

The concept of element interactivity (Sweller, 2010) may provide information that could help us understand the relation between the varying levels of schemas (e.g., lower level schema *versus* higher level schema). According to Sweller (2010), element interactivity acts as an index of complexity of learning material – in other words, the extent to which elements within the learning material interact determines the level of element interactivity. Estimation of the level of element interactivity is made by noting the number of elements involved, as well as assessing the interaction between the elements. Interestingly, in terms of instructional designs, a level of element interactivity accounts for the efficiency of a particular design – for example, a high level of element interactivity imposes high cognitive load and, likewise, vice versa.

Researchers (Blayney et al., 2009; Ngu et al., 2014) have advocated sequencing complex materials to allow the building of a higher level schema upon a lower level schema (i.e., prior knowledge). In relation to linear equations, capitalizing prior knowledge of one-step equations (**Figure 3A**) in order to learn two-step equations (**Figure 3B**) would help ease the burden of the working memory. The learning of the two-step equations can occur in two stages. In the first stage, we can instruct the learner to review the solution procedure of the one-step equation (i.e., Lines 1, 2, and 3, **Figure 3A**). In the second stage, the learner will learn the solution procedure of the two-step equation (i.e., Lines 1, 2, 3, 4 and 5, **Figure 3B**). The recall of prior knowledge, in this case, ensures that the learner is able to identify that $3x = 12$ (i.e., one-step equation) is similar to that of $4x = 8$ (i.e., Line 3, two-step equation) in terms of problem structure, and therefore both share the same solution procedure. Accordingly, the learning of the two-step equation becomes the learning of Lines 1 and 2 only, thus alleviating working memory load. From this understanding, the acquisition of a higher level schema (i.e., two-step equation) is built upon a lower level schema (i.e., one-step equation), which then reduces the working memory load. Based on the same rationale, we can acquire a higher level schema of a multi-step equation (i.e., $5x - 2 = 3x + 8$) by building on the prior knowledge of a lower level schema of a two-step equation (i.e., $4x - 5 = 11$).

Our emphasis of the acquisition of a higher level of schema (i.e. complex equations) by building on a lower level of schema (i.e., simple equations) is expected to generate positive emotions, which, in our view, could serve as a source of energy for students. Nonetheless, the prior knowledge level of students may differentially stimulate the buoyancy of the five or a subset of the mentioned attributes in varying degrees of magnitude. In accordance with research in the area of *expertise reversal effect* (Kalyuga et al., 2003), low prior knowledge students need greater instructional support to strengthen their prior knowledge (e.g., one-step equations). Therefore, instructional design that highlights the capitalization of prior knowledge for learning linear equations would have greater impact on the process of optimization for high rather than low prior knowledge students.

It should be noted that popular mathematics textbooks (e.g., Vincent et al., 2012) advocate the learning of linear equations in a hierarchical order of complexity without explicitly indicating the connection between a lower level schema (e.g., one-step equations) and a higher level schema (e.g., two-step equations). This manner of learning linear equations would impose high cognitive load and cultivate negative emotions. Consequently, this would lower the student's energy level, which is likely to dampen the stimulation of buoyancy of the five mentioned attributes, leading to limited positive arousal and sustainability in the optimization progress.

An Example of Psychosocial Agency: The Importance of Social Relationships

School is a complex place that may impart conflicting, but yet important information and influences on students, teachers, and school administrators. The school social milieu, in this sense, may influence and shape students' cognitive, social, moral, and emotional development. This premise places emphasis on the "situational placement" of a student within a larger sociocultural system of change (Okagaki and Sternberg, 1993;

A

One-step equation

Line 1	$3x = 12$	(÷ 4 on both sides)
Line 2	÷ 3 ÷ 3	
Line 3	$x = 4$	

B

Two-step equation

Line 1	$4x + 2 = 10$	(– 2 on both sides)
Line 2	– 2 – 2	
Line 3	$4x = 8$	(÷ 4 on both sides)
Line 4	÷ 4 ÷ 4	
Line 5	$x = 2$	

FIGURE 3 | (A) One-step equation. **(B)** Two-step equation.

Okagaki, 2001). Okagaki's (2001) proposed *triarchic model of student achievement*, similar to that of Bronfenbrenner's (1989) *bioecological systems* theory, is interesting as it contends that improvement in cognitive development (e.g., academic performance) is not isolated, but rather depends on extraneous social and educational influences.

Our proposition, described in the preceding sections, considers the school social milieu as a possible agency of optimization. The point of contention is that different individual and/or sociocultural attributes within the social milieu, and not the social milieu itself, would act as optimizing agencies. For example, from Goodenow's (1993) research, we note that *teachers' attitudes* toward students and/or *school-based philosophical beliefs* could influence the perceptions of cultural acceptance and diversity, resulting in some students' negative experiences of school belonging. One notable facet of the school social environment, which could impart meaningful contributions to students' academic adjustment and learning experiences is that of *teacher-student relationships* (Cornelius-White, 2007; Bergeron et al., 2011; Allen et al., 2013), commonly known as TSRs (Roorda et al., 2011). What is so unique about the concept of TSR as a potential optimizing agency for change?

Roorda et al.'s (2011) theoretical review delves into an interesting tenet, namely, the explanatory account of the concept of TSR in school settings. According to the authors, there are two interesting perspectives that could account and explain the quality of TSRs: *extended attachment* and *social-motivational* perspectives (Note: consult Roorda et al., 2011 for further detail). The extended attachment perspective postulates that teachers, like caregivers, may provide a security base (e.g., emotional security) from which children feel free, and can explore the school environment and engage in different learning and extracurricular activities (Birch and Ladd, 1997; Pianta et al., 1997; Pianta, 1999). Social-motivational perspectives (e.g., self-determination theory: Deci et al., 1991; Ryan and Powelson, 1991), in contrast, contend that children become motivated when they are able to fulfill three fundamental needs: the needs for *relatedness*, for *competence*, and for *autonomy*. Teachers play a major role, according to Roorda et al. (2011), by showing "involvement (i.e., caring for and expressing interest in the student), providing structure (i.e., setting clear rules and being consequent), and supporting autonomy (i.e., giving students freedom to make their own choices and showing connections between schoolwork and students' interests)". Regardless of which theoretical perspective we align to, it is obvious that teachers play a central role in the schooling process.

Teacher-student relationships, consequently, form an important basis at school for social functioning (e.g., Ladd et al., 1999), school adjustment (e.g., Buyse et al., 2009), academic achievement (e.g., Valiente et al., 2008), and engagement in learning activities (Skinner et al., 1990). We expand on this research testament by proposing that a teacher's role at school could yield a number of meaningful outcomes, which would then transform into a source of energy to differentially stimulate the buoyancy of intrinsic motivation, personal resolve, effective functioning, mental strength, and effort expenditure. In this analysis, from our previous discussion into the operational nature of quality TSRs (Roorda et al., 2011), we consider the importance of the following: (1) a teacher's persona in-class that conveys messages of warmth, care, and nurturing, (2) a teacher's attempts to provide opportunities, pathways, and means for student growth, and (3) a teacher's willingness to facilitate a school social milieu that fosters acceptance, cultural diversity, and a sense of belonging. This development in school, similar to that of self-efficacy and instructional designs, would create a positive learning environment and a strong emotional base for students to learn.

However, the nature of stimulation is subject to different contextual and personal situations. For example, a student's favorable response to a teacher's warmth and caring nature may lead to mental strength whereas another student's response to a teacher's provision of opportunities and pathways may lead to intrinsic motivation that facilitates effective cognitive functioning and personal resolve. In contrast, a student's negative experience of school, especially in the relationship with a number of teachers, may result in a low level of mental strength that thwarts learning.

CONCLUSION

The study of optimal functioning, which emphasizes the maximization of a person's capability, requires understanding into the process of optimization. The theoretical concept of optimization has received some research interests, both theoretically and empirically. A synthesis of the literature in the areas of education, psychology, health, and subjective well-being indicates a number of comparable constructs such as cognitive flow, academic buoyancy, and personal thriving. To date, there is no satisfactory account or explanation as to what constitutes optimization. Capitalizing on recent research progress (e.g., Fraillon, 2004; Phan et al., 2017, 2019a), we develop an in-depth account of optimization for further development. We conceptualize optimization as an "underlying process" that could facilitate the achievement of optimal functioning. Optimization, we contend, is more than just a statistical prediction of a psychological variable (e.g., self-efficacy: Bandura, 1997); rather, optimization reflects the experience of "energy," which then stimulates the buoyancy of intrinsic motivation, effective functioning, personal resolve, mental strength, and effort expenditure.

An important advancement for investigation includes the development of appropriate methodological designs that could test and validate our theoretical contribution of optimization. Our proposed quantification of optimization is useful for assessing a person's level of optimal functioning, self-referenced against his/her current level of functioning. From our theorization, the index of optimization, quantified as a numerical value, helps us to address specific types of functioning that a person may develop over time (e.g., optimal cognitive functioning *versus* optimal emotional functioning).

In sum, our theoretical contribution into the study of optimal functioning has potential to facilitate specific positive outcomes, academically and nonacademically. The three major optimizing

agencies (psychological, educational, and psychosocial) are prevalent as sources of information that enable a person's experience of energization. What practitioners need to consider are specific pathways and means that could instill and sustain a state of energization to achieve optimal functioning.

Despite the aforementioned theoretical account of optimization, we do acknowledge that more progress is needed to truly understand the nature of optimization. Our proposition (e.g., the concept of "optimizing effect"), as described, is theoretical and conceptual, providing grounding for further empirical development. As Merrotsy (2017) recently noted in his book, titled *Pedagogy for creative problem solving*, there are similar theories such as *flow* (Csíkszentmihályi, 1990, 2014) that lack empirical support – "…It is interesting to note that Thomas (2011) conducted a comprehensive search, using every available database search engine, and was unable to locate any independent research on the existence of flow" (p. 163). This testament, in tandem with our own writing, suggests an important need for researchers to consider pathways and means by which we could soundly "measure, assess, and validate" optimization and, hence, the quantification of flourishing. Importantly too, from this analysis, is a focus on the positive association between optimization and academic performance in school contexts. Our previous description proposes a potential correlation between a child's experience of optimal best and his/her achievement of a cognitive test (e.g., a quiz in mathematics). It is achievable, in this case, for us to validate this relationship *via* means of factorial and/or regression techniques. What *is* of perplexity, however, is how does the totality of optimization, as detailed in **Figure 1**, explain a child's academic performance? At present, we are investigating the operational nature of energy using a quantitative, nonexperimental approach. We encourage readers, likewise, to undertake robust scientific inquiries to support our proposed theoretical-conceptual model of optimization. We are mindful of Merrotsy's (2017) caution that we do not simply accept and use a theory and/or a concept in a "passing" manner without concrete, established grounding.

AUTHOR CONTRIBUTIONS

HP was responsible for the original conceptualization and write-up of the manuscript. BN was responsible for the original conceptualization and write-up of the manuscript. AY assisted in the revision of the manuscript.

REFERENCES

ACU and Erebus International (2008). *Scoping study into approaches to student wellbeing: Literature review. Report to the Department of Education, Employment and Workplace Relations*. (Sydney, Australia: Australian Catholic University).

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., and Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: predicting student achievement with the classroom assessment scoring system-secondary. *Sch. Psychol. Rev.* 42, 76–98. ISSN: 0279-6015.

Ardell, D. (1982). *Fourteen days to a wellness lifestyle*. (San Francisco: Whatever Publishing).

Babbie, E. (2014). *The basics of social research*. (Belmont, CA: Thomson).

Bandalos, D. L., Yates, K., and Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on text anxiety. *J. Educ. Psychol.* 87, 611–623. doi: 10.1037/0022-0663.87.4.611

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. (Englewood Cliffs, NJ: Prentice-Hall).

Bandura, A. (1997). *Self-efficacy: The exercise of control*. (New York: W. H. Freeman & Co).

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

Bergeron, J., Chouinard, R., and Janosz, M. (2011). The impact of teacher-student relationships and achievement motivation on students' intentions to dropout according to socio-economic status. *US-China Educ. Rev. B* 2, 273–279. ISSN: 1548-6613.

Betz, N. E., and Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based majors. *J. Vocat. Behav.* 23, 329–345. doi: 10.1016/0001-8791(83)90046-5

Betz, N. E., and Hackett, G. (1986). Applications of self-efficacy theory to understand career choice behavior. *J. Soc. Clin. Psychol.* 4, 279–289. doi: 10.1521/jscp.1986.4.3.279

Birch, S. H., and Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *J. Sch. Psychol.* 35, 61–79. doi: 10.1016/S0022-4405(96)00029-5

Blayney, P., Kalyuga, S., and Sweller, J. (2009). Interactions between the isolated–interactive elements effect and levels of learner expertise: experimental evidence from an accountancy class. *Instr. Sci.* 38, 277–287. doi: 10.1007/s11251-009-9105-x

Bollen, K. A., and Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. (Hoboken, NJ: Wiley).

Bordens, K. S., and Abbott, B. B. (2008). *Research design and methods: A process approach*. (New York, NY: McGraw-Hill).

Bornstein, M. H., Davidson, L., Keyes, C. L. M., and Moore, K. A. (2003). *Well-being positive development across the life course*. (Lawrence, Erlbaum Associates: Mahwah).

Bronfenbrenner, U. (1989). "Ecological systems theory" in *Annals of child development: Theories of child development: Revised formulations and current issues*. ed. R. Vasta, vol. 6 (Greenwich, CT: JAI Press), 187–251.

Buyse, E., Verschueren, K., Verachtert, P., and Van Damme, J. (2009). Predicting school adjustment in early elementary school: impact of teacher-child relationship quality and relational classroom climate. *Elem. Sch. J.* 110, 119–141. doi: 10.1086/605768

Chen, O. H., Kalyuga, S., and Sweller, J. (2017). The expertise reversal effect is a variant of the more general element interactivity effect. *Educ. Psychol. Rev.* 29, 393–405. doi: 10.1007/s10648-016-9359-1

Church, M. A., Elliot, A. J., and Gable, S. L. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *J. Educ. Psychol.* 93, 43–54. doi: 10.1037/0022-0663.93.1.43

Corbin, C. L. R. (1997). *Concepts of fitness and wellness: With laboratories*. (Madison, WI: Brown and Benchmark Publishers).

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: a meta-analysis. *Rev. Educ. Res.* 77, 113–143. doi: 10.3102/003465430298563

Csíkszentmihályi, M. (1990). *Flow: The psychology of optimal experience*. (New York: Harper Perennial).

Csíkszentmihályi, M. (2014). *Flow and the foundations of positive psychology*. (Springer).

Daulta, M. S. N. (2008). Impact of home environment on the scholastic achievement of children. *J. Hum. Ecol.* 23, 75–77. doi: 10.1080/09709274.2008.11906058

Deci, E. L., Vallerand, R. J., Pelletier, L. G., and Ryan, R. M. (1991). Motivation and education–the self-determination perspective. *Educ. Psychol.* 26, 325–346. doi: 10.1080/00461520.1991.9653137

Diener, E., and Biswas-Diener, R. (2002). Will money increase subjective well-being? A literature review and guide to needed research. *Soc. Indic. Res.* 57, 119–169. doi: 10.1023/A:1014411319119

Diener, E., Wirtz, D., Biswas-Diener, R., Tov, W., Kim-Prieto, C., Choi, D.-W., et al. (2009). "New measures of well-being" in *Assessing well-being: The collected works of Ed Diener*. ed. E. Diener (Netherlands: Springer Science+Business Media B.V.), 247–266.

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., et al. (2010). New well-being measures: short scales to assess flourishing and positive and negative feelings. *Soc. Indic. Res.* 97, 143–156. doi: 10.1007/s11205-009-9493-y

Dunn, H. L. (1961). *High level wellness*. (Arlington: Beatty).

Elliot, A. J., and Murayama, K. (2008). On the measurement of achievement goals: critique, illustration, and application. *J. Educ. Psychol.* 100, 613–628. doi: 10.1037/0022-0663.100.3.613

Fast, L. A., Lewis, J. L., Bryant, M. J., Bocian, K. A., Cardullo, R. A., Rettig, M., et al. (2010). Does math self-efficacy mediate the effect of perceived classroom environment on standardized math performance? *J. Educ. Psychol.* 102, 729–740. doi: 10.1037/a0018863

Fraillon, J. (2004). *Measuring student well-being in the context of Australian schooling: Discussion paper*. (Carlton South, Victoria: The Australian Council for Research).

Fraser, K., Huffman, J., Ma, I., Sobczak, M., McIlwrick, J., Wright, B., et al. (2014). The emotional and cognitive impact of unexpected simulated patient death: a randomized controlled trial. *Chest J.* 145, 958–963. doi: 10.1378/chest.13-0987

Fraser, K., Ma, I., Teteris, E., Baxter, H., Wright, B., and McLaughlin, K. (2012). Emotion, cognitive load and learning outcomes during simulation training. *Med. Educ.* 46, 1055–1062. doi: 10.1111/j.1365-2923.2012.04355.x

Fredrickson, B. L. (2000). Cultivating positive emotions to optimize health and well-being. *Prevent. Treat.* 3, 1. doi: 10.1037/1522-3736.3.1.31a

Freund, A. M., and Baltes, P. B. (1998). Selection, optimization, and compensation as strategies of life management: correlations with subjective indicators of successful aging. *Psychol. Aging* 13, 531–543. doi: 10.1037/0882-7974.13.4.531

Goodenow, C. (1993). The psychological sense of school membership among adolescents–scale development and educational correlates. *Psychol. Sch.* 30, 79–90. doi: 10.1002/1520-6807(199301)30:1<79::AID-PITS2310300113>3.0.CO;2-X

Goodenow, C., and Grady, K. (1993). The relationship of school belonging and friends' values to academic motivation among urban adolescent students. *J. Exp. Educ.* 60, 60–71.

Gravetter, F. J., and Forzano, L.-A. B. (2009). *Research methods for the behavioral sciences*. (Belmont, CA: Wadsworth Cengage Learning).

Grice, J. W., Cohn, A., Ramsey, R. R., and Chaney, J. M. (2015). On muddled reasoning and mediation modeling. *Basic Appl. Soc. Psychol.* 37, 214–225. doi: 10.1080/01973533.2015.1049350

Guay, F., Boivin, M., and Hodges, E. V. E. (1999). Predicting change in academic achievement: a model of peer experiences and self-system processes. *J. Educ. Psychol.* 91, 105–115. doi: 10.1037/0022-0663.91.1.105

Hancock, G. R., and Lawrence, F. R. (2006). "Using latent growth models to evaluate longitudinal change" in *Structural equation modeling: A second course*. eds. G. R. Hancock, and R. O. Mueller (Greenwich, Connecticut: Information Age Publishing), 171–196.

Headey, B., and Wooden, M. (2004). The effects of wealth and income on subjective well-being and ill-being. *Econ. Rec.* 80, S24–S33. doi: 10.1111/j.1475-4932.2004.00181.x

Hettler, B. (1984). Wellness: encouraging a lifetime pursuit of excellence. *Health Values* 8, 13–17.

Huppert, F. A., and So, T. T. (2013). Flourishing across Europe: application of a new conceptual framework for defining well-being. *Soc. Indic. Res.* 110, 837–861. doi: 10.1007/s11205-011-9966-7

Kalyuga, S., Ayres, P., Chandler, P., and Sweller, J. (2003). The expertise reversal effect. *Educ. Psychol.* 38, 23–31. doi: 10.1207/S15326985EP3801_4

Keyes, C. L. (2002). The mental health continuum: from languishing to flourishing in life. *J. Health Soc. Behav.* 43, 207–222. doi: 10.2307/3090197

Keyes, C. L., Shmotkin, D., and Ryff, C. D. (2002). Optimizing well-being: the empirical encounter of two traditions. *J. Pers. Soc. Psychol.* 82, 1007–1022. doi: 10.1037/0022-3514.82.6.1007

Kline, R. B. (2015). The mediation myth. *Basic Appl. Soc. Psychol.* 37, 202–213. doi: 10.1080/01973533.2015.1049349

Ladd, G. W., Birch, S. H., and Buhs, E. S. (1999). Children's social and scholastic lives in kindergarten: related spheres of influence? *Child Dev.* 70, 1373–1400. doi: 10.1111/1467-8624.00101

Liem, G. A. D., Ginns, P., Martin, A. J., Stone, B., and Herrett, M. (2012). Personal best goals and academic and social functioning: a longitudinal perspective. *Learn. Instr.* 22, 222–230. doi: 10.1016/j.learninstruc.2011.11.003

Liem, A. D., Lau, S., and Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemp. Educ. Psychol.* 33, 486–512. doi: 10.1016/j.cedpsych.2007.08.001

MacCallum, R. C., and Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annu. Rev. Psychol.* 51, 201–226. doi: 10.1146/annurev.psych.51.1.201

Marsh, H. W., and Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: structural equation models of longitudinal data. *J. Educ. Psychol.* 89, 41–54. doi: 10.1037/0022-0663.89.1.41

Martin, A. J. (2006). Personal bests (PBs): a proposed multidimensional model and empirical analysis. *Br. J. Educ. Psychol.* 76, 803–825. doi: 10.1348/000709905X55389

Martin, A. J. (2011). Personal best (PB) approaches to academic development: implications for motivation and assessment. *Educ. Pract. Theory* 33, 93–99. doi: 10.7459/ept/33.1.07

Martin, A. J., Colmar, S. H., Davey, L. A., and Marsh, H. W. (2010). Longitudinal modelling of academic buoyancy and motivation: do the 5Cs hold up over time? *Br. J. Educ. Psychol.* 80, 473–496. doi: 10.1348/000709910X486376

McCartney, K., Dearing, E., Taylor, B. A., and Bub, K. L. (2007). Quality child care supports the achievement of low-income children: direct and indirect pathways through caregiving and the home environment. *J. Appl. Dev. Psychol.* 28, 411–426. doi: 10.1016/j.appdev.2007.06.010

Merrotsy, P. (2017). *Pedagogy for creative problem solving*. (Abingdon, United Kingdom: Taylor & Francis Ltd).

Ngu, B. H., and Phan, H. P. (2016). Comparing balance and inverse methods on learning conceptual and procedural knowledge in equation solving: a cognitive load perspective. *Pedagogiesl* 11, 63–83. doi: 10.1080/1554480X.2015.1047836

Ngu, B. H., Phan, H. P., Hong, K. S., and Hasbee, U. (2016). Reducing intrinsic cognitive load in percentage change problems: the equation approach. *Learn. Individ. Differ.* 51, 81–90. doi: 10.1016/j.lindif.2016.08.029

Ngu, B. H., and Yeung, A. S. (2013). Algebra word problem solving approaches in a chemistry context: equation worked examples versus text editing. *J. Math. Behav.* 32, 197–208. doi: 10.1016/j.jmathb.2013.02.003

Ngu, B. H., Yeung, A. S., Phan, H. P., Hong, K. S., and Usop, H. (2018). Learning to solve challenging percentage-change problems: a cross-cultural study from a cognitive load perspective. *J. Exp. Educ.* 86, 362–385. doi: 10.1080/00220973.2017.1347774

Ngu, B. H., Yeung, A. S., and Tobias, S. (2014). Cognitive load in percentage change problems: unitary, pictorial, and equation approaches to instruction. *Instr. Sci.* 42, 685–713. doi: 10.1007/s11251-014-9309-6

Noble, T., and McGrath, H. (2008). The positive educational practices framework: A tool for facilitating the work of educational psychologists in promoting well-being. *Educ. Child Psychol.* 25, 119–134.

Okagaki, L. (2001). Triarchic model of minority children's school achievement. *Educ. Psychol.* 36, 9–20. doi: 10.1207/S15326985EP3601_2

Okagaki, L., and Sternberg, R. J. (1993). Parental beliefs and children's school performance. *Child Dev.* 64, 36–56. doi: 10.2307/1131436

Pajares, F. (1996a). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemp. Educ. Psychol.* 21, 325–344.

Pajares, F. (1996b). Self-efficacy beliefs in academic settings. *Rev. Educ. Res.* 66, 543–578.

Pajares, F., and Johnson, M. J. (1996). Self-efficacy beliefs and the writing performance of entering high school students. *Psychol. Sch.* 33, 163–175.

Pajares, F., and Kranzler, J. (1995). Self-efficacy beliefs and general mental-ability in mathematical problem-solving. *Contemp. Educ. Psychol.* 20, 426–443. doi: 10.1006/ceps.1995.1029

Pajares, F., and Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem-solving: a path analysis. *J. Educ. Psychol.* 86, 193–203. doi: 10.1037/0022-0663.86.2.193

Phan, H. P., and Ngu, B. H. (2015). Validating personal well-being experiences: a quantitative approach. *Education* 136, 34–52. ISSN: 0013-1172.

Phan, H. P., Ngu, B. H., and Alrashidi, O. (2018a). Contextualised self-beliefs in totality: an integrated framework from a longitudinal perspective. *Educ. Psychol.* 38, 411–434. doi: 10.1080/01443410.2017.1356446

Phan, H. P., Ngu, B. H., Wang, H.-W., Shih, J.-H., Shi, S.-Y., and Lin, R.-Y. (2018b). "Introducing the theoretical concept of 'profiling': a cross-cultural perspective" in *Progress in education*. ed. R. V. Nata, vol. 53 (New York, NY: Nova Science), 43–74.

Phan, H. P., Ngu, B. H., Wang, H.-W., Shih, J.-H., Shi, S.-Y., and Lin, R.-Y. (2018c). Understanding levels of best practice: an empirical development. *PLoS One* 13:e0198888. doi: 10.1371/journal.pone.0198888

Phan, H. P., Ngu, B. H., Wang, H.-W., Shih, J.-H., Shi, S.-Y., and Lin, R.-Y. (2019a). Achieving optimal best practice: an inquiry into its nature and characteristics. *PLoS One* 14:e0215732. doi: 10.1371/journal.pone.0215732

Phan, H. P., Ngu, B. H., and Williams, A. (2016). Introducing the concept of optimal best: theoretical and methodological contributions. *Education* 136, 312–322. ISSN: 0013-1172.

Phan, H. P., Ngu, B. H., and Yeung, A. S. (2017). Achieving optimal best: instructional efficiency and the use of cognitive load theory in mathematical problem solving. *Educ. Psychol. Rev.* 29, 667–692. doi: 10.1007/s10648-016-9373-3

Phan, H. P., Wang, H.-W., Shih, J.-H., Shi, S.-Y., Lin, R.-Y., and Ngu, B. H. (2019b). "The importance of mindfulness in the achievement of optimal best: conceptualization for research development" in *Educational psychology – Between certitudes and uncertainties*. ed. V. Trif (Rijeka, Croatia: Intech Publishing).

Piaget, J. (1963). *The psychology of intelligence*. (Totowa, New Jersey: Littlefield Adams).

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. (Washington, DC: American Psychological Association).

Pianta, R. C., Nimetz, S. L., and Bennett, E. (1997). Mother-child relationships, teacher-child relationships, and school outcomes in preschool and kindergarten. *Early Child. Res. Q.* 12, 263–280. doi: 10.1016/S0885-2006(97)90003-X

Prilleltensky, I., and Prilleltensky, O. (2006). *Promoting well-being: Linking personal, organizational, and community change*. (Hoboken, NJ: John Wiley & Sons, Inc).

Rogosa, D. (1979) in *Causal models in longitudinal research: Rationale, formulation, and interpretation. Longitudinal research in the study of behaviour and development*. eds. J. R. Nesselroade and P. B. Balles (New York: Academic Press), 263–302.

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., and Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: a meta-analytic approach. *Rev. Educ. Res.* 81, 493–529. doi: 10.3102/0034654311421793

Rosenberg, M. (1965). *Society and the adolescent child*. (Princeton, NJ: Princeton University Press).

Ryan, R. M., and Powelson, C. L. (1991). Autonomy and relatedness as fundamental to motivation and education. *J. Exp. Educ.* 60, 49–66.

Ryff, C. D. (1995). Psychological well-being in adult life. *Curr. Dir. Psychol. Sci.* 4, 99–104.

Schunk, D. H. (1995) in *Self-efficacy and education and instruction. Self-efficacy, adaptation, and adjustment: Theory, research and application*. ed. J. E. Maddux (New York: Plenum Press), 283–301.

Seligman, M. (2010). *Flourish: Positive psychology and positive interventions. The Tanner Lectures on human values at the University of Michigan*. (Ann Arbor, Michigan: The University of Michigan).

Seligman, M. (2011). *Flourish*. (North Sydney, NSW: Random House Australia).

Seligman, M., and Csikszentmihalyi, M. (2000). Positive psychology: an introduction. *Am. Psychol.* 55, 5–14. doi: 10.1037/0003-066X.55.1.5

Skinner, E. A., Wellborn, J. G., and Connell, J. P. (1990). What it takes to do well in school and whether I've got it: a process model of perceived control and children's engagement and achievement in school. *J. Educ. Psychol.* 82, 22–32. doi: 10.1037/0022-0663.82.1.22

Snyder, C. R. (2004). Hope and depression: a light in the darkness. *J. Soc. Clin. Psychol.* 23, 347–351. doi: 10.1521/jscp.23.3.347.35458

Snyder, C. R., Ilardi, S. S., Cheavens, J., Michael, S. T., Yamhure, L., and Sympson, S. (2000). The role of hope in cognitive-behavior therapies. *Cogn. Ther. Res.* 24, 747–762. doi: 10.1023/A:1005547730153

Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K., et al. (2015). Learning from comparison in algebra. *Contemp. Educ. Psychol.* 40, 41–54. doi: 10.1016/j.cedpsych.2014.05.005

Stewart-Brown, S. (2000) in *Parenting, well-being, health and disease. Promoting children's emotional well-being*. eds. A. Buchanan and B. Hudson (New York: Oxford University Press), 28–47.

Straszewski, T., and Siegel, J. T. (2018). Positive emotion infusions: can savoring increase help-seeking intentions among people with depression? *Appl. Psychol. Health Well Being* 10, 171–190. doi: 10.1111/aphw.12122

Su, R., Tay, L., and Diener, E. (2014). The development and validation of the comprehensive inventory of thriving (CIT) and the brief inventory of thriving (BIT). *Appl. Psychol. Health Well Being* 6, 251–279. doi: 10.1111/aphw.12027

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educ. Psychol. Rev.* 22, 123–138. doi: 10.1007/s10648-010-9128-5

Sweller, J. (2012). "Human cognitive architecture: why some instructional procedures work and others do not" in *APA educational psychology handbook*. eds. K. Harris, S. Graham, and T. Urdan, Vol. 1 (Washington DC: American Psychological Association), 295–325.

Sweller, J., Ayres, P., and Kalyuga, S. (2011). *Cognitive load theory*. (New York: Springer).

Tabachnick, B. G., and Fidell, L. S. (2007). *Using multivariate statistics*. (Boston, MA: Pearson Education, Inc).

Tate, C. U. (2015). On the overuse and misuse of mediation analysis: it may be a matter of timing. *Basic Appl. Soc. Psychol.* 37, 235–246. doi: 10.1080/01973533.2015.1062380

Thomas, J. (2011). In search of the optimal learning experience: flow theory and its implications for talent development. *TalentEd* 27, 31–38.

Trafimow, D. (2015). Introduction to the special issue on mediation analyses: what if planetary scientists used mediation analysis to infer causation? *Basic Appl. Soc. Psychol.* 37, 197–201. doi: 10.1080/01973533.2015.1064290

Trautwein, U., Ludtke, O., Koller, O., and Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: how the learning environment moderates the dynamics of self-concept. *J. Pers. Soc. Psychol.* 90, 334–349. doi: 10.1037/0022-3514.90.2.334

Valiente, C., Lemery-Chalfant, K., Swanson, J., and Reiser, M. (2008). Prediction of children's academic competence from their effortful control, relationships, and classroom participation. *J. Educ. Psychol.* 100, 67–77. doi: 10.1037/0022-0663.100.1.67

Vincent, J., Price, B., Caruso, N., McNamara, A., and Tynan, D. (2012). *MathsWorld 8 Australian curriculum edition*. (South Yarra, VIC: Macmillan).

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. (Cambridge, MA: Harvard University Press).

Vygotsky, L. (1981). "The genesis of higher mental functions" in *The concept of activity in Soviet Psychology*. ed. J. V. Wertsch (Armonk, NY: Sharpe), 144–188.

Weisner, T. S. (1998). Human development, child well-being, and the cultural project of development. *New Dir. Child Dev.* 81, 69–85.

WHO. (2007). Fact sheet no. 220. Retrieved 15th December, 2014. Available at: http://www.who.int/mediacentre/factsheets/fs220/en/

Wiese, C. W., Tay, L., Su, R., and Diener, E. (2018). Measuring thriving across nations: examining the measurement equivalence of the comprehensive inventory of thriving (CIT) and the brief inventory of thriving (BIT). *Appl. Psychol. Health Well Being* 10, 127–148. doi: 10.1111/aphw.12119

Witmer, J. M., and Sweeney, T. J. (1998). "Toward wellness: the goal of counseling" in *Adlerian counseling: A practioner's approach*. ed. T. J. Sweeney (Philadelphia: Accelerated Development, Taylor & Francis Group), 43–99.

Yailagh, M. S., Birgani, S. A., Boostani, F., and Hajiyakhchali, A. (2013). The relationship of self-efficacy and achievement goals with metacognitiion in female high school students in Iran. *Procedia Soc. Behav. Sci.* 84, 11–119. doi: 10.1016/j.sbspro.2013.06.520

Ziegelmann, J. P., and Lippke, S. (2007). Use of selection, optimization, and compensation strategies in health self-regulation–interplay with resources and successful development. *J. Aging Health* 19, 500–518. doi: 10.1177/0898264307300197

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

Copyright © 2019 Phan, Ngu and Yeung. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

CONCEPTUAL ANALYSIS
published: 18 July 2019
doi: 10.3389/fpsyg.2019.01692

The Educational Situation Quality Model: A New Tool to Explain and Improve Academic Achievement and Course Satisfaction

Fernando Doménech-Betoret[1], Amparo Gómez-Artiga[2] and Laura Abellán-Roselló[1]*

[1] Developmental and Educational Psychology, Jaume I University, Castellón de la Plana, Spain, [2] Developmental and Educational Psychology, University of Valencia, Valencia, Spain

Students' academic achievement is a major concern among countries. Governments spent a lot of money on education to improve students' competences at all levels of education. Despite the enormous amount of money invested and the reforms made to curricula in many countries in recent years, these measures are not generally producing the desired results according to the data of International Performance Measurement programs for students (e.g., Program for International Student Assessment-PISA by OECD). Given the importance of this issue, this article presents an instructional-motivational model developed in the last decade to explain and improve students' learning outcomes, e.g., academic achievement and course satisfaction, entitled the "The Educational Situation Quality Model" (MOCSE, acronym in Spanish). Unlike other educational models, MOCSE offers an integrative teaching-learning approach to explain learning outcomes. By taking the educational setting as a unit of analysis, this proposal introduces a new perspective into the existing literature to predict students' achievement and course satisfaction by combining contributions from relevant psycho-educational theories, such as: "The Job Demands-Resources Model," "The Expectancy-Value Theory," and "The Achievement Goal Theory." Besides being a conceptual framework to guide research, it also provides a methodological way to improve teacher practice and learning outcomes. In this article we first briefly explain the main model's characteristics and functioning from the student perspective and, second, based on the MOCSE, we offer some keys for teachers to improve academic achievement and students' course satisfaction for a specific curricular subject. Finally, future proposals and challenges are discussed. Questionnaires are provided in the **Annex**.

Keywords: academic achievement, course satisfaction, educational model, research in the classroom, teacher training, learning outcomes

OPEN ACCESS

Edited by:
Juan Luis Castejon, University of Alicante, Spain

Reviewed by:
Miguel Angel Carbonero, University of Valladolid, Spain
Pablo Miñano, University of Alicante, Spain

****Correspondence:***
Fernando Doménech-Betoret
betoret@uji.es

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

Received: *22 May 2019*
Accepted: *04 July 2019*
Published: *18 July 2019*

Citation:
Doménech-Betoret F, Gómez-Artiga A and Abellán-Roselló L (2019) The Educational Situation Quality Model: A New Tool to Explain and Improve Academic Achievement and Course Satisfaction. Front. Psychol. 10:1692. doi: 10.3389/fpsyg.2019.01692

INTRODUCTION

Students' learning outcomes, such as academic achievement, is a major concern for teachers and governments, and one of the most important issues in the field of education and educational psychology, as proven by the large amount of research in the existing literature that focuses on this topic.

However, most of the research conducted to explain or predict learning outcomes which we can find in the literature, have two important limitations. First, no consensus theoretical model

is used as a basis. Each author centers on specific variables according to his/her theoretical tradition, which makes comparing and interpreting the results difficult. Second, many of the models used as a reference to guide research in the classroom lack a scientific and solid theoretical basis by, for instance, providing a conceptual framework to consider the teaching and the learning process independently. These limitations hinder progress from being made in this field.

"The Educational Situation Quality Model" (MOCSE, acronym in Spanish), devised by Doménech-Betoret (2006, 2013, 2018), which we present herein, attempts to overcome the aforementioned shortcomings by providing an integrated and scientific approach to explain students' learning outcomes, such as academic achievement and course satisfaction. Unlike other existing ES models in the literature that lack a solid theoretical basis, MOCSE has been configured by combining contributions from relevant psycho-educational theories, such as: "The Job Demands-Resources Model," "The Expectancy-Value Theory," and "The Achievement Goal Theory." Moreover, previous research that has focused on MOCSE has allowed the original proposal to be refined and has improved the model's predictive capacity. Finally, the model besides being a conceptual framework to guide research, it also provides a methodological way to improve teacher practice and learning outcomes.

As the studies conducted to date have been done exclusively with students (and not with teachers) from University and Secondary Education levels, this is why we present the structural configuration of the model centered on student in the current study. Although more research is needed, the results obtained to date seem to support the MOCSE model's viability with students. In the current study we explain the model's characteristics, how it works and its use to improve learning. Investigating the MOCSE model centered on the teacher is a future challenge.

THEORETICAL FRAMEWORK: THE EDUCATIONAL SITUATION QUALITY MODEL FOCUSED ON STUDENTS

What do a student being more engaged than others, learning more than others and obtaining better academic results depend on? What can I, as a teacher, do so that students engage more, learn and, consequently, obtain better results in a given subject? Answering both questions is crucial to implement actions and programs to improve learning outcomes. This article aims to shed light on the answers to both questions from a new approach provided by the Educational Situation Quality Model (Doménech-Betoret, 2006, 2013, 2018). The studies conducted to date about MOCSE have been done from only the student perspective. The data obtained seem to endorse the configuration of the model centered on the student (Doménech-Betoret, 2006; Doménech-Betoret et al., 2014, 2019; Abellán-Roselló, 2016). The structural configuration of model centered on students is displayed in **Figure 1**.

As we can see in **Figure 1**, students' perception of learning demands, and the perceptions of the supports they are provided with to overcome such demands (Stage I: Appraisal stage), predict intention to learn (Stage II: Intention activation) which, in turn, affect the level of engagement (behavioral, cognitive, relational, and affective) adopted by students during the teaching-learning process (Stage III: Teaching-Learning process) which, in turn, finally has an effect on learning outcomes, such as academic achievement and course satisfaction (Stage IV: Learning outcomes). The whole model pivots around the intention to learn, where the components from Stage I are considered antecedents or predictive variables, whereas those from Stages III and IV are considered consequences or outcome variables. As the course unfolds, students receive continuous feedback about their progress that affects their perception of the demands required and supports received. Consequently, students' perceptions are continuously updated and changing. The model operates as a system, insofar that the changes in one of its five components affect all the others.

This proposal integrates three important motivational theories to explain students' involvement (engagement) and learning outcomes: Job Demands and Resources Model (Demerouti et al., 2001; Bakker and Demerouti, 2007), Expectancy-Value theory (Eccles and Wigfield, 2002), and Achievement goal theory (Dweck and Leggett, 1988; Nicholls, 1989; Ames, 1992; Wigfield and Eccles, 2000).

Stage I: Demands and Supports for Students in the Classroom Context

The *Job Demands-Resources Model* (JD-R) (Demerouti et al., 2001) was traditionally utilized in the job context. Basically, researchers in this tradition were interested in investigating employees' work conditions in terms of demands and resources, and how these conditions could affect positively (e.g., work engagement) and negatively (e.g., stress and burnout) performance in the workplace. According to the (JD-R) (Demerouti et al., 2001), job demands are defined as "physical, psychological, social or organizational aspects of work that require physical and/or psychological effort (cognitive or emotional), and are associated with a certain physiological and/or psychological cost" (Bakker and Demerouti, 2007, p. 312). Job resources/supports refer "to the physical, psychological, social, or organizational aspects of the job that may reduce job demands and the associated physiological and psychological cost, are functional for achieving work goals, and stimulate personal growth, learning and development" (Hakanen et al., 2006, p. 497).

Applying this theory to the school context, first, requires a thorough analysis of what the specific demands required of students and teachers are and, second, which resources/supports are provided to them (students and teachers) to complete those demands. We assume, as a general principle, that students' and teachers' perceptions of the learning and teaching demands, respectively, and the resources/supports they are provided with (or expected to be provided with until the end of the course) in order to overcome those demands, play an important role in the quality of the teaching-learning process (T-L) undertaken, and determine, to a large extent, academic achievement and course satisfaction.

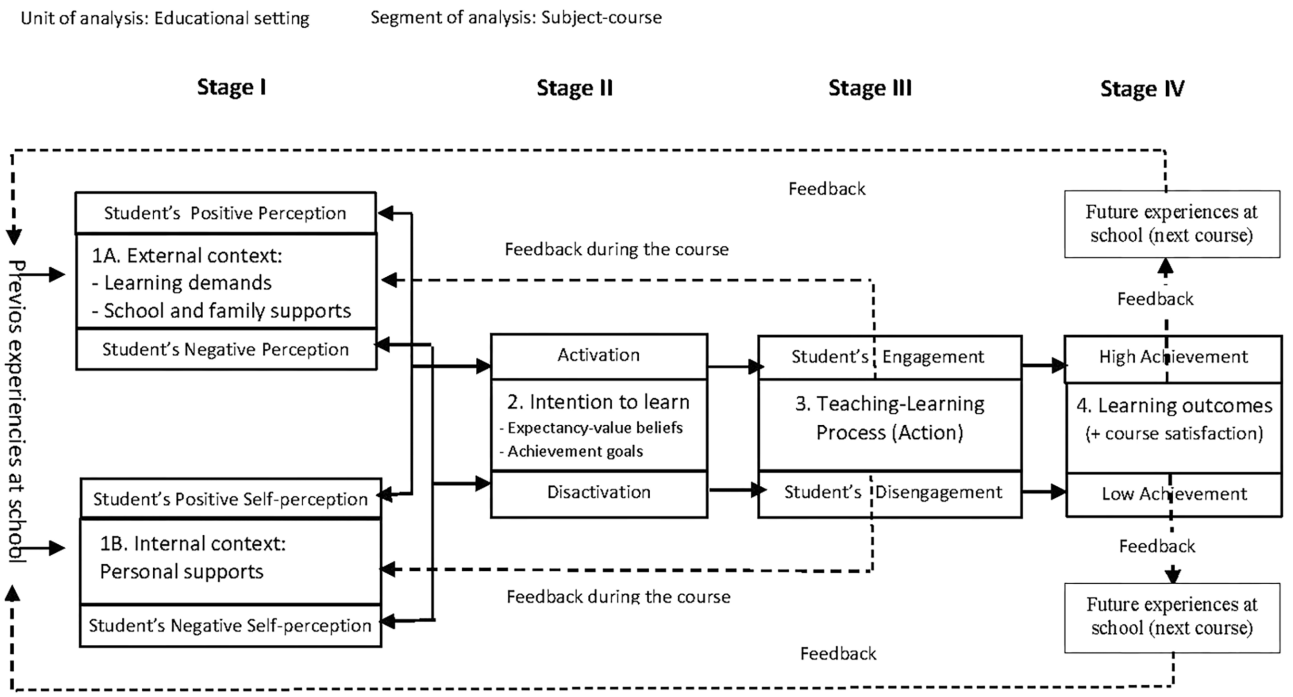


FIGURE 1 | MOCSE focused on students: structural configuration and relationship between variables.

Centered on the students' learning process, the JD-R (Demerouti et al., 2001; Bakker and Demerouti, 2007) was used to study students' perceptions of a specific educational setting in terms of learning demands and the resources/supports (external and internal) they need to overcome such demands. The selected demands or tasks that students have to meet to pass a specific subject matter (e.g., problem solving, assignments, tasks, oral presentations, lab work, etc.) are included and described in the subject's planning. They are all subordinated and aim to fulfill learning objectives (the most important demands). Students may acquire information about their required learning demands at the beginning of the course in the subject syllabus or when the teacher introduces the subject's planning.

We begin with a basic premise (Doménech-Betoret et al., 2014; Abellán-Roselló, 2016; Doménech-Betoret, 2018), the perception that each student forms of; first, the scheduled learning demands they must complete to pass a given subject; second, the support that they receive or perceive they will have mainly from the teacher and family to face these demands is crucial information to activate students' intention to learn. Consequently, they will decide to engage, or not, in learning a given subject.

Demands of Students

The learning demands in the context of a specific subject, refer basically to the tasks that students have to complete and the contents they have to study to achieve the programmed objectives and pass the subject.

Two major features can be distinguished from a learning demand: (1) the typology, related to the competence that the teacher wishes to be developed (cognitive, socio-affective, etc.); (2) the motivational characteristics of demand, which enable students' intention to learn activation. The motivational characteristics of demands (specifically, activities and tasks that students are required to complete to overcome a specific subject) should meet five major requirements to activate students' intentions to learn and encourage students to engage in the teaching learning process (Doménech-Betoret and Abellán-Roselló, 2017): (a) connect with students' living environment; (b) connect with students' interests and needs; (c) connect with what students already know; that is, with their previous knowledge (see Ausubel); (d) have a moderate level of difficulty (neither too easy, nor too difficult); in other words, the level of difficulty should be located, according to Vygotsky, in the zone of proximal development (ZPD); (e) be useful for students at a personal or a professional level; (f) make sense in the context of the subject matter. The more requirements a demand has, the more motivating it will be.

External and Internal Support Resources to Overcome Required Demands

Regarding the external resources/supports provided by teachers, "most authors have usually distinguished between affective/emotional or instructional/instrumental supports, but there is lack of consistency in the terminology used" (Doménech-Betoret, 2018). Instructional support provided by teachers aims to facilitate students' content domain which, in turn, will contribute to achieve learning demands. The affective support provided by teachers aims to meet students' psychological needs and wishes in the classroom context. It will contribute to activate positive emotions and to generate a healthy classroom climate. Students' course satisfaction is related to the emotions experienced by students during the course.

Regarding the external resources/supports provided by the family, previous studies have found that when parents or family members provide academic (e.g., assisting with homework) or affective (e.g., recognizing effort) support, students' academic achievement improves (Jelas et al., 2016). Research has also found that parents' support and involvement positively influence children's perception of their own abilities and also how they value the subjects taught at school (Rodríguez et al., 2017).

Finally, regarding internal support, students' self-beliefs related to their personal identity and self-competence (e.g., self-esteem, self-efficacy, self-concept, self-control, self-confidence, etc.) are considered important internal resources/support to predict motivation and goal setting (Doménech-Betoret, 2006; Doménech-Betoret et al., 2014, 2017; Abellán-Roselló, 2016). Students' positive self-beliefs act as personal-internal support resources that may shape their initial perception of the T-L process in terms of the demands and supports provided to achieve the planned learning objectives (Doménech-Betoret, 2018).

The teaching supports considered to date in the research carried out in the MOCSE context are listed below in **Table 1**. Bearing in mind that three fundamental levels/dimensions exist in an ES situation, namely Academic, Interpersonal and Intrapersonal, which need to be dealt with, we classified teaching supports according to the importance or effect that each specific support has on all three levels.

(a) Interpersonal level: it refers to personal relationships (teacher-students, students-students). The teacher's obligation in this area is to improve the classroom climate by managing these interpersonal relationships. Empathy and respect in dealings between teacher-students and among peers are two fundamental requirements to achieve a good classroom climate.
(b) Intrapersonal level: it refers to the relationship with oneself. Having low self-esteem, low self-concept, undervalued perception of self, etc., generates fears and insecurities (fear of failure, humiliation, etc.) in students, especially adolescents. These fears hinder learning because these students pay more attention to protect themselves than to progress and master the subject.
(c) Academic level: it is a consequence of the two previous levels. It refers to the teaching and learning of curricular content. Successful teachers cover all three levels, but some teachers only cover this level and ignore the above-mentioned two.

The decision to set a goal intention (i.e., choosing a desired end state to strive for) is commonly assumed to depend on both the desirability and the feasibility of a certain outcome (e.g., Fishbein and Ajzen, 2010). Goals are most likely to be set when the anticipated end state is subjectively evaluated as both desirable (I want X!) and feasible (I am confident that I can achieve X!). Thus, from a psychological perspective, a strong desire to attain a goal is not sufficient for the formation of a goal intention; in addition, one must be confident that the chances of attaining the goal are high.

Stage II: Intention to Learn Measured Through Expectancy-Value Beliefs and Achievement Goals

Intentionality is considered the immediate previous step of action. According to the theory of Action-Control (Heckhausen and Kuhl, 1985), intention to learn is a motivational state that is generated in the subject before initiating behavior to achieve a certain goal, and is associated with decision making. The decision to set a goal to strive for is usually assumed to depend on both

TABLE 1 | Classification of the teacher supports considered to date in the research conducted in the MOCSE context.

TEACHER SUPPORTS considered to date in the MOCSE context				
Teacher support	**(1) Acad.**	**(2) Inter.**	**(3) Intra.**	**Item example**
(1) Content comprehension support	*			"The teacher's explanations are clear and understandable".
(2) Motivational support			*	"From the beginning, the teacher made an effort to arouse our curiosity and interest in this subject".
(3) Formative evaluation (teacher feedback)	*			"The evaluation system attaches much importance to students' continued work and the teacher's feedback".
(4) Relational support		*		"The teacher comes over as being willing and open to dialogue".
(5) Competence support			*	"From the beginning, the teacher has conveyed to us the idea that we are all qualified to pass this subject if we propose to do so".
(6) Recognition support			*	"When we do things right, this teacher values it and praises us for it".
(7) Assisting students to improve achievement (study guidance)	*			"The teacher has guided us how to learn more and be better in this subject".
(8) Autonomy support			*	"The teacher gives us a chance to focus and organize the work of the topics as we wish".
(9) Providing didactic resources to study support	*			"The teacher has provided us with enough varied materials to study and work on this subject".
(10) Teacher's accessibility		*		"This teacher quickly and effectively answers the questions raised by students".

**Indicates the classroom level affected by such support. (1) Acad., academic level of the educational setting; (2) Inter., interpersonal level of the educational setting; (3) Intra., intrapersonal level of the educational setting.*

the desirability and the feasibility of reaching that goal (Fishbein and Ajzen, 2010). In other words, a certain goal is more likely to be set when individuals feel a strong desire to attain it and when they are confident that they can achieve it. The formation of a goal intention is governed by motivational principles.

According to a basic educational rule accepted by the majority of authors, successful learning requires "students' intention to learn and the teacher's intention to teach to be activated at the beginning of the educational process, and have to remain active until the process ends" (Doménech-Betoret, 2018). Intention to learn is activated on the first days of the course, and basically depend on students' perception of both the required demands and the supports provided by the teacher. However, it is assumed that this may change and fluctuate during the T-L process as a result of constant (re)appraisals made by students of the support provided to fulfill learning demands.

The Expectancy-Value Theory

The expectancy-value theory (see Eccles and Wigfield, 2002 for a modern version of this theory) is grounded in the social cognitive view of motivation. In this tradition, psychologists claim that individuals' choice, persistence and vigor invested in performance can be basically predicted and explained by their beliefs about how well they will do in the task and the value that the task has for them (Atkinson, 1957; Wigfield and Eccles, 1992; Wigfield, 1994). The three major constructs that are considered important for psychologists in this tradition are listed below:

(a) Expectancy for success (Will I succeed in this subject?). This construct is defined as "individuals' beliefs about how well they will do in upcoming tasks" Eccles and Wigfield (2002, p. 119). Expectancy for success is more future-oriented than simple self-perceptions of competence, and it refers to students' actual beliefs about their future expectancy for success. Expectancy for success usually comprises outcome expectancy and self-efficacy expectancy (Liem et al., 2008). Both terms were introduced by Bandura (1986), who differentiated between "self-efficacy or efficacy expectations" and "outcome expectancy." This author defined the former as an individual's belief in his/her own capability to accomplish a given task, while the latter is considered a person's belief that the effort he/she invests will lead to the desired outcome (Bandura, 1986).

(b) Expectancy for enjoyment (How will I feel studying this subject?). Given the importance of students' affective state for their engagement while learning, this affective component has been considered by expectancy-value theorists to be crucial (Pintrich, 1989; Pintrich and De Groot, 1990). It refers to the feelings that students expect to experience during the course, which derive from the teacher-students, content-students and peer relationships.

(c) Task/subject value (What value does this subject have for me?). Task value refers to students' beliefs about if a task or subject is worth pursuing. According to Liem et al. (2008), students' beliefs about if a task or subject is worth pursuing is a key component for understanding students' behaviors and learning outcomes. The term "value" seems a simple construct, but it is not because it has different understandings. For instance, an object can have an intrinsic, extrinsic and instrumental value. The modern expectancy-value theory (Eccles and Wigfield, 2002; Eccles, 2009) distinguishes four task-value components that we applied to a course subject to assess the subject matter value: utility, importance, interestingness and cost.

Finally, given the importance of the attributional theory in students' motivation, an additional construct was considered and added to the three aforementioned ones.

(d) Expectancy of control (To what extent does it depend on me to pass or fail this subject?). Given the importance of the attributional theory (Rotter, 1966; Weiner, 1992) in students' motivation, the construct of control was taken into account in the way of expectations, and an additional fourth construct, called "expectancy of control," was added to the three above-mentioned ones. The theorists of this tradition stress the idea that causal interpretations or attributions made by students of academic results (successes and failures) determine their motivation and efforts to a great extent. For instance, student motivation will suffer, and students will most certainly not make much effort to study a subject, if they consider that it does not depend on them (no matter how much effort they make) to pass or fail it, but on other factors beyond their control; e.g., if they get on well with the teacher, the teacher's mood when correcting exams, luck, etc. Accordingly, if the attributions that students make of their academic successes and failures are controllable, they will be more motivated to learn than if their causal attributions are uncontrollable.

The Achievement Goal Theory

The achievement goal theory (Dweck and Leggett, 1988; Nicholls, 1989; Ames, 1992; Wigfield and Eccles, 2000) argues that "the purposes that students hold for engaging in a specific learning task or in a learning process followed with a specific subject matter (i.e., their achievement goals) are an important antecedent to their achievement-related processes and outcomes" (Liem et al., 2008, p. 487). Three main goals are usually considered by researchers in this field: mastery goals, performance goals and performance-avoidance goals. The students who adopt mastery goals focus on developing one's competence to achieve a task or to pass a subject. The students who set a performance goal are concerned about others demonstrating their competence. Finally, the students who set a performance-avoidance goal wish to avoid social judgments and humiliation by others, such as the teacher or peers. For more in-depth details, see the study carried out by King and Mclnerney (2014). Previous research has found associations between the achievement goals adopted by students and outcomes variables, such as engagement/disengagement (Harackiewicz et al., 2000; Liem et al., 2008); academic achievement (Pintrich, 2000; Roebken, 2007; Diseth et al., 2012; Wei-Wen and Yi-Lee, 2015); and student satisfaction and enjoying class (Harackiewicz et al., 2002; Roebken, 2007).

In conclusion, we consider expectancy-value beliefs and achievement goals the two main dimensions to assess "intention to learn" (for more details, see Doménech-Betoret, 2018).

Prior research (Plante et al., 2013) seems to indicate that achievement goals are well explained by expectancy-value constructs, but not the other way around. That is why we have placed expectancy-value constructs as the first dimension and achievement goals as the second dimension. A high score in both dimension indicates high intention to learn, whereas a low score in both dimension indicates low intention to learn.

Stage III: The Teaching-Learning Process: Students' Engagement

Students' engagement is crucial for academic outcomes and school success, that's why engagement is one of the most important issues of educational research. A review of the literature reveals that no consensus has been reached by authors about defining this construct (Schaufeli, 2013) and, consequently, about how it should be measured. Broadly speaking, in the school context, it is generally assumed that engagement occurs when students are involved in learning tasks, and is characterized by students' continuous effort, determination and perseverance in learning (Liem et al., 2008). On the contrary, disengaged students are characterized by lack of interest, inaction and the use of avoidance strategies. Avoidance strategies are considered a negative indicator of students' engagement. Students use avoidance strategies when they give up, quit or disengage in their learning tasks related to a specific subject matter. In short, we can roughly state that *engagement* refers to involvement or participation; conversely, non-engagement, or disengagement, refers to withdrawal or apathy.

In recent years, students' engagement has been viewed as a multidimensional concept (Schaufeli, 2013). Centered on an educational setting, engagement is usually examined by considering how students behave (behavioral engagement), feel (affective or emotional engagement), think (cognitive engagement), and socialize or interact (social or relational engagement) in the classroom. Behavioral engagement is more observable and easily measurable. It usually includes actions and efforts made by students (Fredricks et al., 2004; Handelsman et al., 2005), such as, asking questions, taking an active part in class, paying attention and taking notes, participating in learning activities, etc. Cognitive engagement refers to how students feel about themselves and how effective the processing strategies or skills they use to master certain tasks are (Metallidou and Vlachou, 2007), such as, synthesizing information, highlighting the main ideas, etc. Emotional engagement has to do with the positive or negative emotions that students experience in their relations with the teacher, peers, content and school (Davis et al., 2010) such as, I feel I'm in tune with the teacher, I feel that my classmates like me, etc. Social or relational engagement contributes to create a positive and healthy classroom climate depend on the quality of interactions maintained between students and the teacher, and also between peers, during the course to a great extent. Former research works in the literature on motivation provide key notions and aspects of relational engagement, such as autonomy support (Jang et al., 2010) or school belonging (Goodenow, 1993; Ros, 2014).

Stage IV: Academic Achievement and Course Satisfaction

Learning outcomes, specifically student achievement and course satisfaction, are two of the most important indicators of a successful T-L process. "Student satisfaction is both an outcome of the learning process and a requirement for successful learning" (Sinclaire, 2014, p. 2). Accordingly, in this stage, learning outcomes and course satisfaction should be considered and evaluated. The aim of this evaluation centered on the product is, first, to know to what extent the learning objective has been achieved at the end of the course and, second, to know the level of student satisfaction reported about the followed T-L process. Student satisfaction is related to the emotions experienced by students during the course. This evaluation provides the teacher with valuable information and feedback. It allows the teacher to reflect retrospectively to introduce instructional changes for subsequent courses in order to correct failures and, thus, improve students' achievement and course satisfaction. These changes will focus mainly on those variables that are the teacher's responsibility; that is, learning demands and teacher support from components 1A and 1B.

APPLYING MOCSE TO IMPROVE LEARNING OUTCOMES

Actions Centered on the Classroom Level

When students' low achievement is detected in a specific subject matter at any level of education, the teacher is encouraged to use the MOCSE model to improve students' engagement and academic results. So, based on MOCSE postulates, we suggest following a procedure that comprises two phases. Implementing the first phase (Intervention Phase 1) is recommended at the beginning of the course, a few days after the course begins, to diagnose students' initial motivational profiles based on intention to learn indicators. To address this diagnosis, the Intention to Learn Questionnaire is provided in **Annex 1**.

The specific actions to be implemented into Phase 1 are listed below:

First action (Action 1): to assess "intention to learn" constructs (Expectancy-value beliefs and motivational goals) for a *diagnosis evaluation*. It should be carried out at the beginning of the course, some days after the course begins. Intention to learn is the cornerstone of the model on which the remaining components pivot: antecedents (Components 1A and 1B) and consequents (Components 3 and 4). Based on this structure, the first step consists in assessing intention to lean. "This action will provide teachers with valuable information about the extent to which students will engage in studying and working on a specific subject" (Doménech-Betoret, 2018).

Second action (Action 2): *analysis of the results* to detect the strengths and weaknesses of students' initial motivational profiles

based on intention to learn indicators (Expectancy-value beliefs and motivational goals).

Third action (Action 3): *reflection* on students' initial motivational profile at the beginning of the course, and on the actions than I can take as a teacher. Do I continue with the planned schedule or must I introduce changes?

If appropriate, the teacher can continue with his/her established subject planning without making any changes. If the teacher notes any major deficiencies in intention to learn (either at the class or the individual level, only in some students), then a second intervention will be necessary as soon as possible (Intervention Phase 2). This second intervention aims to assess the predictive variables (demands and supports) to detect the causes responsible for students' low motivational level. With this information, from a scientific basis, the teacher is able to initiate improvement actions by introducing the necessary changes, to avoid or reduce students' risk of failure.

The specific actions to be implemented in Phase 2 are explained below:

Fourth action (Action 4). If deficiencies are detected, a fourth action must be implemented. This consists in designing an *Action Plan* to introduce corrective measures that aim to correct the motivational deficiencies found. It will be necessary to specify in the action plan when the improvement actions will be carried out, how, and in what way its efficacy will be evaluated. Note that the aim of the action plan is to activate and increase intention to learn (the model's Component 3). To achieve this, it is necessary to assess the predictive or antecedent variables, and identify the specific variables responsible for students' intention to learn according to students' point of view; related to learning demands and supports from Components 1A and 1B. To address this assessment, the Demands and Teacher Support Questionnaire is provided in **Annex 2**. However, in order to collect more complete and detailed information, on the same dimensions, from students, it is recommendable to use the interview technique as a complementary methodology.

Subsequently, the instructional actions and programs that center on the aforementioned variables, which are the teacher's responsibility and the teacher control, should be implemented. "Besides correcting motivational deficiencies, a diagnosis evaluation also allows the teacher to adjust teacher support to students' characteristics at the beginning of the course" (Doménech-Betoret, 2018). In short, in order to activate intention to learn, appealing and meaningful demands should be planned, and affective and instructional supports (from teachers, peers, and families) should be provided."

A decline in students' motivation to learn (and as a result learning outcomes) has been found in the transition from Primary to Secondary Education, a period that coincides with the first years of adolescence, a difficult stage in a child's development. Accordingly, it is especially important to use this tool at this level of education to improve students' intention to learn and learning outcomes (academic achievement and course satisfaction). By assuming that a standard Secondary Education course comprises three trimesters and that students' progress is evaluated and reported to parents at the end of each trimester, there are basically three time points at which the students involved in the T-L process re-update their perceptions, and almost simultaneously make decisions about what their own role and involvement must be during the course. The most important time point to check students' intention to learn is the period when the course begins (after some days of class), but the time points corresponding "to the start of the second and third trimesters are also key due to the results obtained at the end of each trimester, provided on a report card" (Doménech-Betoret, 2018). So if we wish to check the evolution of students' intention to learn (and perception) throughout the course, it is advisable to make evaluations at the three aforementioned time points. Psychologist, can assist teachers to implement the aforementioned actions.

Actions Centered on the School Level

When students' low achievement is detected in a specific school, the same actions and procedure can be followed at the school level. Previously, teachers should be trained to first understand the conceptual configuration and postulates of MOCSE, and second how it can be applied in class. The "empirical data obtained with MOCSE procedures can provide the scientific basis to design effective programs for different levels and subjects" (Doménech-Betoret, 2018), to active students' intention to learn and, in turn, to learning outcomes like academic achievement and satisfaction at school. Finally, we wish to point out that teacher training is a fundamental element for teachers to implement changes and improvements into class. In the MOCSE context, teacher training would take two main directions; one, to train them so they are able to formulate stimulating learning demands that match the needs and interests of students in today's society; two, so they are capable of offering students the supports they need at all times while learning.

DISCUSSION

In accordance with the aforementioned rationale, and based on previous research, we assert that the MOCSE model explains coherently how an educational setting operates, and provides a scientific and useful framework to be used by both researchers and teachers. On the one hand, it can be used by researchers to guide their research conducted on educational settings to explain and predict academic achievement and course satisfaction from a new approach. On the other hand, it can be used by teachers as a methodological procedure to diagnose and intervene in the classroom to improve students' intention to learn and, consequently, learning outcomes.

The model can be applied at any level of education, to improve students' engagement and academic results. The data obtained to date seem to indicate the model's viability given its capacity to explain students' engagement and academic achievement in undergraduate students (Doménech-Betoret, 2006; Doménech-Betoret et al., 2019) and Secondary Education (Abellán-Roselló, 2016). However, further research is needed at different levels of education and cultural contexts to obtain more reliable findings. After providing the model's validity, we wish to use this tool

to detect regularities in similar educational contexts, that is, at the same level of education and/or in specific curricular contents or degrees. This is a future challenge, for instance, questions such as: What are the best specific teacher supports for primary, secondary or undergraduate students? What are the best specific teacher supports for secondary students in specific curricular subjects? What is the predictive role of learning demands in primary, secondary or undergraduate students? etc. These questions and others like them are still unsolved, but are important to implement efficient instructional actions and programs to improve students' engagement and academic achievement at a specific level or in a given curricular content.

AUTHOR CONTRIBUTIONS

FD-B wrote the manuscript. AG-A and LA-R reviewed the whole manuscript, checked the references, and made significant contributions. All the authors reread the manuscript and approved its submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01692/full#supplementary-material

REFERENCES

Abellán-Roselló, L. (2016). *Motivación y Aprendizaje Escolar en la Educación Secundaria Obligatoria: Un Estudio Empírico Desde el Modelo de Calidad de Situación Educativa [Scholar Learning and Motivation: Un Empirical Study from the Quality Educational Setting Model]*. Ph.D. dissertation, University of Valencia, Valencia.

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *J. Educ. Psychol.* 84, 261–271. doi: 10.1037//0022-0663.84.3.261

Atkinson, J. W. (1957). Motivacional determinants of risk taking behavior. *Psychol. Rev.* 64, 359–372. doi: 10.1037/h0043445

Bakker, A. B., and Demerouti, E. (2007). The job demands resources model: state of the art. *J. Manag. Psychol.* 22, 309–328. doi: 10.1108/02683940710733115

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Davis, H. A., Chang, M., Andrzejewski, C. E., and Poirier, R. R. (2010). Examining behavioral, relational, and cognitive engagement in smaller learning communities: a case study of reform in one suburban district. *J. Educ. Change* 11, 345–401.

Demerouti, E., Bakker, A. B., Nachreiner, F., and Schaufeli, W. B. (2001). The job demands resources model of burnout. *J. Appl. Psychol.* 86, 499–512. doi: 10.1037/0021-9010.86.3.499

Diseth, A., Danielsen, A. G., and Samdal, O. (2012). A path analysis of basic need support, self-efficacy, achievement goals, life satisfaction and academic achievement level among secondary school students. *Educ. Psychol.* 32, 335–354. doi: 10.1080/01443410.2012.657159

Doménech-Betoret, F. (2006). Testing an instructional model in a university educational setting from the student's perspective. *Learn. Instr.* 16, 450–466. doi: 10.1016/j.learninstruc.2006.09.005

Doménech-Betoret, F. (2013). An instructional model for guiding reflection and research in the classroom: the educational situation quality model. *Electron. J. Res. Educ. Psychol.* 11, 239–260.

Doménech-Betoret, F. (2018). The educational situation quality model: recent advances. *Front. Psychol.* 9:328. doi: 10.3389/fpsyg.2018.00328

Doménech-Betoret, F., and Abellán-Roselló, L. (2017). *Guía Práctica para Mejorar la Motivación de los Estudiantes de Educación Secundaria y Formación Profesional [Practical Guide to Improve Motivation in Secondary School]*. Castellón: Publicaciones de la Universitat Jaume I.

Doménech-Betoret, F., Abellán-Roselló, L., and Gómez-Artiga, A. (2017). Self-efficacy, satisfaction and academic achievement: the mediator role of students' expectancy-value beliefs. *Front. Psychol.* 8:1193. doi: 10.3389/fpsyg.2017.01193

Doménech-Betoret, F., Gómez-Artiga, A., and Lloret-Segura, S. (2014). Personal variables, motivation and avoidance learning strategies in undergraduate students. *Learn. Individ. Differ.* 35, 122–129. doi: 10.1016/j.lindif.2014.06.007

Doménech-Betoret, F., Gómez-Artiga, A., Rocabert-Beut, E., Abellán-Roselló, L., and Descals-Tomás, A. (2019). "Influencia de los apoyos docentes y familiares en la motivación de los estudiantes universitarios [Influence of Teaching and Family Supports on University Students' Motivation]," in *Proceedings of the X Internacional Congress on Psychology and Education*, 17–19 June 2020, Cordoba.

Dweck, C. S., and Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychol. Rev.* 95, 256–273. doi: 10.1037//0033-295x.95.2.256

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

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

Fishbein, M., and Ajzen, I. (2010). *Predicting and Changing Behavior: The Reasoned Action Approach*. New York, NY: Psychology Press.

Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* 74, 59–109. doi: 10.3102/00346543074001059

Goodenow, C. (1993). The psychological sense of school membership among adolescents: scale development and educational correlates. *J. Psychol. Sch.* 30, 79–90. doi: 10.1002/1520-6807(199301)30:1<79::aid-pits2310300113>3.0.co;2-x

Hakanen, J. J., Bakker, A. B., and Schaufeli, W. B. (2006). Burnout and work engagement among teachers. *J. Sch. Psychol.* 43, 495–513. doi: 10.1016/j.jsp.2005.11.001

Handelsman, M. M., Briggs, W. L., Sullivan, N., and Towler, A. (2005). A measure of college student engagement. *J. Educ. Res.* 98, 184–191. doi: 10.3200/JOER.98.3.184-192

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., and Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: predicting interest and performance over time. *J. Educ. Psychol.* 92, 316–330. doi: 10.1037/0022-0663.92.2.316

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., and Elliot, A. J. (2002). Predicting success in college: a longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *J. Educ. Psychol.* 94, 562–575. doi: 10.1037/0022-0663.94.3.562

Heckhausen, H., and Kuhl, J. (1985). "From wishes to action: the dead ends and short cuts on the long way to action," in *Goal-Directed Behaviour: The Concept of Action in Psychology*, eds M. Frese and J. Sabini (Hillsdale, NJ: Lawrence Erlbaum), 134–160.

Jang, H., Reeve, J., and Deci, E. L. (2010). Engaging students in learning activities: it is not autonomy support or structure but autonomy support and structure. *J. Educ. Psychol.* 102, 588–600. doi: 10.1037/a0019682

Jelas, Z. M., Azman, N., Zulnaidi, H., and Ahmad, N. A. (2016). Learning support and academic achievement among Malaysian adolescents: the mediating role of student engagement. *Learn. Environ. Res.* 19, 221–240. doi: 10.1007/s10984-015-9202-5

King, R. B., and Mclnerney, D. M. (2014). The work avoidance goal construct: examining its structure, antecedents and consequences. *Contemp. Educ. Psychol.* 39, 42–58. doi: 10.1016/j.cedpsych.2013.12.002

Liem, A. D., Lau, S., and Nie, Y. (2008). The role of self-efficacy, task value, and achievement goals in predicting learning strategies, task disengagement, peer relationship, and achievement outcome. *Contemp. Educ. Psychol.* 33, 486–512. doi: 10.1016/j.cedpsych.2007.08.001

Metallidou, P., and Vlachou, A. (2007). Motivational beliefs, cognitive engagement, and achievement in language and mathematics in elementary school children. *Int. J. Psychol.* 42, 2–15. doi: 10.1080/00207590500411179

Nicholls, J. (1989). *The Competitive Ethos and Democratic Education*. Cambridge, MA: Harvard University Press.

Pintrich, P. R. (1989). "The dynamic interplay of student motivation and cognition in the college classroom," in *Advances in Motivation and Achievement*, Vol. 6, eds C. Ames and M. L. Maher (Greenwich, CT: JAI Press), 117–160.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: the role of goal orientation in learning and achievement. *J. Educ. Psychol.* 92, 544–555. doi: 10.1037//0022-0663.92.3.544

Pintrich, P. R. Y., and De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom performance. *J. Educ. Psychol.* 82, 33–40.

Plante, I., O'keefe, P. A., and Theoret, M. (2013). The relation between achievement goal and expectancy-value theories in predicting achievement-related outcomes: a test of four theoretical conceptions. *Motiv. Emot.* 37, 65–78. doi: 10.1007/s11031-012-9282-9

Rodríguez, S., Piñeiro, I., Gómez, M. L., Regueiro, B., Estévez, I., and Valle, A. (2017). An explanatory model of Maths achievement: perceived parental involvement and academic motivation. *Psicothema* 29, 184–190. doi: 10.7334/psicothema2017.32

Roebken, H. (2007). The influence of goal orientation on student satisfaction, and academic achievement. *Electron. J. Res. Educ. Psychol.* 5, 679–704.

Ros, I. (2014). El sentimiento de pertenencia de los estudiantes por curso y género en una cooperativa escolar de trabajo asociado [The sense of belonging of the students per course and gender in a school worker cooperative]. *Rev. Psicol. Educ.* 9, 201–218.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monogr.* 80, 1–28. doi: 10.1037/h0092976

Schaufeli, W. B. (2013). "What is engagement?," in *Employee Engagement in Theory and Practice*, eds C. Truss, K. Alfes, R. Delbridge, A. Shantz, and E. Soane (London: Routledge).

Sinclaire, J. K. (2014). An Empirical Investigation of Student Satisfaction with College Courses. *Res. Higher Educ. J.* 22, 1–21.

Weiner, B. (1992). *Human Motivation: Metaphors, Theories, and Research*. Newbury Park, CA: Sage.

Wei-Wen, C., and Yi-Lee, W. (2015). The Relationship between goal orientation and academic achievement in Hong Kong: the role of context. *Asia Pac. Educ. Res.* 24, 169–176. doi: 10.1007/s40299-013-0169-7

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: a developmental perspective. *Educ. Psychol. Rev.* 6, 49–78.

Wigfield, A., and Eccles, J. (1992). The development of achievement task values: a theoretical analysis. *Dev. Rev.* 12, 265–310. doi: 10.1016/0273-2297(92)90011-p

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

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

Copyright © 2019 Doménech-Betoret, Gómez-Artiga and Abellán-Roselló. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 10 October 2019
doi: 10.3389/fpsyg.2019.02284

The Influence of Self-Regulation Behaviors on University Students' Intentions of Persistance

Ana Bernardo, María Esteban, Antonio Cervero, Rebeca Cerezo and Francisco Javier Herrero*

Department of Psychology, University of Oviedo, Oviedo, Spain

The implementation of the European higher education area (EHEA) is a true paradigm change in university education in which the student, with particular consideration given to autonomous work, takes the place of the teacher as the central element of the teaching-learning process. In this autonomous work, the strategies the students regularly use become particularly important, given the supposition that doing that work will lead to academic success. The objective of this study is to analyze the variables that influence students' expectations of success, measured through their intention to persist on the course they are doing. A questionnaire designed *ad hoc* was given to a sample of 1037 university students. It included aspects related to reasons for choosing the course, institutional integration, use of self-regulation strategies, and intention to drop out. Data analysis allowed the identification of satisfaction with the course chosen and appropriate study skills acquired in secondary education as predictors of expectations of academic persistance, with some differences in terms of gender. Other strategies such as class attendance or going deeply into course content did not figure. These results are at odds with the principles underlying the EHEA and show that they have not yet been interiorized by the students, who continue to perceive their studies more traditionally.

Keywords: higher education, university, university dropout, academic success, academic self-regulation

OPEN ACCESS

Edited by:
Raquel Gilar,
University of Alicante, Spain

Reviewed by:
Susana Rodríguez,
University of A Coruña, Spain
Leandro S. Almeida,
University of Minho, Portugal

****Correspondence:***
María Esteban
estebangmaria@uniovi.es;
maria_esteban_garcia@hotmail.com

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

Received: *09 July 2019*
Accepted: *24 September 2019*
Published: *10 October 2019*

Citation:
Bernardo A, Esteban M, Cervero A, Cerezo R and Herrero FJ (2019) The Influence of Self-Regulation Behaviors on University Students' Intentions of Persistance. Front. Psychol. 10:2284. doi: 10.3389/fpsyg.2019.02284

INTRODUCTION

The continuous, rapid technological, and social advances in the last fifty years have led to the new social paradigm of the "knowledge society" (Pérez et al., 2018), basing economic growth on people's intellectual capital. It seeks to improve citizen education and training, making the most of people's capacity for continuous learning, producing better qualified individuals, and so improving the number and quality of jobs available.

Universities play a fundamental role in this context, as they are the prime bodies for the production of knowledge through scientific research, transmission of knowledge through education and training, and diffusion of knowledge by different channels (Comisión de las Comunidades Europeas, 2003).

If we add to that the growing process of globalization, it is no surprise that in education at the European level there is a plan for convergence that would allow universities to join forces, and unite educational policies. This has given us the European higher education area (EHEA), with the objective of modernizing higher education teaching and institutions across Europe (Alonso-Sáez and Arandia-Loroño, 2017).

The European Higher Education Area is not only about structural and organizational change, but rather a real paradigm shift with implications in the way we understand the teaching-learning process (Esteve and Gisbert, 2011) which affects institutions at all levels: economic, methodological, social, and evaluational. However, various studies have highlighted that neither students nor teachers, nor the institutions themselves, are adequately prepared or equipped with the means to properly enact this change in educational paradigm (López et al., 2015; Alonso-Sáez and Arandia-Loroño, 2017).

One of the most significant changes is the consideration of the student as the central element of the teaching-learning process. The teacher, up to now the fundamental pillar of teaching from the more behavioral point of view, cedes ground to the student, who is established as an autonomous, self-regulated learner. So students are the protagonist, responsible for their own educational process, in line with the constructivist paradigm (Conole, 2013; Muñoz-Cantero and Mato-Vásquez, 2014; Tirado-Morueta and Aguaded-Gómez, 2014).

This autonomous character, present in the educational tenets of the EHEA, was most fully realized in the adoption of the European credit transfer system (ECTS) as the unit of measure for academic credit. ECTS credit system gives importance to classroom activities but also take into account offsite activities. So, for example, a subject with 6 ECTS credits will include in the plan 60 h of classroom work and 90 h of autonomous student work, making up the 150 actual hours of work in the subject, as generally 1 ECTS credit equals 25 hours of effective student work (Art. 4. Boletín Oficial del Estado [BOE], 2003). Boletín Oficial del Estado [BOE] (2003) stated in its explanatory preamble that this system was a conceptual reformulation of higher education curricula via the adoption of new teaching models focusing on student work. It also defined the extent to which theoretical and practical teaching would be incorporated, as well as other academic activities students were required to carry out to reach the learning objectives in each of the subjects of syllabuses (Art. 3. Boletín Oficial del Estado [BOE], 2003).

In this context, educational quality is a principal aim for European Higher Education institutions. In this sense, quality is mainly assed in terms of graduation rates, quality of instruction and excellence of research (European Association for Quality Assurance in Higher Education [ENQA], 2015). Thus, student dropout is a great problem that in Europe reach rates between 20 and 40% of university students (Vossensteyn et al., 2015). Since academic performance has showed to be the main predictor of university dropout it is important to extend the research about it (Gairín et al., 2014; Soria-Barreto and Zúñiga-Jara, 2014; Cerezo et al., 2015), particularly in the new EHEA context.

Academic performance is a fundamental variable in student progress in an institution (Casanova et al., 2018), especially in the early stages of adapting to the university system. Literature clearly shows the huge number of variables that can influence student performance, and those that may be subject to intervention have been the object of particular study, for instance psycho-educational variables such as prior training, study habits and interest or engagement in the course. In addition, the level of prior knowledge is an academic variable which is generally related to performance, especially when this knowledge is insufficient or inadequate as the basis for new learning (Soria-Barreto and Zúñiga-Jara, 2014; López et al., 2016). In fact, the influence of this variable in later academic performance in university is so great that researchers such as Miranda et al. (2013) note it as a highly influential variable and the prime institutional variable influencing students' academic failure.

Knowledge and application of appropriate study techniques have also been shown to directly influence the decision to continue with a course of study (Arriaga et al., 2011; Tuero et al., 2018), as has satisfaction with the chosen program (Bethencourt et al., 2008). Academic success requires not only a good choice of program, a good base level of knowledge and adequate study methods, it also requires regular study. Daily or periodic study is another widely researched variable related to academic performance and success (García, 2014; Bakker et al., 2015; Cerezo et al., 2016). This study engagement is easier when the student is interested in the content (Ordóñez and Rodríguez, 2015; Garrote et al., 2016) and so, indicators such as more in-depth personal study of course content contribute significantly to successfully completing subjects and programs (Carbonero et al., 2013).

Most of these variables are indicators of self-regulation of learning (de la Fuente et al., 2017). Hence, in the new European educational paradigm self-regulation of learning is encouraged in order to promote academic success and persistance (Álvarez and López, 2011). In fact, the EHEA assigns a prime role to self-regulation strategies because of their influence on the teaching-learning process and on academic results. However, it seems paradoxical that despite personal autonomy and learning-skill acquisition being part of the Spanish educational curriculum in primary (Boletín Oficial del Estado [BOE], 2014) and secondary Education (Boletín Oficial del Estado [BOE], 2015), a large proportion of students at university fail when facing the demands of self-regulation of learning (Gil-Flores, 2015; Cerezo et al., 2017; Klemenčič, 2017). This is not exclusive to Spain, it is an international problem, both in traditional and virtual environments (Broadbent and Poon, 2015; Trevors et al., 2016).

Faced with this, it is worth asking ourselves whether the cause may be found in a lack of preparation (in terms of prior knowledge or study habits) or whether it is a consequence of a discrepancy between students' perceptions of study requirements and reality, or an insufficient understanding of those requirements. In the context of the EHEA, variables that are traditionally considered to be influential in academic performance and success, such as regular class attendance, gain particular importance, as the indications teachers gather from students in those sessions are essential to orient autonomous work, as demonstrated in research by Bernardo et al. (2015); Esteban et al. (2016), and Muñoz-Cantero and Mato-Vásquez (2014).

The aim of this study is to analyze the influence of the variables outlined above on expectations of academic persistence. A better fit between prior achievement and subsequent achievement may function as a predictor of satisfaction with results and continuation with the course of study (Khattab, 2015; Velázquez and González, 2017).

To that end, the objective of this study is to examine the possible influence that study habits and personal baggage may have on students' expectations of their academic success and persistence on the institution. Specifically, we aim to see whether those variables related to the implementation of EHEA are perceived by students as precursors of satisfactory academic progress and persistence. Thus, we draw two hypothesis:

H1: There will be higher expectations of persistence, in those students who consider their prior training (in terms of prior knowledge of and mastery of study techniques) to be sufficient to the demands of the course that they are on.

H2: The students will consider those variables related to self-regulated learning behaviors important for they academic persistence.

MATERIALS AND METHODS

Sample

The sample was made up of 1037 first-year students in the University of Oviedo. The majority (73.9%) were women, and the average age was 19.94 years old (SD = 4.17). The sampling method used was non-probability intentional selection, based on the working-group teachers' access to the sample.

The students were doing various undergraduate degree courses. The most common were primary education (22.6% of the students), nursing (22.2%), infant education (16.9%), and psychology (12%). Students were doing other degree courses to a lesser extent (less than 10% of students on each course): Economics; Law; Law, Management and Business Administration (double degree[1]); English; Chemistry; Speech Therapy; Physics; Physics and mathematics (double degree[1]); and business and marketing.

Instruments

An *ad hoc* questionnaire was created for data collection in this study about university experience, self-regulation strategies applyed in higher education, dropout intentions and reasons for dropping out of university (Tuero et al., 2018). It had a Cronbach alpha of 0.79.

It was made up of eleven classification variables and many other variables grouped in eight dimensions. The classification variables refer to factors such as: identifying data, sex, age, availability of grants, branch of secondary education, final secondary education grade, university entrance exam grade, mother's educational qualifications, father's educational qualifications, whether they are doing subjects in the first course they enrolled on, whether it is their first chosen degree, whether they do any paid work and if so, their working hours, and whether they do any non-curricular activities outside class-time and if so, what type of activity and how long they spend on it (sports, academic or social activities, paid work, etc).

[1]There are several double degrees in the University of Oviedo, where pupils study two different programs at the time.

The rest of the questionnaire corresponded to 8 dimensions that contain 66 items about: (1) reason for choosing the program; (2) prior knowledge; (3) finances; (4) current situation; (5) interest in the program; (6) integration; (7) institutional variables; and (8) self-regulation strategies.

Apart from the classification variables, which were dichotomous, multiple choice or open response questions, the responses for the remaining dimensions were via a five-point Likert-type scale with the following scoring: (1) completely disagree; (2) disagree; (3) neither agree nor disagree; (4) agree; and (5) completely agree.

Procedure

The questionnaire process began initially with contact with teachers who were signed up to a university teaching innovation project, This teaching innovation project sought to analyze the motivations behind drop-out intentions and university students' self-regulation strategies.

The questionnaires were administered, on paper, in the classroom to be completed in the teachers' own classes by freshmen, 3 months after starting the course. This was to allow an evaluation before the first exams in the school year.

The procedure include written consent of participation in the study and agree with the criteria stablished by our university ethics committee.

Data Analysis

In order to examine the possible relationships between student self-regulatory behavior and expectations of academic persistance we ask students about their persistence intentions. Thus, through students' intentions to continue on the course that they started, we looked into students expectations of success. We used categorical regression techniques to evaluate the impact that the variables described previously could have on the probability that a student would stay on their current course or drop out.

Independent variables included in the analysis were categorical so we applied a categorical regression model were students expectations of persist on the program was the dependent variable and there were nine independent variables; prior knowledge, adequacy of prior acquired study techniques, interest in study, satisfaction with the chosen program, class attendance, daily study, interest in course content, performance orientation, and deepen in course contents.

Data analysis were performed using the IBM SPSS Statistics v.24 package.

RESULTS

The categorical regression model was applied first, given that it is the best fit to the mix of ordinal Likert-type variables making up the questionnaire and the dichotomized criterion variable. This model explains 22.3% of the variance in the participants' expectations of remaining on their current program.

The analysis of variance of the model, which is significant ($p < 0.005$), ensure its validity [$F(21) = 15.0713447$].

TABLE 1 | Regression coefficients for students' expectations.

	Standardized coefficients		df	*F*	Sig.
	B	Estimation of sample simulation (1000) of standard error			
I feel that what I learned previously in secondary school is sufficient to deal with this first university year	0.098	0.063	1	2.431	0.119
I think that the study techniques used up to now have been adequate	0.106	0.036	2	8.806	0.000
I am more interested in studying now than I was in secondary school	0.047	0.060	2	0.616	0.540
I feel satisfied with my choice of program	0.364	0.044	3	68.924	0.000
I have good attendance, I attend most of the classes in the university	0.093	0.067	2	1.938	0.144
I keep up to date with my subjects	−0.051	0.046	3	1.229	0.298
I am very interested in the program content	0.081	0.056	2	2.103	0.122
I try to get the best marks possible	0.059	0.039	4	2.211	0.066
I look into the topics we deal with in class in order to know more about the subject	−0.049	0.063	2	0.605	0.546

Only two variables significantly contribute to the model: the opinion that study techniques used to date are adequate, and satisfaction with the choice of program. **Table 1** shows that satisfaction is more important ($B = 0.364$, $p < 0.005$) than positive opinions about proper study techniques ($B = 0.106$, $p < 0.005$).

We found statistically significant differences in the first predictor (related to students' appropriate use of study techniques), such that those students who were thinking of dropping out tended to respond more negatively to the item (completely disagree and disagree) than those who were not thinking of dropping out. This means that using appropriate study techniques prevent students' intentions of droping out. Nonetheless, in regard to IBM (2019) the effect size of this variable is small ($\chi^2 = 30.865$; df = 4; $p < 0.000$; VCramer = 0.173).

Similarly, with the second predictor (satisfaction with choice of program), there were also significant differences. Students who were considering dropping out were less satisfied with their choices (completely disagree, disagree, and neither agree nor disagree) than those who were not thinking of dropping out (completely agree). Following the guidelines of IBM (2019) we can cathegorize the effect size as higher than for the other predictor, in the moderate effect range ($\chi 2 = 206.108$; df = 8; $p < 0.000$;VCramer = 0.446).

Following the results in the contingency tables, we carried out a correspondence analysis to visualize where the differences lay, including the gender variable. We found that while differences in the variable about study techniques were inter-gender and the differences in the satisfaction variable were intra-gender.

For the first variable (use of appropriate study techniques) we found the values shown in **Table 2**, where the variability would be almost completely explained by a single dimension, with an inertia of 0.41 out of 0.45 (91.9%).

Figure 1 shows how men who are thinking of dropping out are associated with low evaluations of their use of study techniques in contrast to women who are not thinking of dropping out, who tend to score them as adequate.

TABLE 2 | Summary of correspondence analysis: Study techniques vs. sex.

Dimension	Singular value	Inertia	Chi-squared	Sig.	Proportion of inertia
1	0.203	0.041			0.919
2	0.049	0.002			0.054
3	0.035	0.001			0.027
Total		0.045	46.523	0.000*	1.000

Df = 12, N = 1037.

With the second variable (satisfaction with the chosen degree), the results are shown in **Table 3**. As with the previous case, the variability is mostly explained by a single dimension, with an inertia of 0.21 out of 0.23 (88.1%).

As **Figure 2** shows, women who are thinking of dropping out are associated with values of completely disagree and disagree when it comes to satisfaction with their choice of course, whereas women who are not thinking of dropping out give more positive evaluations (completely agree).

CONCLUSION

The process of transition from secondary education to university is not an easy one for students, as it requires adaptation to an unknown, demanding environment regardless of what they might have been taught in prior educational and guidance processes. In this context, academic performance and expectations surrounding it are particularly interesting variables (Stinebrickner and Stinebrickner, 2014; Wolters and Hussain, 2015; Honicke and Broadbent, 2016).

In particular, in the European context, the EHEA brings along the requirement for students to develop an autonomous learning (Boletín Oficial del Estado [BOE], 2003). Thus, it is important to know whether the students understand the obligations that they need to match in order to accomplish this goal (McCardle et al., 2017).

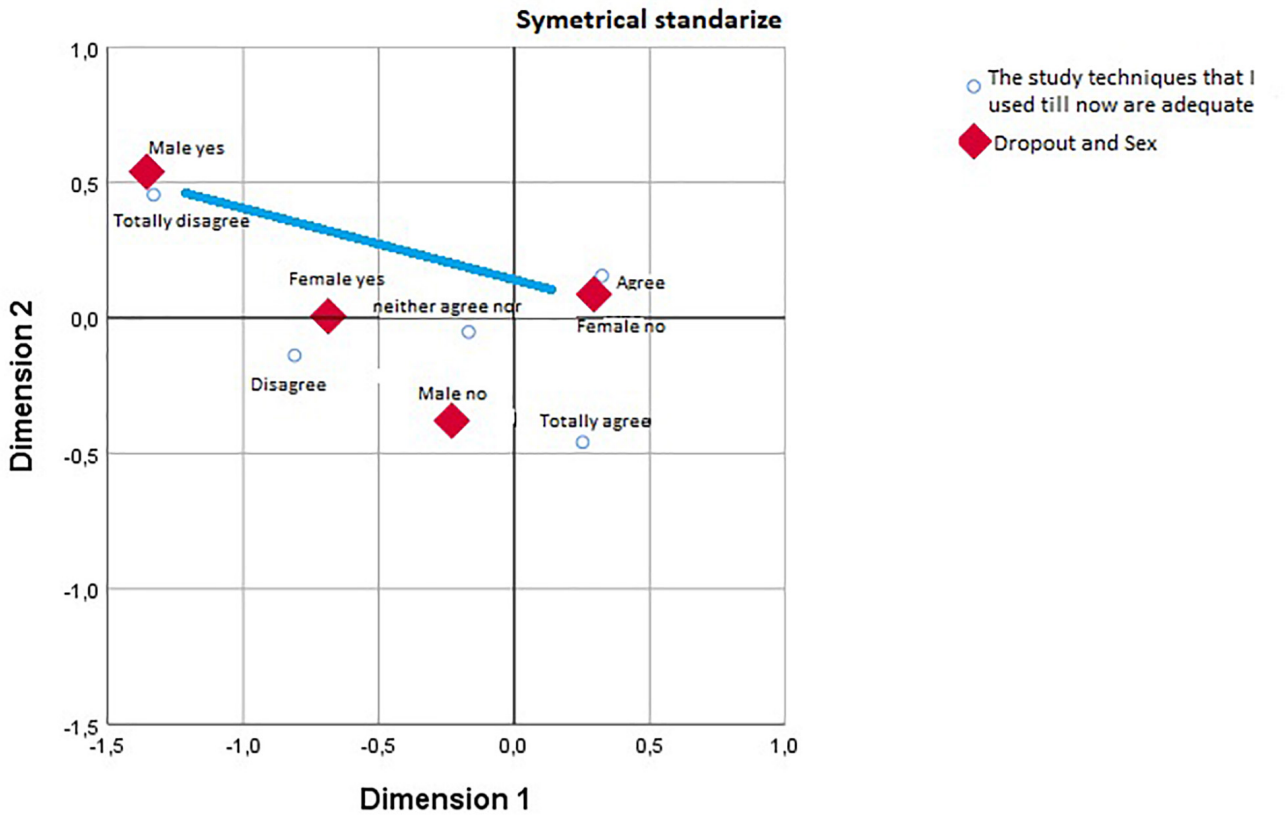


FIGURE 1 | Study techniques vs. sex.

TABLE 3 | Summary of correspondence analysis: Satisfaction with choice of program vs. sex.

Dimension	Singular value	Inertia	Chi-squared	Sig.	Proportion of inertia
1	0.458	0.041			0.881
2	0.147	0.002			0.091
3	0.082	0.001			0.028
Total		0.045	46.523	0.000*	1.000

Df = 12, N = 1037.

Therefore, in this study we analyzed those variables which, according to student expectations, influence academic persistance. Thus, we assumed that it will be those expectations which can condition their behavior for proper performance.

Our results provide evidence, in line with research in this field, of the importance students place on study techniques, an indicator which is widely related to satisfactory achievement (Navarro et al., 2015; Ng, 2018). In our case, we did not confirm the weight given to prior knowledge, in contrast to other research (Albalate et al., 2011; Roksa et al., 2017). Having prior knowledge and study techniques depends on the itinerary of prior studies (Martínez et al., 2016). Hence, the results confirm our first hypothesis, but only partially, demonstrating the need to ensure that students starting different programs do so by the appropriate selection of a specific, individualized academic itinerary (Álvarez and López, 2017; Tuero et al., 2017). This would lead to have an appropriate prior knowledge and to have acquired appropriate study techniques, which in turn would lead students to have higher expectations of persistence (Bennett, 2003).

It does seem paradoxical that student's perceptions and expectations of persistence are not related to other variables of significant learning and self-regulation, which does not support our second hypothesis. EHEA sift the educational paradigm, giving more protagonism to the student, who is supposed to be an autonomous learner. This is particularly important to bear in mind that when planning subjects, as can be seen in any teaching guide that follows the premises of EHEA, one must consider not only classroom activities such as lectures, practical classes, laboratory classes and tutorial groups, but also non-classroom activities such as individual and team work that occasionally require more time, and always the added need to learn autonomously and with self-regulation (Art. 4. Boletín Oficial del Estado [BOE], 2003).

Thus, some of the variables that we have studied – like class attendance or daily study- are important to succeed in the EHEA (Tomlinson, 2017). Despite that, our results are consistent with other studies and show how the participants

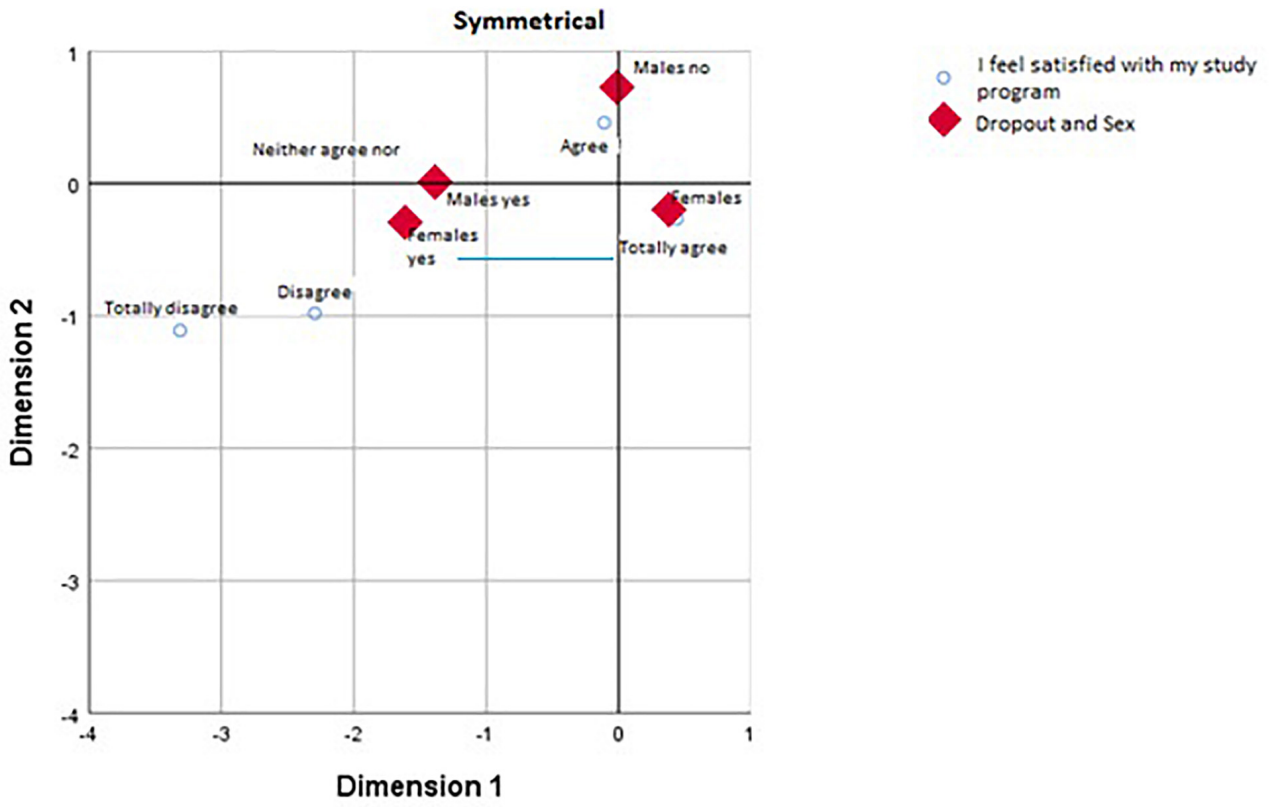


FIGURE 2 | Satisfaction with program choice vs. sex.

do not feel that these variables are important with regard to achieving satisfactory academic success. So, variables as fundamental as interest in the subject being studied (Ghasemi and Dowlatabadi, 2018), more in-depth personal study of course content (Montes, 2012; Bogarín et al., 2018), and academic engagement in terms of attendance or being up to date with work (Cerezo et al., 2017; Rissanen, 2018) are not perceived as important for success by students, in opposition to EHEA principles. Since these variables are indicators of the three dimensions of learning self-regulation of learning -motivational, behavioral, and cognitive-, we can conclude that students do not consider important to be a self-regulated learner.

These results seem to show that the postulates that gave rise to the creation and implementation of the EHEA, particularly the ECTS system of credits, have not yet been interiorized by students, who continue to perceive their study more traditionally. It is necessary to continue improving effective interventions regarding learning self-regulation; in this sense training programs in higher education such eTRAL (Cerezo et al., 2017) or Metatutor (Bouchet et al., 2016) have shown to have positive impacts on academic performance and success (Esteban et al., 2017) and can encourage better fit between students' characteristics and the requirements from EHEA based study plans.

Finally, future research should be aimed at increasing sample heterogeneity in different university years to understand whether these results apply to other programs or knowledge areas and whether there are significant differences between them. Intervention policies may be proposed to provide a better student guidance, able to guarantee a better adjustment to the context and demands of EHEA.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Secretaría General de la Universidad de Oviedo. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors contributed to the writing of the manuscript and development of the research. AB organized the sample gathering

and designed the research instrument. ME and RC helped to design the research instrument and gather the sample. AC and FH coded the data and performed the analysis.

FUNDING

This research was funded by the Spanish Ministry of Education (EDU2014-57571-P), the Spanish Ministry of Economy and Competitiveness (BES-2015-072470), and the European Union on behalf of the Principality of Asturias (FC-GRUPINIDI/2018/000199 and BP16014).

ACKNOWLEDGMENTS

We thank the support of our university, and particularly to the teachers who took part in the teaching innovation project. Without them we would not have been able to carry out this research.

REFERENCES

Albalate, D., Fageda, X., and Perdiguero, J. (2011). Éxito académico, características personales y proceso de Bolonia: una aplicación econométrica. *Revista d'Innovació Docent Universitària* 3, 11–25.

Alonso-Sáez, I., and Arandia-Loroño, M. (2017). 15 años desde la declaración de bolonia. desarrollo, situación actual y retos del espacio europeo de educación superior. *Revista iberoamericana de educación superior* 8, 199–213.

Álvarez, P., and López, D. (2011). El absentismo en la enseñanza universitaria: un obstáculo para la participación y el trabajo autónomo del alumnado. *Bordón* 63, 43–56.

Álvarez, P. R., and López, D. (2017). Estudios sobre deserción académica y medidas orientadoras de prevención en la universidad de la Laguna. *Paradigma* 38, 48–71.

Arriaga, J., Burillo, V., Carpeño, A., and Casaravilla, A. (2011). "Caracterización de los tipos de abandono. dividamos el problema y venceremos más fácilmente," in *Proceeding of the I Conferencia Latinoamericana sobre el Abandono en la Enseñanza Superior*, ed. J. Arriaga, (Madrid: Dpto. de Publicaciones de la E.U.I.T. de Telecomunicación), 75–82.

Bakker, A. B., Vergel, A. I. S., and Kuntze, J. (2015). Student engagement and performance: a weekly diary study on the role of openness. *Motiv. Emot.* 39, 49–62. doi: 10.1007/s11031-014-9422-5

Bennett, R. (2003). Determinants of undergraduate student drop out rates in a university business studies department. *J. Further High. Educ.* 27, 123–141. doi: 10.1080/030987703200065154

Bernardo, A., Cerezo, R., Rodríguez-Muñiz, L. J., Núñez, J. C., Tuero, E., and Esteban, M. (2015). Predicción del abandono universitario: variables explicativas y medidas de prevención. *Revista Fuentes* 16, 63–84. doi: 10.12795/revistafuentes.2015.i16.03

Bethencourt, J., Cabrera, L., Hernández, J. A., Álvarez, P., and González, M. (2008). Variables psicológicas y educativas en el abandono universitario. *Revista Electrónica de Investigación Psicoeducativa* 6, 603–622.

Bogarín, A., Cerezo, R., and Romero, C. (2018). Discovering learning processes using inductive miner: a case study with learning management systems (LMSs). *Psicothema* 30, 322–330. doi: 10.7334/psicothema2018.116

Boletín Oficial del Estado [BOE] (2003). *Real Decreto 1125/2003, De 5 De Septiembre, Por El Que Se Establece El Sistema Europeo De Créditos Y El Sistema De Calificaciones En Las Titulaciones Universitarias De Carácter Oficial Y Validez En Todo El Territorio Nacional*. Madird: Boletín Oficial del Estado (BOE).

Boletín Oficial del Estado [BOE] (2014). *Real Decreto 126/2014, De 28 De Febrero, Por El Que Se Establece El Currículo Básico De La Educación Primaria*. Madird: Boletín Oficial del Estado (BOE).

Boletín Oficial del Estado [BOE] (2015). *Real Decreto 1105/2014, De 26 De Diciembre, Por El Que Se Establece El Currículo Básico De La Educación Secundaria Obligatoria Y Del Bachillerato*. Madird: Boletín Oficial del Estado (BOE).

Bouchet, F., Harley, J. M., and Azevedo, R. (2016). "Can adaptive pedagogical agents' prompting strategies improve students' learning and self-regulation?," in *International Conference on Intelligent Tutoring Systems*, eds A. Micarelli, J. Stamper, and K. Panourgia, (Cham: Springer), 368–374. doi: 10.1007/978-3-319-39583-8_43

Broadbent, J., and Poon, W. L. (2015). Self-regulated learning strategies y academic achievement in online higher education learning environments: a systematic review. *Int. High. Educ.* 27, 1–13. doi: 10.1016/j.iheduc.2015.04.007

Carbonero, M. A., Román, J. M., and Ferrer, M. (2013). Programa para "aprender estratégicamente" con estudiantes universitarios: diseño y validación experimental. *Anales de Psicología* 29, 876–885. doi: 10.6018/analesps.29.3.165671

Casanova, J., Cervero, A., Núñez, J. C., Almeida, L., and Bernardo, A. (2018). Factors that determine the persistence and dropout of university students. *Psicothema* 30, 408–414. doi: 10.7334/psicothema2018.155

Cerezo, R., Bernardo, A. B., Esteban, M., Sánchez, M., and Tuero, E. (2015). programas para la promoción de la autorregulación en educación superior: un estudio de la satisfacción diferencial entre metodología presencial y virtual. *Eur. J. Educ. Psychol.* 8, 30–36. doi: 10.1016/j.ejeps.2015.10.004

Cerezo, R., Esteban, M., Sánchez-Santillán, M., and Núñez, J. C. (2017). Procrastinating behavior in computer-based learning environments to predict performance: a case study in moodle. *Front. Psychol.* 8:1403. doi: 10.3389/fpsyg.2017.01403

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., and Núñez, J. C. (2016). Students' LMS interaction patterns and their relationship with achievement: a case study in higher education. *Comput. Educ.* 96, 42–54. doi: 10.1016/j.compedu.2016.02.006

Comisión de las Comunidades Europeas (2003). *El papel de las Universidades en la Europa del Conocimiento*. Bruselas: COM.

Conole, G. (2013). *Designing for learning in an open world*. New York, NY: Springer.

de la Fuente, J., Sander, P., Martínez-Vicente, J. M., Vera, M., Garzón, A., and Fadda, S. (2017). Combined effect of levels in personal self-regulation and regulatory teaching on meta-cognitive, on meta-motivational, and on academic achievement variables in undergraduate students. *Front. psychol.* 8:232. doi: 10.3389/fpsyg.2017.00232

Esteban, M., Bernardo, A., and Rodríguez-Muñiz, L. J. (2016). Permanencia en la universidad: la importancia de un buen comienzo. *Aula Abierta* 44, 1–6. doi: 10.1016/j.aula.2015.04.001

Esteban, M., Cerezo, R., Bernardo, A., Cervero, A., and Núñez, J. C. (2017). Entornos hipermedia: posibilidades y retos. *Revista de Estudios e Investigación en Psicología y Educación* 13, 358–362. doi: 10.17979/reipe.2017.0.13.2105

Esteve, F. M., and Gisbert, M. (2011). El nuevo paradigma de aprendizaje y las nuevas tecnologías. *Revista de docencia universitaria* 9, 55–73.

European Association for Quality Assurance in Higher Education [ENQA] (2015). *Standards And Guidelines for Quality Assurance in the European Higher Education Area (ESG)*. Brussels: ENQA.

Gairín, J., Triadó, X. M., Feixas, M., Figuera, P., Aparicio, P., and Torrado, M. (2014). Student dropout rates in catalan universities: profile and motives for disengagement. *Qual. High. Educ.* 20, 165–182. doi: 10.1080/13538322.2014.925230

García, A. M. (2014). Rendimiento académico y abandono universitario: Modelos, resultados y alcances de la producción académica en la Argentina. *Revista Argentina de Educación Superior* 8, 9–38.

Garrote, D., Garrote, C., and Jiménez, S. (2016). Factores influyentes en motivación y estrategias de aprendizaje en los alumnos de grado. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación* 14, 31–44.

Ghasemi, A. A., and Dowlatabadi, H. R. (2018). Investigating the role of task value, surface/deep learning strategies, and higher order thinking in predicting self-regulation and language achievement. *J.AsiaTEFL* 15, 664–681. doi: 10.18823/asiatefl.2018.15.3.664

Gil-Flores, J. (2015). La evaluación del aprendizaje en la universidad según la experiencia de los estudiantes. *Estudios sobre educación* 22, 133–153.

Honicke, T., and Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: a systematic review. *Educ. Res. Rev.* 17, 63–84. doi: 10.1016/j.edurev.2015.11.002

IBM (2019). *V de Cramer*. Available at: https://www.ibm.com/support/knowledgecenter/en/SSEP7J_11.1.0/com.ibm.swg.ba.cognos.ug_ca_dshb.doc/cramersv.html

Khattab, N. (2015). Students' aspirations, expectations and school achievement: what really matters? *Br. Educ. Res. J.* 41, 731–748. doi: 10.1002/berj.3171

Klemenčič, M. (2017). From student engagement to student agency: conceptual considerations of European policies on student-centered learning in higher education. *High. Educ. Policy* 30, 69–85. doi: 10.1057/s41307-016-0034-4

López, M. C., Pérez-García, P., and Rodríguez, M. J. (2015). Concepciones del profesorado universitario sobre la formación en el marco del espacio europeo de educación superior. *Revista de Investigación Educativa* 33, 179–194.

López, N., Salcedo, A., Casaravilla, A., and Diconca, B. (2016). "percepción social del abandono en latinoamérica y españa. resultados de una encuesta-sondeo de opinión en línea," in *Proceedings of the III Conferencia Latinoamericana sobre el Abandono en la Educación Superior* en la Universidad Nacional Autónoma de México, Mexico.

Martínez, A. M., Zurita-Ortega, F., Castro-Sánchez, M. C., Chacon-Cuberos, R., Hinojo-Lucena, M. A., and Espejo-Garcés, T. (2016). La elección de estudio superiores universitarios en estudiantes de último curso de bachillerato y ciclos formativos. *Revista Electrónica Educare* 20, 1–18.

McCardle, L., Webster, E. A., Haffey, A., and Hadwin, A. F. (2017). Examining students' self-set goals for self-regulated learning: goal properties and patterns. *Stud. High. Educ.* 42, 2153–2169. doi: 10.1080/03075079.2015.1135117

Miranda, J. A., Montaño, B., and Sáez, M. A. (2013). "Causas del bajo rendimiento de los alumnos en la asignatura historia económica mundial y de españa," in *X Jornadas de Redes de Investigación en Docencia Universitaria: la participación y el compromiso de la comunidad universitaria*, eds M. T. Tortosa, J. D. Álvarez, and Y. N. Pellín, (Alicante: Universidad de Alicante), 1631–1647.

Montes, I. (2012). Investigación longitudinal de los hábitos de estudio en una cohorte de alumnos universitarios. *Revista Lasallista de Investigación* 9, 96–110.

Muñoz-Cantero, J., and Mato-Vásquez, M. D. (2014). El proyecto docente en la universidad española según el espacio europeo de educación superior. *Calidad en la educación* 40, 319–334. doi: 10.4067/S0718-45652014000100011

Navarro, I., González, C., López, F., Fernández, F., and Heliz, J. (2015). "Desarrollo de competencias en técnicas de estudio y relación con el rendimiento académico," in *XIII Jornadas de Redes de Investigación en Docencia Universitaria. Nuevas estrategias organizativas y metodológicas en la formación universitaria para responder a la necesidad de adaptación y cambio*, eds M. T. Tortosa, J. D. Álvarez, and Y. N. Pellín, (Alicante: Universidad de Alicante), 1633–1641.

Ng, E. M. (2018). Integrating self-regulation principles with flipped classroom pedagogy for first year university students. *Comput. Educ.* 126, 65–74. doi: 10.1016/j.compedu.2018.07.002

Ordóñez, R., and Rodríguez, M. R. (2015). Docencia en la universidad: valoraciones de los estudiantes de la universidad de sevilla. *Bordón: revista de pedagogía* 3, 85–101.

Pérez, R., Mercado, P., Martínez, M., Mena, E., and Partida, Á (2018). La sociedad del conocimiento y la sociedad de la información como la piedra angular en la innovación tecnológica educativa. *RIDE* 8, 847–870. doi: 10.23913/ride.v8i16.371

Rissanen, A. (2018). Student engagement in large classroom: the effect on grades, attendance and student experiences in an undergraduate biology course. *Can. J. Sci. Math. Technol. Educ.* 18, 136–153. doi: 10.1007/s42330-018-0015-2

Roksa, J., Trolian, T. L., Blaich, C., and Wise, K. (2017). Facilitating academic performance in college: understanding the role of clear and organized instruction. *High. Educ.* 74, 283–300. doi: 10.1007/s10734-016-0048-2

Soria-Barreto, K., and Zúñiga-Jara, S. (2014). Aspectos determinantes del éxito académico de estudiantes universitarios. *Formación universitaria* 7, 41–50. doi: 10.4067/S0718-50062014000500006

Stinebrickner, R., and Stinebrickner, T. (2014). Academic performance and college dropout: using longitudinal expectations data to estimate a learning model. *J. Labor Econ.* 32, 601–644. doi: 10.1086/675308

Tirado-Morueta, R., and Aguaded-Gómez, J. I. (2014). Influencias de las creencias del profesorado sobre el uso de la tecnología en el aula. *Revista de Educación* 363, 230–255. doi: 10.4438/1988-592X-RE-2012-363-179

Tomlinson, M. (2017). Student perceptions of themselves as 'consumers' of higher education. *Br. J. Sociol. Educ.* 38, 450–467. doi: 10.1542/peds.2014-2195

Trevors, G., Feyzi-Behnagh, R., Azevedo, R., and Bouchet, F. (2016). Self-regulated learning processes vary as a function of epistemic beliefs and contexts: mixed method evidence from eye tracking and concurrent and retrospective reports. *Learn. Instr.* 42, 31–46. doi: 10.1016/j.learninstruc.2015.11.003

Tuero, E., Cervero, A., Esteban, M., Arriaga, J., and Bernardo, A. (2017). "La transición a la universidad: eficacia de un programa de orientación en Bachillerato apoyado en las TIC," in *Proceedings of the VII Conferencia Latinoamericana sobre el Abandono en la Educación Superior en la Universidad de Córdoba*, Argentina.

Tuero, E., Cervero, A., Esteban, M., and Bernardo, A. (2018). ¿Por qué abandonan los alumnos universitarios? variables de influencia en el planteamiento y consolidación del abandono. *Educación* 21, 131–154. doi: 10.5944/educxx1.21.1

Velázquez, Y., and González, M. A. (2017). Factores asociados a la permanencia de estudiantes universitarios. *Revista de la Educación Superior* 46, 117–138. doi: 10.1016/j.resu.2017.11.003

Vossensteyn, J. J., Kottmann, A., Jongbloed, B. W., Kaiser, F., Cremonini, L., Stensaker, B., et al. (2015). *Dropout and Completion in Higher Education in Europe: Main Report*. Brussels: European Union

Wolters, C. A., and Hussain, M. (2015). Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacogn. Learn.* 10, 293–311. doi: 10.1007/s11409-014-9128-9

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.

Copyright © 2019 Bernardo, Esteban, Cervero, Cerezo and Herrero. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 22 October 2019
doi: 10.3389/fpsyg.2019.02225

Children's Academic, Artistic, and Athletic Competencies: Successes Are in the Eye of the Beholder

Sarah J. Racz[1], Diane L. Putnick[2], Gianluca Esposito[3,4] and Marc H. Bornstein[2,5]**

[1] University of Maryland, College Park, College Park, MD, United States, [2] Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD, United States, [3] Nanyang Technological University, Singapore, Singapore, [4] University of Trento, Trento, Italy, [5] Institute for Fiscal Studies, London, United Kingdom

A significant challenge to fully understanding children's academic and other competencies is dependency of the determination on the method of study, including notably who makes the assessment. This study examined similarities and differences in child, mother, father, and teacher reports of children's competencies across multiple domains of math, reading, music, and sports from two separate perspectives of rater agreement, mean level and order association. Two hundred and sixty-seven European American families were recruited from the mid-Atlantic region of the United States, and children, mothers and fathers, and teachers completed a commonly used rating measure of children's competencies when the children were 10 years of age. Results showed (1) high levels of order agreement (perhaps reflecting the observable nature of children's competencies), (2) some systematic mean level differences across raters, and (3) little inter-domain agreement (except among teachers, which may reflect teachers' unique perspectives on children's competencies). The educational, developmental, and methodological implications of the findings are discussed in the context of children's school performance. Who makes the determination of children's several different competencies matters.

Keywords: academic competencies, mean differences, rank order agreement, perceptions of academic performance, teacher perspectives

OPEN ACCESS

Edited by:
José Carlos Núñez, Universidad de Oviedo, Spain

Reviewed by:
Pedro Rosário, University of Minho, Portugal
Alejandro Veas, University of Alicante, Spain

****Correspondence:***
Gianluca Esposito
gianluca.esposito@ntu.edu.sg
Marc H. Bornstein
marc.h.bornstein@gmail.com

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

Received: *17 May 2019*
Accepted: *17 September 2019*
Published: *22 October 2019*

Citation:
Racz SJ, Putnick DL, Esposito G and Bornstein MH (2019) Children's Academic, Artistic, and Athletic Competencies: Successes Are in the Eye of the Beholder. Front. Psychol. 10:2225. doi: 10.3389/fpsyg.2019.02225

INTRODUCTION

Individual perceptions of children's academic and other competencies have important implications for their scholastic achievements and classroom adjustment as well as their overall feelings of self-worth (Wigfield et al., 1991; McGrath and Repetti, 2000; Wentzel et al., 2016). Importantly, systematic differences in reports of children's competencies in the school setting and elsewhere may create conflicts, biases, and expectations regarding children's academic and other abilities (Rosenthal, 1994). Prior research suggests that these perceptions tend to become self-fulfilling prophecies, such that children perceived as more or less competent "live up or down" to those expectations and subsequently perform well or poorly in those domains (i.e., the "Pygmalion Effect"; Jussim and Harber, 2005). What is not clear, however, is whether children's performance is due to actual individual differences in their abilities or to variability in how different observers judge their scholastic and other competencies (i.e., inter-rater agreement) or to how their competences are evaluated across different domains of performance (inter-domain agreement). Few studies

have examined multi-reporter similarities and differences in perceptions of children's performance in multiple domains commonly experienced during elementary school (i.e., math, reading, music, and sports). One aim of the current study is to address this gap in the literature.

Consider academics. Understanding how different reporters rate children's scholastic competencies across different domains is important as it may help identify areas where there is a mismatch between expectations for how a student will perform academically. That is, disagreement in reporters' ratings of children's academic competences may lead to conflict or other difficulties in navigating educational settings. For instance, if a parent perceives a child's academic competency to be higher than what the child's teacher perceives it to be, the parent may consider any poor grades or academic difficulties to reflect unfair or biased perceptions of the child's teacher. Often this conflict is based around interpersonal differences in educational values or expectations, and understanding how perceptions of academic competencies may contribute to such conflict may help identify ways in which to strengthen or remediate key parent–teacher learning partnerships to better support children's academic success (Crozier, 1999; Creech and Hallam, 2003). Similarly, differences in perceptions of children's competencies across various academic domains may help identify areas where children may need more support, or where their competencies are less readily observable to multiple reporters. For areas where children's competency may be less observable, it may be important for teachers, parents, and children to communicate openly and regularly (e.g., parent–teacher conferences, home-school notes). Awareness of differences and similarities in these inter-rater and inter-domain perceptions may help reduce potential interpersonal conflicts and enhance the ability of parents, children, and teachers to work together to best support children's academic progress. Thus, a second aim of the current study was to provide concrete information on agreement of children, parents, and teachers about multiple domains of children's competencies.

Several theoretical frameworks have been advanced to understand children's academic competencies. For instance, differences in children's competencies across reporters and academic domains have been conceptualized within dimensional comparison theory (Möller and Marsh, 2013; Marsh et al., 2014), whereby ratings of competencies in one domain are determined, in part, by how achievement in that domain is perceived compared to achievements in other areas. Ratings of children's academic competencies may also reflect both internal (comparing children's competencies in one domain with their competencies in another domain) and external (comparing the child's competencies in one domain to the competencies demonstrated by other children in that domain) processes (Marsh, 1986). More salient to the current study is Expectancy-Value Theory (EVT; Wigfield and Eccles, 2000), which asserts that expectations and values regarding achievement directly influence academic choices, performance, effort, and persistence. Such expectancies and values are influenced by various social and cognitive factors including beliefs about general abilities and perceived difficulty of tasks, which in turn are influenced by individuals' perceptions of their own academic experiences as well as perceptions from others (e.g., parents and teachers; socialization influences). In many ways this theory proposes a mediational model of academic achievement, whereby perceptions of academic competencies (both by the self and others) lead to social and cognitive factors which lead in turn to academic expectations and values which then lead to academic achievement outcomes. Much research has documented links among expectations, values, and outcomes (e.g., Eccles and Wigfield, 1995; Wigfield et al., 1997; Simpkins et al., 2012), but less work has explored the potential role of academic perceptions of the self and other socializers. As noted above, these self and other perceptions may be particularly important to supporting student successes in an educational setting. An essential step in highlighting how these factors fit within the broader EVT model is to understand reporting patterns in perceptions of children's academic competencies across individuals and domains. This understanding constituted a third aim of the current study.

Ratings of Academic Competencies Across Reporters and Domains

When considering individual differences in ratings of children's academic competencies, it is important to acknowledge that findings documenting any unique perspectives show that the degree of inter-rater disagreement often depends on the reporters being compared. Specifically, agreement between mothers' and fathers' ratings tends to be higher than that between other reporters (e.g., parents versus teachers), which may be due in part to the fact that mothers and fathers tend to observe children in the same contexts (i.e., home; Kerr et al., 2007; Schroeder et al., 2010; De Los Reyes et al., 2015). Between different settings (i.e., school versus home), teachers tend to report significantly fewer child difficulties as compared to parents and children (Salbach-Andrae et al., 2009). Given that teachers are exposed to a range of students, the abilities exhibited by any one student may not seem as noteworthy to teachers as they do to parents, who tend to view their children's capabilities in more isolated and individualized settings. Teachers may therefore have a unique perspective on children's academic competencies due to their extensive interactions with children in academic settings.

In addition to inter-reporter differences, the informant discrepancy literature indicates that ratings of children's functioning tend to vary depending on whether the domain being assessed is easily observable or not (De Los Reyes et al., 2015). It may therefore be the case that greater inter-rater agreement is observed within academic domains that are easy to track and observe through homework assignments and test scores (e.g., math and reading) as compared to those areas with less observable and quantifiable outcomes (e.g., music and sports). Furthermore, differences in reporters' ratings often reveal meaningful information about children's behavior in various contexts (e.g., home versus school) and domains (e.g., internal versus external behaviors), and therefore have important implications for the identification of difficulties in children's cognitive, academic, social, emotional, and behavioral competencies (De Los Reyes, 2011; De Los Reyes et al., 2013).

For instance, children's academic functioning may only become evident in the context of the cognitive and behavioral demands of core academic classrooms (Rimm-Kaufman et al., 2001), further highlighting the importance of comparing ratings of children's academic competence across reporters and domains. Multi-informant assessment approaches therefore constitute a promising avenue to fully evaluate children's competencies across both raters and domains. Overall, extensive research findings indicate that differences in reports of children's functioning tend to vary depending on both the raters and domains assessed. The question of which raters and which domains are so influenced is only answered by a multi-informant multi-domain design with the same children as we use here.

Differential associations among various raters' reports of children's academic competencies, for example, underscore the importance of considering differences in perceptions of children's competencies across reporters and domains. Specifically, children's self-perceptions of their academic competencies have been reported to be more strongly associated with their mothers' than their teachers' perceptions of their academic competencies (Wigfield et al., 1997). However, teachers' ratings of children's literacy and math competencies in early elementary school predict children's perceptions of their literacy and math competencies, respectively, in later elementary school, whereas only parents' judgments of children's math competence predict children's later perceptions of their math competence (Herbert and Stipek, 2005). These differences in teachers' perceptions of children's academic competencies may be due to several factors noted above, including teachers' unique perspectives in the classroom, the availability of official examination standards which may inform their perceptions, and the potential for their expectations to create a self-fulfilling prophecy for their students. More attention to these differential associations in reports of children's academic competencies may identify additional explanatory factors.

Of note, most previous studies have investigated perceptions of children's academic competencies when children are in the early elementary school years (i.e., Kindergarten and 1st grade; Wigfield et al., 1997; Herbert and Stipek, 2005). It is therefore not known if these findings extend to the later elementary school years, which represent a crucial educational transition point as children prepare to enter middle/secondary school. The current study therefore provides a developmental extension of the extant literature by examining these associations when children are in the later elementary school years (i.e., approximately 5th grade). The current study also extends findings regarding agreement in raters' reports of children's academic and other competencies to incorporate perspectives of children's competencies from *four* reporters (i.e., children, mothers, fathers, and teachers) and across *four* academic domains (i.e., math, reading, music, and sports) commonly experienced in the school setting.

Furthermore, the current study provides a methodological contribution to the study of individual differences in perceptions of children's competencies. Reporter agreement and disagreement can be evaluated using (at least) two different approaches to analysis – mean level and rank order. *Mean level* analyses assess similarities and differences in ratings for the whole group. For example, if children tend to view themselves as more competent in a given domain than parents view them, we would find a significant mean level difference between raters. Rank order analyses assess similarities and differences in correlations between raters. For example, agreement would be considered strong if two raters assess one individual's math competence high and another individual's math competence low relative to others in a group. Agreement would be less strong if one rater ranks individuals' competence relatively high and another rater ranks competence relatively low in the group (and vice versa). Clearly, the two analytical approaches provide different sorts of information about rater agreement. In the current study, we assess both approaches to analysis to provide a more complete picture of inter-rater and inter-domain judgments of children's competencies.

Overview of the Current Study

The current study examines child, mother, father, and teacher ratings of children's competencies in four domains commonly experienced in the school setting (i.e., math, reading, music, and sports). Our methodological approach allowed us to disentangle rater agreement that is general (e.g., similar mother-father agreement holds across all domains) from rater agreement that is specific to particular domains (e.g., raters agree about competence in music but not sports). Likewise, we disentangle rating patterns that are general to all raters (e.g., correlations between math and reading competence are similar for all raters) from rating patterns that are specific to a particular rater (e.g., child-rated competence in all domains is correlated, but mother-rated competence differs across domains). Thus, the current study adds to our understanding of how children's academic, artistic, and athletic competencies are perceived by different raters and across different domains.

The first specific aim of the current study was to evaluate both correlational and mean level similarities and differences in ratings of perceptions of children's math, reading, music, and sports competencies according to children, mothers, fathers, and teachers. Consistent with the broader literature reviewed above, for inter-rater comparisons (Hypothesis 1a), we hypothesized higher agreement (i.e., larger correlations and fewer mean level differences) regarding children's competencies between mothers and fathers than between the other pairs of reporters. We also hypothesized that teachers' perceptions of children's competencies would be the most different from the other reporters. Furthermore, we hypothesized that teachers would provide the highest mean competence ratings out of the four reporters. Based on literature highlighting differences in reports of children's functioning according to the domain assessed (De Los Reyes et al., 2015), we also hypothesized about several inter-domain comparisons among math, reading, music, and sports competencies (Hypothesis 1b). Specifically, we hypothesized higher agreement (i.e., larger correlations and fewer mean differences) regarding reports of children's competencies in math and reading than those for music and sports, given that performance in these academic domains is more easily observable through homework and test grades. This hypothesis is supported by literature indicating that children

are viewed as more competent in core academic courses as compared to extracurricular activities (Wigfield et al., 1997; Jacobs et al., 2002).

Our second specific aim was to compare patterns of inter-reporter agreement and disagreement about children's competencies across all four domains (Hypothesis 2a) and patterns of inter-domain agreement and disagreement across all four reporters (Hypothesis 2b). Most previous studies have evaluated reporter agreement at the bivariate level, which may mask more nuanced variations in ratings of children's competencies. For instance, it may be that correlation matrices of math and reading differ, suggesting potential differences in the overall degree of reporter agreement depending on the competence domain assessed. We therefore compared correlation matrices of child, mother, father, and teacher reports of children's competencies in math, reading, music, and sports across reporters and domains (e.g., comparing the matrix of correlations of children's competencies according to mother report to the matrix of correlations according to teacher report; comparing the correlation matrix of children's math competencies to the matrix of correlations of children's reading competencies). Given the novelty of these analyses, we did not have any specific directional hypotheses for this study goal.

MATERIALS AND METHODS

Participants

European American families with healthy first-born and second-born children were recruited through newspaper advertisements and mass mailings from the mid-Atlantic region of the United States. The current analyses focus on data collected from children, mothers, fathers, and teachers when children were 10 years old. There were no competence data available for 84 (23.93%) of the 351 families who contributed data to the 10-year assessment. We therefore only analyzed data from families where at least one reporter provided information regarding children's competencies in at least one of the four domains (i.e., math, reading, music, and sports), yielding a total sample size of 267 for the current study. Children were on average 10.27 years of age (*SD* = 0.18, *range* = 9.76 – 10.90) and approximately half (*n* = 136; 50.94%) were boys. Mothers were on average 41.33 years old (*SD* = 5.18, *range* = 26.94 – 55.87), and fathers were 43.58 years (*SD* = 6.41, *range* = 27.78 – 67.93). The majority of the mothers and fathers were married and living together (*n* = 214; 80.15%) and had obtained at least a bachelor's degree (mothers: 67.79%, *n* = 181; fathers: 63.67%, *n* = 170). Families were on average from a middle to upper socioeconomic status (SES) on the Hollingshead (1975) Four-Factor Index of Social Status, and ranged from lower to upper socioeconomic status (*M* = 54.77, *SD* = 9.90, *range* = 25 – 66). Of the teachers who provided data to the 10-year assessment, the majority were 4th (37.45%, *n* = 100) or 5th (27.72%, *n* = 74) grade teachers (6 [2.25%] from combined 3rd–4th or 4th–5th grade classrooms; 87 [32.6%] missing/no response). Written informed consent was obtained from the parents of child participants and from parents and teachers for their participation.

Procedure

Measures collected for the current study were part of a larger assessment battery that included home and laboratory visits. Children's competencies were assessed at the home visit. Packets of questionnaires were mailed to mothers and fathers prior to their visits. Mothers were also asked to provide their children's teachers with a packet of questionnaires to complete. The packet contained a letter explaining the study, a consent form, and a self-addressed stamped envelope to return the completed questionnaires directly to the research team. Informed consent was obtained from mothers, fathers, and teachers, and assent was obtained from children. Study procedures were approved and monitored by our Institutional Review Board.

Measures

Children's Competencies

Children, mothers, fathers, and teachers each completed the Children's Competence Beliefs and Subjective Task Values (CBTV; Eccles et al., 1993; Eccles and Wigfield, 1995; Wigfield et al., 1997) measure to assess individual perceptions of children's competencies in the domains of math, reading, music, and sports. Reporters are asked to rate children's competencies in these domains on a 7-point rating scale, with higher scores indicating better competence. Subscale scores used in the current study were based on a factor analysis conducted by Wigfield et al. (1997). Items used to assess children's, mothers', fathers', and teachers' perceptions of children's competencies in math, reading, music, and sports are included in the **Appendix**. Subscale scores were calculated by taking the mean of the items within each domain according to each reporter. Summary means across raters and domains were calculated by taking the average of these subscales (e.g., the child-report summary mean was calculated by taking the average of the child-reported math, reading, music, and sports subscales). The subscales used in the current study demonstrated good reliability according to child (α_{Math} = 0.86, $\alpha_{Reading}$ = 0.86, α_{Music} = 0.76, α_{Sports} = 0.91), mother (α_{Math} = 0.90, $\alpha_{Reading}$ = 0.92, α_{Music} = 0.89, α_{Sports} = 0.92), father (α_{Math} = 0.93, $\alpha_{Reading}$ = 0.93, α_{Music} = 0.91, α_{Sports} = 0.93), and teacher (α_{Math} = 0.89, $\alpha_{Reading}$ = 0.89, α_{Music} = 0.81, α_{Sports} = 0.89) reports. In keeping with our current focus on perceptions of children's academic competencies, we only examined items relevant to children's competencies and not items related to subjective value reports.

Analysis Plan

Analyses were conducted in SPSS 21 (for Hypothesis 1) and *Mplus* 7.2 (for Hypothesis 2; Muthén and Muthén, 2014). To account for the nested structure of the data from families where two children participated (i.e., first- and second-borns), we clustered based on family and used the Huber-White adjustment of the standard errors to account for non-independence. We used a maximum likelihood estimator that calculated robust standard errors (MLR; Little and Rubin, 2002). Missing data were due to non-response of one or more family members and to non-response to individual items. Specifically, data regarding 10-year competencies were available from all 267 children, 225 (84.27%)

mothers, 202 (75.66%) fathers, and 198 (74.16%) teachers. Missing data were handled with full-information maximum likelihood (FIML) in *Mplus*. Data were missing completely at random, as evidenced by a non-significant Little's MCAR test, $\chi^2(336) = 338.10$, $p = 0.458$.

For our first aim investigating mean and correlational similarities and differences in reports of children's competencies, we first examined inter-rater and inter-domain bivariate correlations to examine degrees of agreement in ratings of children's competencies in math, reading, music, and sports. Specifically, we tested for differences in dependent inter-rater and inter-domain correlations via a web utility developed by Lee and Preacher (2013) based on methods identified by Steiger (1980). We then conducted a series of paired-samples *t*-tests to determine if, at the bivariate level, the mean-level ratings of children's competencies were different among individual reporters and domains. We also examined if the summary means differed across reporters and domains.

For our second aim comparing patterns of inter-reporter and inter-domain agreement about children's competencies, all competence variables were standardized prior to analyses to aid in interpretation of the coefficients. We first estimated two correlation matrices (e.g., the correlation matrices of children's math and reading competencies) simultaneously and allowed all bivariate correlations within those matrices to be freely estimated (i.e., an unconstrained, fully free model). We then constrained all paired bivariate associations within the correlation matrices to be equal to each other (i.e., a fully constrained model). For instance, when comparing the correlation matrices of math and reading, we constrained the correlation of child and mother reports of math to be equal to the correlation of child and mother reports of reading, and so on for all six paired correlations (refer to **Table 1**). If this fully constrained model provided a significantly worse fit to the data than the unconstrained model, we examined modification indices and freed constraints with the highest modification indices (Yoon and Kim, 2014). Individual constraints were therefore freed progressively, until the nested model test reached non-significance (i.e., a partially constrained model). Nested model tests (comparing the constrained versus unconstrained models) were conducted via the Satorra-Bentler scaled $\Delta\chi^2$-difference test (Satorra and Bentler, 2001). Non-significant $\Delta\chi^2$-difference tests indicated that a more restrictive model (i.e., the model with more parameter equality constraints) did not provide a significantly worse fit to the data than a less restrictive model (i.e., the model with fewer parameter constraints).

RESULTS

Hypothesis 1: Rank-Order and Mean-Level Similarities and Differences in Ratings of Children's Competencies

Hypothesis 1a: Inter-Rater Comparisons

Correlations in **Table 1** (above the diagonal) indicate significant positive associations between all reporters, such that higher

TABLE 1 | Means, standard deviations, and inter-rater correlations of child math, reading, music, and sports competencies.

	A *Math Competence*				B *Reading Competence*				C *Music Competence*				D *Sports Competence*				Mean (*SD*)
Variable	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*	
(1) Child	–	0.57***	0.48***	0.36***	–	0.63***	0.55***	0.40***	–	0.45***	0.49***	0.38***	–	0.67***	0.65***	0.36***	5.26 (0.67)
(2) Mother	0.68	–	0.70***	0.48***	0.60	–	0.78***	0.52***	0.80	–	0.53***	0.24**	0.55	–	0.74***	0.43***	5.27 (0.74)
(3) Father	0.77	0.51	–	0.40***	0.70	0.39	–	0.53***	0.76	0.72	–	0.30***	0.58	0.45	–	0.53***	5.19 (0.75)
(4) Teacher	0.87	0.77	0.84	–	0.84	0.73	0.72	–	0.86	0.94	0.91	–	0.87	0.82	0.72	–	5.18 (1.04)
Mean (*SD*)	5.34 (1.08)	5.70 (1.09)	5.67 (1.10)	5.55 (1.19)	5.57 (1.02)	5.74 (1.24)	5.73 (1.25)	5.46 (1.20)	4.80 (1.29)	4.69 (1.35)	4.66 (1.26)	4.60 (1.32)	5.38 (1.30)	4.91 (1.40)	4.71 (1.33)	4.95 (1.57)	

*Numbers above the diagonal in matrices are correlations; numbers below the diagonal indicate proportions of unexplained variance between the two variables. Possible scores on the CBTV range from 1 to 7. **$p < 0.01$. ***$p < 0.001$.*

ratings of children's competencies according to one reporter are related to higher competence ratings according to the other reporters. Testing for differences in these dependent correlations revealed that mothers and fathers tended to agree with each other more than the other reporters for all four domains (all $zs > 2.43$; all $ps < 0.05$), except when compared to child–mother ($z = 1.77$, $p = 0.076$) and child–father ($z = 0.87$, $p = 0.384$) agreement in music competency. Children and mothers tended to agree more than children and fathers on math ($z = 2.63$, $p = 0.008$) and reading ($z = 2.89$, $p = 0.004$) competencies; more than children and teachers on math ($z = 4.57$, $p < 0.001$), reading ($z = 5.47$, $p < 0.001$), and sports ($z = 6.93$, $p < 0.001$) competencies; and more than mothers and teachers on reading ($z = 2.50$, $p = 0.013$), music ($z = 3.87$, $p < 0.001$), and sports ($z = 5.23$, $p < 0.001$) competencies. Children and fathers also tended to agree more than children and teachers on competencies in all four domains (all $zs > 2.02$; all $ps < 0.05$), and more than fathers and teachers on music ($z = 3.60$, $p < 0.001$) and sports ($z = 2.71$, $p = 0.007$) competencies. Mothers and teachers agreed more than children and teachers on math ($z = 2.74$, $p = 0.006$) and reading ($z = 3.02$, $p = 0.002$) competencies, and children and teachers agreed more than mothers and teachers on music competency ($z = 2.67$, $p = 0.008$). Mothers and teachers also agreed more than fathers and teachers on math competency ($z = 2.19$, $p = 0.028$), and fathers and teachers agreed more than children and teachers on reading competency ($z = 3.00$, $p = 0.003$). Fathers and teachers also agreed more than children and teachers ($z = 4.38$, $p < 0.001$) and mothers and teachers ($z = 3.02$, $p = 0.003$) on sports competency. Hence, as hypothesized, mothers and fathers tended to have the strongest agreement and teachers tended to have lower agreement with all other reporters.

Overall tests of mean-level inter-rater agreement (aggregating across domains) indicated that children, mothers, fathers, and teachers provided similar ratings of children's competencies.

However, within-domain paired-samples *t*-tests indicated some degree of systematic inter-reporter disagreement (see **Table 2**), such that children and mothers rated children's math, reading, and sports competencies differently (children rated themselves lower than mothers in math and reading but higher in sports competencies). Children and fathers rated children's math, music, and sports competencies differently (children rated themselves lower than fathers in math but higher in music and sports competencies). Children and teachers and mothers and fathers only differed in ratings of sports competence, with children rating themselves higher than teachers and mothers providing higher ratings than fathers. Mothers also provided higher ratings of children's math and reading competencies compared to teachers, and fathers provided higher ratings of children's reading competencies compared to teachers.

Hypothesis 1b: Inter-Domain Comparisons

Compared to the inter-rater comparisons as hypothesized, fewer significant correlations (66.67%) obtained between domains (see **Table 3**). Specifically, only ratings of children's reading and music competencies were associated according to child report. According to mother and father reports, children's math and reading competencies, math and sports competencies,

TABLE 2 | Average inter-rater difference scores of child math, reading, music, and sports competencies.

	Math Competence			*Reading Competence*			*Music Competence*			*Sports Competence*			*Overall (Across Domains)*		
Variable	**Mean (*SD*)**	***t* (*df*)**	***d***	**Mean (*SD*)**	***t* (*df*)**	***d***	**Mean (*SD*)**	***t* (*df*)**	***d***	**Mean (*SD*)**	***t* (*df*)**	***d***	**Mean**	***t* (*df*)**	***d***
Children vs. mothers	−0.35 (0.99)	−5.34 (223)***	−0.33	−0.16 (0.98)	−2.50 (215)*	−0.15	0.13 (1.39)	1.35 (218)	0.10	0.46 (1.10)	6.13 (214)***	0.34	−0.01	−0.14 (223)	−0.01
Children vs. fathers	−0.34 (1.11)	−4.32 (197)***	−0.32	−0.15 (1.11)	−1.91 (195)	−0.13	0.21 (1.29)	2.31 (197)*	0.17	0.67 (1.10)	8.49 (194)***	0.51	0.07	1.56 (199)	0.12
Children vs. teachers	−0.14 (1.27)	−1.53 (192)	−0.12	0.09 (1.22)	1.03 (187)	0.08	0.19 (1.47)	1.67 (174)	0.14	0.51 (1.63)	4.16 (177)***	0.35	0.08	1.40 (194)	0.12
Mothers vs. fathers	0.04 (0.87)	0.65 (180)	0.04	0.03 (0.82)	0.45 (181)	0.02	−0.06 (1.23)	−0.62 (178)	−0.05	0.16 (1.00)	2.19 (180)*	0.12	0.08	0.89 (181)	0.06
Mothers vs. teachers	0.24 (1.14)	2.81 (170)**	0.22	0.30 (1.16)	3.42 (171)**	0.26	0.04 (1.67)	0.32 (153)	0.03	0.07 (1.57)	0.54 (162)	0.05	0.09	1.65 (172)	0.14
Fathers vs. teachers	0.16 (1.05)	1.05 (152)	0.09	0.22 (1.16)	2.39 (154)*	0.19	0.07 (1.54)	0.51 (141)	0.05	−0.13 (1.45)	−1.11 (147)	−0.09	0.01	0.12 (155)	0.01

*t-statistics indicate if the means are significantly different from each other (paired-samples t-test). *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.*

TABLE 3 | Means, standard deviations, and inter-domain correlations of child math, reading, music, and sports competencies according to child, mother, father, and teacher reports.

	A Child Report				B Mother Report				C Father Report				D Teacher Report				Across-rater Mean (*SD*)
Variable	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***	***1***	***2***	***3***	***4***	
(1) Math	–	0.09	0.01	0.12	–	0.40***	0.13	0.14*	–	0.42***	0.16*	0.15*	–	0.68***	0.50***	0.43***	5.51 (0.92)
(2) Reading	0.99	–	0.31***	-0.10	0.84	–	0.17*	−0.14*	0.82	–	0.18*	-0.08	0.54	–	0.53***	0.24**	5.60 (1.03)
(3) Music	1.00	0.90	–	0.09	0.98	0.97	–	0.02	0.97	0.97	–	0.18*	0.75	0.72	–	0.43***	4.68 (1.02)
(4) Sports	0.99	0.99	0.99	–	0.98	0.98	1.00	–	0.98	0.99	0.97	–	0.82	0.94	0.82	–	5.02 (1.18)
Mean (*SD*)	5.34 (1.08)	5.57 (1.02)	4.80 (1.29)	5.38 (1.30)	5.70 (1.09)	5.74 (1.24)	4.69 (1.35)	4.91 (1.40)	5.67 (1.10)	5.73 (1.25)	4.66 (1.26)	4.71 (1.33)	5.55 (1.19)	5.46 (1.20)	4.60 (1.32)	4.95 (1.57)	

*Numbers above the diagonal in matrices are correlations; numbers below the diagonal indicate proportions of unexplained variance between the two variables. Possible scores on the CBTV range from 1 to 7. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.*

and reading and music competencies were positively related. However, the association between children's reading and sports competencies was negative according to mother report, indicating that mothers viewed more competence in reading as related to less competence in sports. According to fathers' reports, associations between children's math and music competencies, and music and sports competencies, were positive. All correlations according to teacher reports were positive, indicating that teachers perceived more child competence in one domain as related to more competence in other domains.

Testing for differences in these dependent correlations indicated that mothers, fathers, and teachers all viewed children's math and reading competencies as more strongly related with each other than relations among the other domains (all $zs > 3.45$; all $ps < 0.01$). Children viewed their reading and music competencies as more strongly correlated with each other than relations among the other domains (all $zs > 2.88$; all $ps < 0.01$), and mothers, fathers, and teachers perceived children's reading and music competencies as more strongly related than their reading and sports competencies (all $zs > 3.84$; all $ps < 0.001$). All reporters viewed music and sports competencies as more highly related than reading and sports competencies (all $zs > 2.33$; all $ps < 0.05$). Last, mothers, fathers, and teachers viewed child math and sports competencies as more closely related than reading and sports competencies (all $zs > 4.02$; all $ps < 0.001$), and children saw their reading and sports competencies as more highly associated than their math and sports competencies ($z = 3.07$, $p = 0.002$).

Consistent with the lower inter-domain correlations, paired-samples *t*-tests indicated more systematic inter-domain disagreement (see **Table 4**) than was observed among raters. Overall tests of mean-level inter-domain agreement (aggregating across reporters) indicated that ratings of children's math competence were not significantly different from overall ratings of reading competence, but that overall math and reading competencies were rated higher than music and sports. Overall ratings of children's sports competence were also higher than overall ratings of children's music competence. In within-reporter paired *t*-tests, competency ratings of math compared to reading were different only according to child-report (with reading rated higher than math); however, children's math and reading competencies were consistently rated higher than their music competency according to all four reporters. Children's math and reading competencies were rated as higher than their sports competency according to mothers, fathers, and teachers, whereas children's sports competency was rated higher than their music competency according to children and teachers.

Hypothesis 2: Comparing Patterns of Agreement and Disagreement in Ratings of Children's Competencies

Hypothesis 2a: Inter-Rater Patterns Across All Four Domains

Comparing the matrices of domain correlations (e.g., comparing panels A, B, C, and D in **Table 3**) revealed differences in inter-rater patterns of agreement. Only the fully constrained

TABLE 4 | Average inter-domain difference scores of child math, reading, music, and sports competencies.

Variable	Child Report			Mother Report			Father Report			Teacher Report			Overall (Across Reporters)		
	Mean (SD)	t (df)	d	Mean (SD)	t (df)	d	Mean (SD)	t (df)	d	Mean (SD)	t (df)	d	Mean	t (df)	d
Math vs. reading	−0.22 (1.42)	−2.45 (248)*	−0.21	−0.04 (1.28)	−0.43 (224)	−0.03	−0.07 (1.27)	−0.72 (200)	−0.06	0.11 (0.96)	1.56 (193)	0.09	−0.09	−1.10 (263)	−0.07
Math vs. music	0.54 (1.68)	5.22 (260)***	0.46	1.02 (1.62)	9.31 (220)***	0.83	1.01 (1.53)	9.38 (200)***	0.86	0.91 (1.26)	9.64 (177)***	0.73	0.83	10.54 (266)***	0.86
Math vs. sports	−0.04 (1.59)	−0.37 (247)	−0.03	0.79 (1.65)	7.22 (224)***	0.63	0.96 (1.59)	8.51 (199)***	0.79	0.61 (1.51)	5.55 (185)***	0.44	0.49	5.97 (263)***	0.47
Reading vs. music	0.74 (1.37)	8.52 (248)***	0.64	1.06 (1.68)	9.41 (220)***	0.82	1.07 (1.61)	9.49 (201)***	0.86	0.80 (1.23)	8.72 (177)***	0.64	0.92	12.27 (263)***	0.90
Reading vs. sports	0.19 (1.74)	1.71 (248)	0.16	0.83 (1.99)	6.24 (224)***	0.63	1.02 (1.89)	7.63 (200)***	0.79	0.52 (1.73)	4.08 (184)***	0.37	0.58	5.55 (263)***	0.52
Music vs. sports	−0.55 (1.75)	−4.94 (247)***	−0.42	−0.19 (1.92)	−1.50 (220)	−0.14	−0.04 (1.65)	−0.36 (200)	−0.03	−0.29 (1.56)	−2.49 (175)*	−0.20	−0.34	−3.71 (263)***	−0.31

*t-statistics indicate if the means are significantly different from each other (paired-samples t-test). *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.*

model comparing mother and father intercorrelations did not provide a significantly worse fit to the data than the fully free (i.e., unconstrained) model, $\Delta\chi^2(6) = 4.04$, $p = 0.671$. This finding points to global similarities in the associations between mother and father reports of children's math, reading, music, and sports competencies. The fully constrained models comparing inter-domain correlation matrices for child and mother, child and father, child and teacher, mother and teacher, and father and teacher reports provided a significantly worse fit to the data than the fully free models, $\Delta\chi^2_{child-mother}(6) = 22.43$, $p = 0.001$; $\Delta\chi^2_{child-father}(6) = 20.71$, $p = 0.002$; $\Delta\chi^2_{child-teacher}(6) = 51.67$, $p < 0.001$; $\Delta\chi^2_{mother-teacher}(6) = 36.18$, $p < 0.001$; $\Delta\chi^2_{father-teacher}(6) = 19.50$, $p = 0.003$. Constraints were progressively freed until the nested model test reached non-significance.

The partially constrained models comparing child and mother, and child and father, reports did not provide a significantly worse fit to the data than the fully free models, $\Delta\chi^2_{child-mother}(5) = 8.64$, $p = 0.124$; $\Delta\chi^2_{child-father}(5) = 7.78$, $p = 0.169$. In these models the associations between math and reading competencies were freed, and all other inter-domain associations were constrained to be equal, indicating similarities in the associations among these domains according to child and mother, and child and father, reports. However, the association between child-reported math competence and child-reported reading competence ($r = 0.10$) was lower than the associations between mother- and father-reported math competence and mother- and father-reported reading competence (both rs = 0.40; both zs = −3.71, $p < 0.001$).

Nested model tests did not reach non-significance after progressively freeing all constraints for the models comparing child and teacher, mother and teacher, and father and teacher reports, and therefore for these comparisons the fully free models provided the best fit to the data. All inter-domain associations were freely estimated, indicating global differences in ratings of children's math, reading, music, and sports competencies between these reporters. Specifically, child-, mother-, and father-reported associations were consistently lower (rs_{child} ranged from −0.10 to 0.31; rs_{mother} ranged from −0.14 to 0.40; rs_{father} ranged from −0.08 to 0.41) than teacher-reported associations (rs ranged from 0.25 to 0.68) across all domains (there is some overlap in these ranges, but the individual comparisons between child, mother, and father associations were all lower than the teacher associations; see **Table 3**; all zs < −3.86, ps < 0.001).

Hypothesis 2b: Inter-Domain Patterns Across All Four Raters

Comparing the matrices of reporter correlations (e.g., comparing panels A, B, C, and D in **Table 1**) revealed fewer differences in inter-domain patterns of agreement compared to the inter-rater patterns. The fully constrained models comparing ratings of children's competencies in math and reading, math and music, math and sports, reading and sports, and music and sports did not provide a significantly worse fit to the data than the fully free (i.e., unconstrained) models, $\Delta\chi^2_{math-reading}(6) = 2.82$, $p = 0.831$; $\Delta\chi^2_{math-music}(6) = 9.95$, $p = 0.127$; $\Delta\chi^2_{math-sports}(6) = 6.18$, $p = 0.403$; $\Delta\chi^2_{reading-sports}(6) = 4.17$, $p = 0.653$;

$\Delta\chi^2_{music-sports}(6) = 11.29$, $p = 0.080$. These findings indicated similarities in the associations between math and reading, math and music, math and sports, reading and sports, and music and sports across all four reporters (e.g., the correlation between child-reported and mother-reported math was similar to the correlation between child-reported and mother-reported reading).

The fully constrained model comparing ratings of children's reading and music competence provided a significantly worse fit to the data than the fully free model, $\Delta\chi^2(6) = 14.84$, $p = 0.021$. Constraints were progressively freed until the nested model test reached non-significance. The partially constrained model did not provide a worse fit to the data than the fully free model, $\Delta\chi^2(3) = 6.35$, $p = 0.096$. For the reading-music comparison, the associations between children and mothers, mothers and teachers, and fathers and teachers were progressively released and freely estimated. Examination of the correlation coefficients indicated that the associations between child and mother, mother and teacher, and father and teacher reports for reading competence were higher ($r_{child-mother} = 0.64$; $r_{mother-teacher} = 0.50$; $r_{father-teacher} = 0.52$) than those associations for music competence ($r_{child-mother} = 0.47$; $r_{mother-teacher} = 0.25$; $r_{father-teacher} = 0.29$), respectively ($z_{child-mother} = 2.85$, $p = 0.004$; $z_{mother-teacher} = 3.38$, $p = 0.001$; $z_{father-teacher} = 3.19$, $p = 0.001$). Associations between ratings from children and fathers, children and teachers, and mothers and fathers were constrained to be equal across reading and music competencies, indicating similarity in these associations.

DISCUSSION

The current study examined child, mother, father, and teacher perceptions of children's competencies in math, reading, music, and sports and utilized two revealing analytic approaches to evaluate inter-rater and inter-domain agreement, rank order and mean level. In doing so, this study extends previous findings with Expectancy-Value Theory (Eccles and Wigfield, 1995; Wigfield and Eccles, 2000) and supports the need to consider perceptions of the self and other socializers across multiple domains when conceptualizing children's overall academic, artistic, and athletic functioning. EVT posits that children's academic performance and behavior are influenced by relations among beliefs about how well children will complete an academic task and how much value is placed on those tasks. Several academic constructs are conceptualized in this theoretical model, including children's interests, affect, values, choices, effort, task difficulty, and *competence*. EVT proposes that perceptions of children's academic competencies constitute a first step in this process, whereby these views influence how much a child will value and express interest in an academic task. It stands to reason then that children who do not perceive themselves as having high levels of academic competence will subsequently demonstrate little interest and see little value in academic tasks. Other studies with EVT document that low levels of academic interest, expectancies, and subjective task value lead to low academic performance and outcomes (Wigfield, 1994; Eccles and Wigfield, 1995; Wigfield et al., 1997; Simpkins et al., 2012; Trautwein et al., 2012) as well as poor career attainment (Lauermann et al., 2017). As such, initial perceptions of children's academic competencies may have far-reaching effects on their later academic and educational success.

Our results suggest that academic perceptions of the self and other socializers tend to vary systematically. Extending the literature on EVT, perceptions of academic competence from other socializers are also important to consider, such that if children assume their parents or teachers perceive their competency in a certain domain to be relatively low, children may subsequently "live up" to those expectations (Jussim and Harber, 2005; Tomasetto et al., 2015). Perceptions of children's academic competencies from other socializers may also "spill over" into children's own views and expectations of how much they should value certain academic domains. For instance, if a child's parent perceives the child to have high levels of math competency but low levels of reading competency, the child may place more value and effort in math and less in reading. The child may then eventually develop difficulties in reading as a result of those perceptions. Discrepancies in perceptions are also important to consider. That is, perceptions of high levels of math competency according to the child, but low levels of math competency according to the parent, may create conflict in expectations and values, which may in turn negatively impact children's academic performance. Borrowing from the broader reporter discrepancy literature, the most negative outcomes are observed when two reporters harbor differing viewpoints on a child's functioning and behavior (De Los Reyes et al., 2015). Similar findings may be observed when applied to children's academic competencies and the broader EVT model.

Clearly, understanding how to incorporate inter-reporter and inter-domain discrepancies into theoretical models of children's academic competencies is a crucial direction for future research. More work is also needed to understand how these differences in perceptions relate to social cognitive factors that influence academic performance. Models examining EVT should therefore consider the potential for discrepancies in reports of children's academic competencies across raters and domains, and several methodological advances have provided the statistical modeling techniques necessary for such investigations (e.g., polynomial regression, multitrait-multimethod models, latent interaction models; Jager et al., 2012; Trautwein et al., 2012; Laird and De Los Reyes, 2013).

Are There Differences in How Children's Competencies Are Perceived by Different Raters?

Findings suggested a relatively high level of agreement across reporters' ratings of children's competencies within the four domains assessed in the current study. The correlational findings coincided with our results examining difference scores, such that there were no significant overall differences among child, mother, father, and teacher reports of children's competencies. This documented agreement among reporters' ratings of children's competencies may be attributable to the observable nature

of children's abilities. For instance, homework assignments, instructor/coach feedback, test performance, and grades on report cards may provide children, parents, and teachers with a consistent sense of how children perform in a given domain.

However, testing for differences between dependent correlations revealed a more nuanced picture of cross-reporter perceptions of children's diverse competencies. Consistent with our hypotheses, as well as the extant literature (Schroeder et al., 2010; De Los Reyes et al., 2015), mothers and fathers tended to demonstrate the most agreement in perceptions of children's competencies. Generally higher levels of agreement were also noted between children and mothers, and mothers and teachers, as compared to other reporter dyads, perhaps reflecting mothers' stereotypically greater involvement in and knowledge of their children's performance and functioning (Grolnick and Slowiaczek, 1994). Additionally, agreement between teachers and other reporters tended to be lower in several domains, perhaps reflecting teachers' unique perspectives of children's academic functioning. Notably, we observed lower agreement between children and teachers, as compared to mothers and teachers and fathers and teachers, perhaps reflecting home-school communication between parents and teachers through parent–teacher conferences and report cards of which children in the elementary school years may be mostly unaware.

The mean differences on about half of the inter-rater comparisons also point to some systematic differences in how children, mothers, fathers, and teachers rate children's competencies, but they were not consistent across domains (as indicated by non-significant global *t*-tests). No differences were noted between mothers' and fathers', and children's and teachers', average ratings of child competency in any domain except sports. Furthermore, no difference was noted in average ratings of math and reading competencies between children and teachers, but mothers tended to rate these competencies higher than children and teachers. Despite child–teacher agreement in mean level, child–teacher correlations tended to be lower than child–mother and mother–teacher correlations. Hence, the rank-order alignment of children and teachers was less consistent (relative to child–mother and mother–teacher alignment), but their overall assessments of competence in the group were similar.

Whereas the bivariate correlations and overall *t*-tests pointed to relatively high levels of agreement across all four reporters, analyses comparing the correlation matrices provided a further differentiated picture of degrees of reporter agreement. Specifically, child-reported math-reading associations were lower than those according to both mother and father reports, suggesting that parents perceive stronger relations between children's competencies in these two domains than do children. Differences also emerged between teacher-reported associations and those reported by the other three reporters, such that all teacher-reported associations were higher than those reported by children and by parents. These strong correlations may be attributable to teachers having a unique perspective on children's school functioning, as they may see more consistencies in children's competencies across domains than do children or parents who are normally limited to seeing individual children's grades on report cards. This finding may also reflect the fact that teachers may form a more global view of children's competencies that generalizes across domains (Bornstein and Putnick, 2019), whereas children and parents may see more differentiated competencies. Specifically, core classroom teachers may have less information about children's performance in music and sports given that they do not generally teach those subjects to their students. Teachers may therefore not observe students in these settings, perhaps leading teachers to generalize or stereotype children's competencies in those areas based on children's performance in core academic subjects. These correlations may also be due to biases in teachers' perceptions of children's competencies (den Brok et al., 2004), as it may be that a teacher who views a student as struggling in one domain assumes that the student is having similar difficulties in other domains. It is therefore important to consider the influence of teachers' undifferentiated expectations and perceptions of children's competencies.

This set of findings has clear educational implications, suggesting that teacher input should be carefully considered when discussing children's academic functioning. The findings also suggest that many procedures already utilized in schools (e.g., midterm reports, parent–teacher conferences) provide parents with information regarding children's academic competencies, which may in turn decrease potential areas of conflict between parents and teachers. Likewise, it may be important to include children in some of these information procedures to provide children with additional perspectives about how they are functioning in the educational setting. Such discussions may provide children with developmentally appropriate feedback on strengths and weaknesses and motivation to improve in any needed areas. Incorporating parents' perceptions of their children's academic competencies in broader, standard educational assessments may also help calibrate expectations for children's academic performance, highlight academic strengths and weaknesses, and suggest potential educational interventions to support children's academic success.

Are There Differences in How Children's Competencies Are Perceived Across Different Domains?

Inter-domain correlations tended to be smaller and less consistent than inter-rater correlations, and significant inter-domain disagreement emerged in the difference scores of children's competencies of math, reading, music, and sports. Taken together, these results were consistent with our hypotheses, as well as the broader literature (Jacobs et al., 2002), that children would be perceived as more competent in core academic subjects (e.g., math and reading) than in extracurricular ones (e.g., music and sports). It is important to note consistent agreement between ratings of children's math and reading competencies (except according to child report), supporting the strong links between math and reading abilities as well

as the common cognitive skills underlying these competencies (Hart et al., 2009; Bornstein and Putnick, 2019). Additionally, ratings of children's music competence were lower than ratings for their math, reading, and sports competencies. This finding may reflect a general lack of exposure to and practice with musical instruments in elementary schools in the United States (Wigfield et al., 1997). The lower ratings of children's music competence may also reflect the fact that fewer questions assessing music competence are included in the CBTV, and therefore the music subscale may be missing some crucial aspects of children's competencies in this domain (e.g., confidence in abilities, expectations for performance; see **Appendix**).

Dependent correlation tests also supported stronger associations between math and reading as compared to associations between the other domains. Stronger correlations were also revealed between music and sports, which are frequently seen as extracurricular activities, as compared to the other core academic domains. Children also consistently perceived reading and music as more strongly correlated than the other domains, and this was the only significant inter-domain correlation according to children's reports (**Table 2**). Reading and music were often more highly related than other domains according to other reporters as well. This finding is supported by literature indicating a strong association between reading skills and the ability to discriminate musical sounds, as musical abilities often stem from well-developed phonemic awareness (Lamb and Gregory, 1993).

The bivariate analyses indicated differences among the four domains, but analyses comparing inter-domain correlation matrices highlighted more similarities, except between reading and music. Specifically, and consistent with the bivariate findings discussed above, associations were stronger for reading than for music competence, suggesting more consistency among reporters' ratings of children's reading competence than their music competence. This difference may again be attributable to the lack of exposure to music in elementary school (Wigfield et al., 1997) or to measurement differences.

It may not be surprising that more differences than similarities were noted in perceptions of children's competencies across various domains, as children naturally demonstrate patterns of strengths and weaknesses (Bornstein and Putnick, 2019). However, the fact that more agreement was seen in core academic subjects as compared to extracurricular subjects suggests that children might require more support in and exposure to these areas. Agreement levels may also indicate that additional means of communicating about children's competencies in special academic areas like music and sports is needed. Perhaps music teachers and sport coaches should be included in parent-teacher conferences, or other means of providing feedback to children and parents in these areas seems warranted. Taken together, the findings from the current study point to consistency in the strengths of the inter-rater associations among the four domains assessed and more variation in the inter-domain associations across the four reporters. Our results also highlight the need to consider additional methodological techniques (e.g., actual grades) to fully assess the degree of individual differences in reports of children's academic competencies.

Strengths, Limitations, and Future Directions

The current study has several strengths, including a multi-reporter and multi-domain design to fully evaluate differences in reporters' perceptions of children's elementary school competencies in several arenas, domains which to date have not been examined in the literature on reporter agreement. We also extended our analyses beyond the bivariate level, thereby providing a methodological contribution to this literature. However, we also note several limitations to the current study. Our sample is relatively sociodemographically diverse, but it is not representative of the entire range of ethnic backgrounds; therefore, these findings may not generalize to other samples (e.g., minority populations). Child development and parenting are known to vary with ethnicity (Bornstein and Lansford, 2010; Bornstein, 2015; Murry et al., 2015); by including only European American families, we intentionally avoided the ethnicity-socioeconomic status confound that has vexed the existing literature and would also cloud our findings with respect to children's competencies and parents' and teachers' perceptions (Bornstein et al., 2013; Jager et al., 2017). It is also not clear if the associations and disagreements documented in the current study would be observed in at-risk or clinical populations. For instance, differences in reports of children's competencies may be even lower for children with intellectual and cognitive disabilities, as their functioning may be less variable and more readily known to both parents and teachers due to increased academic testing and accommodations.

Additionally, our measure of children's competencies varied by the reporter and by the domain assessed in terms of the number and content of items (see **Appendix**). It is possible that the correlations obtained in the current study would be even stronger if the same numbers of items and similar contents had been used, as similarity in measurement may have reduced error and variation in ratings. It will be important for future studies to consider the impact of measurement differences on the degree of documented reporter disagreement. Furthermore, we cannot confirm or assume equal intervals between response options on the CBTV. Utilization of analytic techniques that relax this assumption, including many-facet Rasch models (which also evaluate severity across reporters; Myford and Wolfe, 2003), is warranted in future studies with the CBTV. We also did not have an objective measure of children's academic competencies in the current study. Instead, the current study focuses on patterns in perceptions of children's competencies across reporters and domains, providing a contribution to Eccles and colleagues' EVT model (Wigfield and Eccles, 2000) of children's functioning in educational settings.

Last, our study is cross-sectional as ratings of children's competencies were only included at one point in time. It is important to consider that agreement in reports of children's competencies may change over time, particularly given that

child and parent reports tend to vary greatly from each other. Research also suggests that ratings of children's competencies in math, reading, and sports decline across elementary school and then level off or even increase across the high school years (Jacobs et al., 2002). Consideration of these developmental differences is of paramount importance to understanding the extent and impact of reporter agreement and disagreement about children's competencies. Examination of perceptions of academic competencies during adolescence (i.e., when students are in high school) is a particularly important direction for future research, as any discrepancies in these perceptions across reporters and domains may have significant implications for adolescents' college readiness and career trajectories.

Implications

This study provided a novel look into perceptions of child competencies by children, mothers, fathers, and teachers. The mixed agreement found between reporters has implications for educational measurement, theory, and practice. Regarding measurement, does the high level of agreement among reporters suggest that different perspectives should be combined (as in a factor score) to make a more valid measure of child competence? Perhaps not. Mother and father reports were the most consistent in both mean level and rank order (although still not interchangeable), but we found several systematic differences between other reporters, suggesting that reporter differences are not simply a result of random error. Furthermore, recent work has shown that the unshared variance between reporters (which is usually relegated to the error term in a factor model) may be meaningfully predictive of child and dyad functioning (e.g., Jager et al., 2012). Consequently, the reporter of record should be chosen based on study goals.

Regarding theory, the moderate positive relations between perceptions of child math and reading competencies for mothers, fathers, and teachers, but not children, indicates that children may be more likely to use dimensional comparisons to inform their competence ratings about themselves, but adults (i.e., parents and teachers) ratings of children may not. Adults may be more likely to draw on a working model of general (*g*) intelligence that supports positive relations between math and reading competence (Furnham et al., 2002). A hierarchical model of intelligence, where various academic competencies load on a single general factor, has been supported in studies of child intelligence and performance (McGrew, 2009; Castejon et al., 2010; Bornstein and Putnick, 2019). Hence, adults may be more accurate reporters of child competencies than children themselves. Future research assessing the predictive validity of each reporters' competence ratings are needed to validate this deduction.

Finally, regarding educational practice, our study suggests that success is in the eye of the beholder. Mothers and fathers tend to agree about children's competencies across domains, but teachers and children have lower levels of agreement with one another as well as with parents. Reporters' ratings of competence also vary by domain. Teachers, administrators, and school counselors should be aware of these differences in ratings of children's competencies by reporter as well as domain. Assessing competence from multiple sources may provide the most well-rounded picture of child competence. Furthermore, ratings of children's academic and school competencies may develop into inherent expectations and biases, which may alter how individuals perceive children's long-term functioning in an academic setting. Such biased evaluations (both positive and negative, and by the self and others) are related to important long-term academic outcomes, including self-regulation, social bonding, and school achievement (Jussim and Harber, 2005; Leduc and Bouffard, 2017). Acknowledging that individual ratings of children's academic competencies are but one piece of data that needs to be verified or calibrated with other data sources may help prevent the development of any biases or false expectations in children's academic functioning.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

SR reviewed the literature. DP and MB supervised the testing. DP undertook the analysis. MB and DP conceived and designed the study. All the authors wrote and reviewed the manuscript. MB and GE commented and submitted the manuscript.

FUNDING

This research was supported by the Nanyang Assistant Professor Start-Up Grant, the Intramural Research Program of the NIH/NICHD, United States, and an International Research Fellowship at the Institute for Fiscal Studies (IFS), London, United Kingdom, funded by the European Research Council (ERC) under the Horizon 2020 Research and Innovation Programme (grant agreement no. 695300-HKADeC-ERC-2015-AdG).

ACKNOWLEDGMENTS

We thank the participating families and our collaborators and research assistants, Joan T. D. Suwalsky, Charlene Hendricks, Lauren Henry, and Hannah Simon.

REFERENCES

Bornstein, M. H. (2015). "Children's parents," in *Ecological settings and processes in Developmental Systems. Handbook of child Psychology and Developmental Science*, Vol 4, 7th Edn, eds R. M. Lerner, M. H. Bornstein, and T. Leventhal (Hoboken, NJ: Wiley), 55–132.

Bornstein, M. H., Jager, J., and Putnick, D. L. (2013). Sampling in developmental science: situations, shortcomings, solutions, and standards. *Devel. Rev.* 33, 357–370. doi: 10.1016/j.dr.2013.08.003

Bornstein, M. H., and Lansford, J. E. (2010). "Parenting," in *The Handbook of Cultural dEvelopmental Science. Part 1. Domains of Development Across cultures*, ed. M. H. Bornstein (New York, NY: Psychology Press), 259–277.

Bornstein, M. H., and Putnick, D. L. (2019). *The Architecture of the Child Mind: g, Fs, and their Hierarchy*. New York: Routledge.

Castejon, J. L., Perez, A. M., and Gilar, R. (2010). Confirmatory factor analysis of Project Spectrum activities. A second-order g factor or multiple intelligences? *Intelligence* 38, 481–496. doi: 10.1016/j.intell.2010.07.002

Creech, A., and Hallam, S. (2003). Parent-teacher-pupil interactions in instrumental music tuition: a literature review. *Br. J. Music Educ.* 20, 29–44. doi: 10.1017/s0265051702005272

Crozier, G. (1999). Is it a case of "We know when we're not wanted"? The parents' perspective on parent-teacher roles and relationships. *Educ. Res.* 41, 315–328. doi: 10.1080/0013188990410306

De Los Reyes, A. (2011). Introduction to the special section: more than measurement error: discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *J. Clin. Child Adolesc. Psychol.* 40, 1–9. doi: 10.1080/15374416.2011.533405

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., et al. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychol. Bull.* 141, 858–900. doi: 10.1037/a0038498

De Los Reyes, A., Ehrlich, K. B., Swan, A. J., Luo, T. J., Van Wie, M., and Pabón, S. C. (2013). An experimental test of whether informants can report about child and family behavior based on settings of behavioral expression. *J. Child Family Stud.* 22, 177–191. doi: 10.1007/s10826-012-9567-3

den Brok, P., Brekelmans, M., and Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *Sch. Eff. Sch. Improv.* 15, 407–442. doi: 10.1080/09243450512331383262

Eccles, J. S., and Wigfield, A. (1995). In the mind of the actor: the structure of adolescents' achievement and task values and expectancy-related beliefs. *Personal. Soc. Psychol. Bull.* 21, 215–225. doi: 10.1177/0146167295213003

Eccles, J. S., Wigfield, A., Harold, R., and Blumenfeld, P. (1993). Age and gender differences in children's achievement self-perceptions during the elementary school years. *Child Dev.* 64, 830–847. doi: 10.2307/1131221

Furnham, A., Rakow, T., and Mak, T. (2002). The determinants of parents' beliefs about the intelligence of their children: a study from Hong Kong. *Int. J. Psychol.* 37, 343–352. doi: 10.1080/00207590244000151

Grolnick, W. S., and Slowiaczek, M. L. (1994). Parents' involvement in children's schooling: a multidimensional conceptualization and motivational model. *Child Dev.* 65, 237–252. doi: 10.1111/j.1467-8624.1994.tb00747.x

Hart, S. A., Petrill, S. A., Thompson, L. A., and Plomin, R. (2009). The ABCs of math: a genetic analysis of mathematics and its links with reading ability and general cognitive ability. *J. Educ. Psychol.* 101, 388–402. doi: 10.1037/a0015115

Herbert, J., and Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Appl. Dev. Psychol.* 26, 276–295. doi: 10.1016/j.appdev.2005.02.007

Hollingshead, A. B. (1975). *Four Factor Index of Social Status*. New Haven, CT: Yale University.

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

Jager, J., Bornstein, M. H., Putnick, D. L., and Hendricks, C. (2012). Family members' unique perspectives of the family: examining their scope, size, and relations to individual adjustment. *J. Family Psychol.* 26, 400–410. doi: 10.1037/a0028330

Jager, J., Putnick, D. L., and Bornstein, M. H. (2017). II. More than just convenient: the scientific merits of homogeneous convenience samples. *Monogr. Soc. Res. Child Dev.* 82, 13–30. doi: 10.1111/mono.12296

Jussim, L., and Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: knowns and unknowns, resolved and unresolved controversies. *Personal. Soc. Psychol. Rev.* 9, 131–155. doi: 10.1207/s15327957pspr0902_3

Kerr, D. C. R., Lunkenheimer, E. S., and Olson, S. L. (2007). Assessment of child problem behaviors by multiple informants: a longitudinal study from preschool to school entry. *J. Child Psychol. Psychiatry* 48, 967–975. doi: 10.1111/j.1469-7610.2007.01776.x

Laird, R. D., and De Los Reyes, A. (2013). Testing information discrepancies as predictors of early adolescent psychopathology: why difference scores cannot tell you what you want to know and how polynomial regression may. *J. Abnorm. Child Psychol.* 41, 1–14. doi: 10.1007/s10802-012-9659-y

Lamb, S. J., and Gregory, A. H. (1993). The relationships between music and reading in beginning readers. *Educ. Psychol.* 13, 19–27. doi: 10.1080/0144341930130103

Lauermann, F., Tsai, Y.-M., and Eccles, J. S. (2017). Math-related career aspirations and choices with eccles et al.'s expectancy-value theory of achievement-related behaviors. *Dev. Psychol.* 53, 1540–1559. doi: 10.1037/dev0000367

Leduc, C., and Bouffard, T. (2017). The impact of biased self-evaluations of school and social competence on academic and social functioning. *Learn. Individ. Differ.* 55, 193–201. doi: 10.1016/j.lindif.2017.04.006

Lee, I. A., and Preacher, K. J. (2013). *Calculation for the Test of the Difference Between two Dependent Correlations with one Variable in Common [Computer Software]*. Available at: http://quantpsy.org (accessed May 25, 2017).

Little, R. J., and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*, 2nd Edn. New York: Wiley Interscience.

Marsh, H. W. (1986). Verbal and math self-concepts: an internal/external frame of reference model. *Am. Educ. Res. J.* 23, 129–149. doi: 10.3102/00028312023001129

Marsh, H. W., Kuyper, H., Seaton, M., Parker, P. D., Morin, A. J. S., Möller, J., et al. (2014). Dimensional comparison theory: an extension of the internal/external frame of reference effect on academic self-concept formation. *Contemp. Educ. Psychol.* 39, 326–341. doi: 10.1016/j.cedpsych.2014.08.003

McGrath, E. P., and Repetti, R. L. (2000). Mothers' and fathers' attitudes toward their children's academic performance and children's perceptions of their academic competence. *J. Youth Adolesc.* 29, 713–723. doi: 10.1023/A:1026460007421

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

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

Murry, V. M., Hill, N. E., Witherspoon, D., Berkel, C., and Bartz, D. (2015). "Children in diverse social contexts," in *Ecological Settings and Processes in Developmental systems. Volume 4 of the Handbook of child psychology and developmental science*, 7th Edn, eds R. M. Lerner, M. H. Bornstein, and T. Leventhal (Hoboken, NJ: Wiley), 416–454.

Muthén, B., and Muthén, L. (2014). *Mplus (version 7.2)*. Los Angeles, CA: Authors.

Myford, C. M., and Wolfe, E. W. (2003). Detecting and measuring rater effects using many-facet Rasch measurement: part I. *J. Appl. Meas.* 4, 386–422.

Rimm-Kaufman, S., Pianta, R. C., and Cox, M. (2001). Teachers' judgments of problems in the transition to school. *Early Child. Res. Q.* 15, 147–166. doi: 10.1016/S0885-2006(00)00049-1

Rosenthal, R. (1994). Interpersonal expectancy effects: a 30-year perspective. *Curr. Dir. Psychol. Sci.* 3, 176–179. doi: 10.1111/1467-8721.ep10770698

Salbach-Andrae, H., Lenz, K., and Lehmkuhl, U. (2009). Patterns of agreement among parent, teacher and youth ratings in a referred sample. *Eur. Psychiatry* 24, 345–351. doi: 10.1016/j.eurpsy.2008.07.008

Satorra, A., and Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* 66, 333–343. doi: 10.1007/BF02296192

Schroeder, J. F., Hood, M. M., and Hughes, H. M. (2010). Inter-parent agreement on syndrome scales of the child behavior checklist (CBCL): Correspondence and discrepancies. *J. Child and Family Stud.* 19, 646–653. doi: 10.1007/s10826-010-9352-0

Simpkins, S. D., Fredricks, J. A., and Eccles, J. S. (2012). Charting the Eccles' expectancy-value model from mothers' beliefs in childhood to youths' activities in adolescence. *Dev. Psychol.* 48, 1019–1032. doi: 10.1037/a0027468

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychol. Bull.* 87, 245–251. doi: 10.1037/0033-2909.87.2.245

Tomasetto, C., Mirisola, A., Galdi, S., and Cadinu, M. (2015). Parents' math-gender stereotypes, children's self-perception of ability, and children's appraisal of parents' evaluations in 6-year-olds. *Contemp. Educ. Psychol.* 42, 186–198. doi: 10.1016/j.cedpsych.2015.06.007

Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., and Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy-value theory: a Latent variable interaction modeling study. *J. Educ. Psychol.* 104, 763–777. doi: 10.1037/a0027470

Wentzel, K. R., Russell, S., and Baker, S. (2016). Emotional support and expectations from parents, teachers, and peers predict adolescent competence in school. *J. Educ. Psychol.* 108, 242–255. doi: 10.1037/edu0000049

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: a developmental perspective. *Educ. Psychol. Rev.* 6, 49–78. doi: 10.3389/fpsyg.2017.01193

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

Wigfield, A., Eccles, J. S., Mac Iver, D., Reuman, D. A., and Midgley, C. (1991). Transitions during early adolescence: changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Dev. Psychol.* 27, 552–565. doi: 10.1037/0012-1649.27.4.552

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., et al. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: a 3-year study. *J. Educ. Psychol.* 89, 451–469. doi: 10.1037/0022-0663.89.3.451

Yoon, M., and Kim, E. S. (2014). A comparison of sequential and nonsequential specification searches in testing factorial invariance. *Behav. Res. Methods* 46, 1199–1206. doi: 10.3758/s13428-013-0430-2

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.

Copyright © 2019 Racz, Putnick, Esposito and Bornstein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

TABLE A1 | Items to assess child, mother, father, and teacher perceptions of children's competencies in math, reading, music, and sports.

Child	Mother/Father	Teacher
Math		
(1) How good in math are you?	(1) How good is your child in math?	(1) Compared to other children, how hard does this child try in math?
(2) If you were to list all the students in your class from the worst to the best in math, where would you put yourself?	(2) In comparison to other children, how would you evaluate your child's performance in math?	(2) How well is this child performing in math compared to how well you believe s/he could?
(3) Compared to most of your other school subjects, how good are you in math?	(3) Compared to other children, how much innate ability or talent does this child have in math?	(3) Compared to other children, how much innate ability or talent does this child have in math?
(4) How well do you expect to do in math this year?	(4) How well do you think your child will do in math next year?	(4) How well do you expect this child to do next year in math?
(5) How good would you be at learning something new in math?		(5) Compared to other children, to what extent does this child give up when faced with a difficult problem in math?
(6) In general, how hard is math for you? (reversed)		
Reading		
(1) How good in reading are you?	(1) How good is your child in reading?	(1) Compared to other children, how hard does this child try in reading?
(2) If you were to list all the students in your class from worst to best in reading, where would you put yourself?	(2) In comparison to other children, how would you evaluate your child's performance in reading?	(2) How well is this child performing in reading compared to how well you believe s/he could?
(3) Compared to most of your other school subjects, how good are you in reading?	(3) Compared to other children, how much innate ability or talent does this child have in reading?	(3) Compared to other children, how much innate ability or talent does this child have in reading?
(4) How well do you expect to do in reading this year?	(4) How confident is your child in his/her ability to do well in reading?	(4) How well do you expect this child to do next year in reading?
(5) How good would you be at learning something new in reading?		(5) Compared to other children, to what extent does this child give up when faced with a difficult problem in reading?
(6) In general, how hard is reading for you? (reversed)		
Music		
(1) How good are you at music?	(1) How good is your child in music?	(1) Compared to other children, how hard does this child try in music?
(2) Compared to most of your other activities, how good would you be at playing a musical instrument?	(2) Compared to other children, how much innate ability or talent does this child have in music?	(2) Compared to other children, how much innate ability or talent does this child have in music?
(3) How good would you be at learning to play a new musical instrument?	(3) In comparison to other children, how would you evaluate your child's performance in music?	(3) Compared to other children, to what extent does this child give up when faced with a difficult problem in music?
(4) In general, how hard would learning to play a musical instrument be for you? (reversed)		
Sports		
(1) How good at sports are you?	(1) How good is your child in sports?	(1) Compared to other children, how hard does this child try in sports?
(2) If you were to list all the students in your class from the worst to the best in sports, where would you put yourself?	(3) Compared to other children, how much innate ability or talent does this child have in sports?	(2) Compared to other children, how much innate ability or talent does this child have in sports?
(3) Compared to most of your other activities, how good are you at sports?	(3) In comparison to other children, how would you evaluate your child's performance in sports?	(3) Compared to other children, to what extent does this child give up when faced with a difficult problem in sports?
(4) How well do you expect to do in your favorite sport this year?	(4) How confident is your child in his/her ability to do well in sports?	
(5) How good would you be at learning a new sport?		
(6) In general, how hard are sports for you? (reversed)		

frontiers in Education

ORIGINAL RESEARCH
published: 01 November 2019
doi: 10.3389/feduc.2019.00127

School Readiness Profiles and Growth in Academic Achievement

Qianqian Pan[1], Kim T. Trang[1], Hailey R. Love[2] and Jonathan Templin[3]*

[1] Department of Educational Psychology, University of Kansas, Lawrence, Kansas, KS, United States, [2] Department of Early Childhood, Multilingual, and Special Education, University of Nevada, Las Vegas, Las Vegas, NV, United States, [3] Department of Psychological and Quantitative Foundations, University of Iowa, Iowa City, IA, United States

The purpose of this research was to identify the presence of different school readiness profiles and to determine whether profiles could differentially predict academic growth. The Early Childhood Longitudinal Study: 2010-11 (ECLS-K: 2011) public data set was used, and participants were 14,954 first-time kindergarteners. The age of entering kindergarten ranged from 44.81 to 87.98 months with a mean of 76.13 months. In Study 1, a six-dimensional construct of school readiness was used: health, self-regulation, social and emotional development, language development, cognitive development, and approaches to learning. Results revealed 41 profiles with the top six school readiness profiles covering 85% of the sample: (1) Positive Development (28%); (2) Comprehensive At-Risk (24%); (3) Personal and Social Strengths (20%); (4) Cognitive and Language Strengths (5%), (5) Health Strength (5%); and (6) Cognitive, Personal and Social Strengths (3%). Study 2 examined whether school readiness profiles could predict children's reading and math achievement growth using growth curve models. Results showed that different school readiness profile membership had unique academic growth patterns and could predict academic growth above and beyond child and family background variables. Moreover, children with the Positive Development profile had higher academic achievement over time. Children with the Personal and Social Strengths profile had the largest growth rates. In sum, findings support the inclusion of self-regulation as another dimension of school readiness and the important role of personal and social skills in the development of reading and math achievement.

Keywords: school readiness, ECLS-K:2011, log-linear cognitive diagnostic models, growth curve models, academic achievement

OPEN ACCESS

Edited by:
Raquel Gilar,
University of Alicante, Spain

Reviewed by:
Angeliki Mouzaki,
University of Crete, Greece
Yong Luo,
Educational Testing Service,
United States

****Correspondence:***
Qianqian Pan
panqianqian2013@gmail.com

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

Received: *22 August 2019*
Accepted: *16 October 2019*
Published: *01 November 2019*

Citation:
Pan Q, Trang KT, Love HR and Templin J (2019) School Readiness Profiles and Growth in Academic Achievement. Front. Educ. 4:127. doi: 10.3389/feduc.2019.00127

INTRODUCTION

School readiness skills, including cognitive, social, attentional, and self-regulation skills, lay the foundation for future school success. Considerable research has demonstrated a link between kindergarten cognitive skills and later elementary school achievement (La Paro and Pianta, 2000; Bodovski and Farkas, 2007; Duncan et al., 2007; Claessens et al., 2009). Children entering kindergarten with stronger math and literacy skills tend to have higher math and reading achievement in later grades. There is also evidence that social skills, attention skills, and self-regulation skills are important predictors of academic and behavioral outcomes (La Paro and Pianta, 2000; Trentacosta and Izard, 2007; Claessens et al., 2009; Morgan et al., 2016). Furthermore, prosocial skills predict adult outcomes, such as high school graduation, college degree, and employment (Jones et al., 2015). Presumably, children who are able to regulate their emotions

and attention have an easier time attending to academic tasks, which then promotes academic competence. In sum, there is a large body of research evidence demonstrating the importance of both cognitive and personal and social school readiness skills for future school and adult outcomes.

However, not all children enter kindergarten with the necessary school readiness skills. Using data from the Early Childhood Longitudinal Study (ECLS), Kindergarten Class of 1998–99, Wertheimer et al. (2003) estimated that 56% of the nation's young children, or 2.2 million, showed challenges in at least one key area of development before entering kindergarten. More recent data from the ECLS-K:2010-11 illustrated differences in school readiness skills across racial/ethnic groups (Mulligan et al., 2012). Accordingly, many suggest that improving children's school readiness through early childhood education (ECE) is key to reducing the racial achievement differences seen later in education (Heckman, 2006; Duncan et al., 2007). Specifically, publicly-funded early education programs have shown benefits for children who come from socioeconomically-disadvantaged backgrounds and minority children, and these programs help reduce the achievement gap over time (Lee, 2002; Magnuson and Waldfogel, 2016). Moreover, studies have shown that participating in high-quality preschool programs and early interventions can improve all children's health and development (Anderson et al., 2003; Puma et al., 2010).

Yet, while a large body of research supports investment in ECE, there is great disagreement and ambiguity around the underlying theory and conceptualization of school readiness goals within ECE (Snow, 2006; Pretti-Frontczak, 2014). The concept of school readiness has specifically been challenged as often being too narrowly focused on certain literacy and math skills. It has been argued that this narrow focus has led to policies and practices that actually undermine ECE (Pretti-Frontczak, 2014) and have deleterious effects on children with possible special needs (Carlton and Winsler, 1999). These critiques beg the question, what do we mean when we say children need to be ready for school? What skills, knowledge, and abilities are involved in school readiness? To inform interventions or programs that target school readiness skills, we need to better understand the construct of school readiness and its influence on children's later achievement.

SCHOOL READINESS THEORY AND RESEARCH

School Readiness Theoretical Framework

School readiness can be defined generally as the skills, knowledge, and abilities that children need to succeed in formal schooling, which, for most, begins at kindergarten (Snow, 2006). In the last 40 years, research on school readiness has produced many different theories and perspectives (Snow, 2006; Winter and Kelley, 2008). From a maturational perspective, readiness to learn depends on the child's skills and cognitive maturity level (Kagan, 1992). However, more recent perspectives have shifted to a holistic, multidimensional definition of school readiness, emphasizing the importance of personal and social skills and the roles of families and communities (e.g., Diamond, 2010). That is, children need to be ready for school, but schools and communities also need to be ready to support children's future success across multiple developmental domains (Elizabeth Graue, 1992; Pretti-Frontczak, 2014).

Although there is no consensus on an operational definition for school readiness, most researchers rely on the five domains developed by the National Education Goals Panel: health and physical development; emotional well-being and social competence; approaches to learning; communicative skills; and cognition and general knowledge (National Education Goals Panel, 1991). Recent frameworks continue to rely on these general domains and have expanded on them (e.g., U.S. Department of Health Human Services, 2015; Altun, 2018). *Health and physical development* refers to children's health and motor development that support engagement and learning in their environments. *Emotional well-being and social competence* refers to the development of key social skills and attitudes that help build and maintain positive relationships with others. *Approaches to learning* refers to children's attitudes, habits, and learning styles that characterize how they learn. *Communicative skills* refers to language and literacy skills that promote effective communication with others. Lastly, *cognition and general knowledge* refers to ways of thinking and acquiring knowledge that promotes learning. Each domain is a unique aspect of school readiness that needs to be measured and investigated (National Education Goals Panel, 1991).

In recent years, researchers have emphasized the importance of self-regulation in conceptualizing school readiness (Blair, 2002; Blair and Raver, 2015). Self-regulation is used broadly to describe aspects of emotion and behavior regulation that relies in part on development in the prefrontal cortex (Blair and Raver, 2015). In some frameworks, self-regulation is combined with approaches to learning as a general domain (e.g., the Early Childhood Learning and Knowledge Center). However, we argue that the former set of skills can be distinguished from the latter. Positive learning attitudes and beliefs (e.g., growth-mindset, motivation, creativity) are important predictors of child outcomes (Wigfield and Eccles, 2000; Grant and Dweck, 2003). But, these learning attitudes and beliefs are conceptually different from the ability to self-regulate one's emotions, behaviors, and attention in a way that is responsive to specific tasks and demands, which primarily connotes effortful control of cognition and behavior (Liew, 2012). Children can have positive learning orientations but still lack the ability to follow through, inhibit impulses, and focus on achieving goals. In this study, we were interested in investigating the unique contribution of self-regulation processes to children's school readiness ability, and how self-regulation interacts with other domains of school readiness.

Self-regulation offers an important addition to the conceptualization of school readiness because it addresses children's ability to attend to information, use it appropriately, and inhibit behavior that interferes with learning. However, like the broader concept of school readiness, theories and perspectives on self-regulation have focused on various

priorities. Within the fields of early childhood and elementary education, three domains of self-regulation are most consistently studied: attentional flexibility, inhibitory control, and working memory (Blair, 2002; McClelland and Cameron, 2011; Lerner, 2015). *Attentional flexibility* is the ability to focus and shift attention. *Working memory* is the ability to work on and actively process information. *Inhibitory control* is the ability to inhibit prepotent responses and activate adaptive responses. Indeed, numerous studies have demonstrated positive social and academic outcomes for children who possess these abilities (e.g., Eiesnberg et al., 1997; Gilliom et al., 2002; Trentacosta and Izard, 2007; Welsh et al., 2010; Curby et al., 2015). For example, Trentacosta and Izard (2007) found that emotion regulation abilities were positively correlated with academic competence. In addition, Curby et al. (2015) found that preschooler's emotion regulation predicted their preliteracy skills. Children who can regulate their emotions may be more able to pay attention to academic tasks and therefore perform better academically. Thus, significant work suggests that school readiness outcomes are dependent on self-regulation abilities.

School Readiness and Academic Trajectories

Understanding the skills children need to support their early learning is important because children's academic trajectories are associated with the skills they have upon kindergarten entry. Previous studies found both main and interaction effects for school readiness skills on children's academic achievement in later grades. A large body of research demonstrates that kindergarten entry cognitive skills (e.g., literacy and mathematics skills) were positively related to later academic performance (McCoach et al., 2006; Duncan et al., 2007; Li-Grining et al., 2010). In contrast, findings on the predictive contributions of children's personal and social skills have been mixed. Some studies suggested no effects of social skills on standardized achievement outcomes (Duncan et al., 2007; Claessens et al., 2009; Romano et al., 2010). However, other studies provided evidence that self-control, approaches to learning (e.g., task persistence, attention), and executive functions contributed to standardized math and reading scores (Bodovski and Farkas, 2007; DiPerna et al., 2007; Li-Grining et al., 2010). Children's approaches to learning (when conceptualized as including self-control, persistence, and attentiveness) were also associated with linear increases in math and reading from kindergarten to fifth grade (Li-Grining et al., 2010).

Furthermore, beyond the main effects of personal and social skills on later academic achievement, there is evidence that personal and social skills interact with cognitive skills to predict later achievement. For example, Cooper et al. (2014) found that among children with low reading skills in kindergarten, those with higher social skills were more likely to have higher reading scores in fifth grade compared to those with lower social skills. Studies that classified children into different school readiness profiles based on the possession of certain cognitive and social skills found that school readiness skills interacted in distinct patterns to predict divergent achievement outcomes (Hair et al., 2006; Halle et al., 2012). Although they used different statistical methods, both Halle et al. (2012) and Hair et al. (2006) found that preschool children can be classified into four distinct profiles based on their strengths and challenges in the social-emotional, cognitive, approaches to learning, and health domains (e.g., profiles included cognitive strength and socio-emotional risk). Importantly, Hair et al. (2006) found that the profiles differentially predicted academic and social outcomes in early elementary school. Children who were classified as having a comprehensive profile (above average for health, socioemotional, language, cognition) were more likely to be rated as having better approaches to learning, self-control, and general health in first grade; they also scored better on standardized math and reading tests. Meanwhile, children in health risk and socio-emotional risk profiles scored the lowest on various outcomes measures (Hair et al., 2006). Overall, previous research provides evidence for school readiness as a multi-dimensional domain and indicates that school readiness profiles can differentially predict child outcomes. In this study, we extend the previous work to include self-regulation as a separate domain of school readiness skills to investigate its unique contributions within school readiness profiles.

Statistical Methods in School Readiness Studies

Traditionally, researchers used linear regression to demonstrate the main effects of school readiness skills on later academic performance (e.g., Duncan et al., 2007). However, as more school readiness skills are added to models, there could be interaction effects between school readiness skills (e.g., a model with four school readiness skills could result in 6 two-way, 3 three-way, and 1 four-way interaction effects) that would be difficult to interpret using linear regression. Therefore, recently, researchers started applying Latent Class Analysis (LCA; e.g., McCutcheon, 1987) in school readiness studies. LCA allows children to be classified into different school readiness profiles, where each profile illustrates a unique pattern of school readiness skills, accounting for the possible interaction effects among these skills. For example, Halle et al. (2012) used LCA to classify children into four profiles: cognitive strength, cognitive risk, approaches to learning strength, and socio-emotional risk. Similarly, Hair et al. (2006) identified four different school readiness profiles (i.e., comprehensive positive development, social/emotional and health strengths, social/emotional risk, and health risk) in the ECKS-K:1998 sample based on a five-dimensional construct of school readiness (physical well-being and motor development, social and emotional development, language development, cognitive and general knowledge, and approaches to learning). Results showed that children with comprehensive positive profiles (strengths in cognitive, health, social/emotional development) had better first-grade math and reading scores compared to those who were classified as at risk in any one area via OLS regression.

However, there is a need for research that uses more advanced statistical methods to identify both the main effects of school readiness skills and interactions between skills. For

example, when using LCA to classify children into school readiness profiles, researchers applied several models with different numbers of profiles and compared the model fit indices [e.g., Bayesian information criteria (BIC), likelihood ratio test (LMR-LRT), and bootstrap likelihood ratio test (BLRT)], and the model with the best model fit statistics would be selected as the final model. Then, researchers named profiles by reviewing the patterns of profiles. Therefore, results depended heavily on model fit statistics and samples, which might be subjective and arbitrary. The use of such methods could pose challenges in deciding the number of school readiness profiles, labeling classes, and making inference across studies (Abenavoli et al., 2017). Also, the OLS regression used in previous studies cannot demonstrate the relationships between the school readiness profiles and children's academic trajectories, in terms of the initial status and growth rates (e.g., Hair et al., 2006).

CURRENT STUDY

Building from previous research, the present study aimed to extend current knowledge by (1) conceptualizing school readiness as a multi-dimensional construct that includes self-regulation skills in addition to the five previous dimensions that have been used (health, socioemotional development, language development, cognitive development, and approaches to learning), (2) applying Diagnostic Classification Models (DCMs; Rupp et al., 2010) to classify children into different school readiness profiles, and (3) adopting growth curve models (GCMs; e.g., Hoffman, 2015) to investigate the association between school readiness profile memberships and academic growth, above and beyond background variables.

This research was divided into two studies. Study 1 investigated the school readiness profiles of kindergarteners in the ECLS-K:2011 sample by adopting a six-dimensional construct of school readiness via DCMs. Study 2 investigated how school readiness profiles were associated with children's later academic achievement growth by fitting GCMs.

Study 1

Study 1 addressed the following research question: Using six dimensions of school readiness, what school readiness profiles exist among first-time kindergarteners in the ECLS-K:2011 cohort?

Method

Dataset

The current study used data from the Early Childhood Longitudinal Study-Kindergarten Cohort of 2010-2011 (ECLS-K:2011). The ECLS-K:2011 used a multistage probability sample design to select a nationally representative sample of U.S. children who attended kindergarten during the 2010-2011 school year. A total of six waves of data were released to the public (K-fall, K-spring, Grade 1-fall, Grade 1-spring, Grade 2-fall, and Grade 2-spring) when the current study was conducted. The ECLS-K:2011 provides nationally representative data on children's development, learning and performance at school. Background variables included family, school, and community characteristics, which provided opportunities to investigate the relations among these variables and children's development. More details about this database can be found in the user's manual of ECLS-K:2011 (Tourangeau et al., 2009).

Sample

In Study 1, only the first wave, fall of 2010 kindergarten data was used. The participants were limited to first-time kindergarteners to focus on children's status upon entering formal schooling, reflecting the current investigations' focus on school readiness. Also, children who were one of a set of twins were excluded to remove the potential of dependency within families. A total of 14,954 first-time kindergarteners were included in the data analysis (7,330 females, 7,591 males, and 33 with gender missing). The age of kindergarten entry ranged from 44.81 months to 87.98 months with a mean of 76.13 months. The racial/ethnic distribution of the sample was White (49.0%), African-American (12.8%), Asian (7.5%), Hispanic (24.5%), Others and Multi-Racial (6.1%) and missing (0.2%).

Measures

Table 1 shows the six dimensions of school readiness. The construct, variables, and the re-coding rules designed for the current study was based on previous research (i.e., Hair et al., 2006) and the authors' conceptual knowledge. See Appendix A in **Supplementary Material** for more details on the variables used.

Health. There were four indicators of health. Parents reported on the child's overall health using a scale from 1 to 4. Healthy weight was determined using guidelines from the Center for Disease Control and Prevention (e.g., 5th percentile to less than the 85th percentile) (National Center for Health Statistics, 2000). Low birth weight was defined as <5.5 pounds at birth. Premature was defined as more than two weeks before the due date.

Self-regulation. There were five indicators of self-regulation, which were five scale scores including cognitive flexibility scores measured by Dimensional Change Card Sort tasks (Zelazo, 2006), working memory scores measured by Numbers Reversed tasks (Woodcock et al., 2001), attentional focus scores and inhibitory control scores measured by Children's Behavior Questionnaire (Putnam and Rothbart, 2006), and self-control scores measured by Social Skills Rating System (Gresham and Elliott, 1990). Higher scores indicated higher ability in this area.

Social and emotional development. There were four indicators of social and emotional development. The scales were teacher report and were adapted from the Social Skills Rating Systems, measured on a scale from 1 (Never) to 4 (Very Often). Interpersonal skills measured children's ability to relate and interact with others. Externalizing problem behavior measured children's acting out behaviors. Internalizing problem behavior measured the presence of anxiety, loneliness, low self-esteem, and sadness. Finally, impulsive/overactive measured the presence of child behavior that was considered sudden or excessive given a certain situation.

TABLE 1 | The construct of school readiness.

Dimension	Item Number	
Health		
	1	Overall health
	2	Health weight
	3	Low birth weight
	4	Premature
Self-regulation		
	5	Dimensional change card sort
	6	Numbers reversed
	7	Attentional focus
	8	Inhibitory control
	9	Self-control
Social and Emotional Development		
	10	Interpersonal skills
	11	Externalizing problems
	12	Internalizing problems
	13	Impulsive/overactive
Language and Literacy Development		
	14	Reading achievement
	15	Language and literacy (story)
	16	Language and literacy (letters)
	17	Language and literacy (read)
	18	Language and literacy (writing)
	19	Language and literacy (print)
Cognition and General Knowledge		
	20	Mathematics achievement
	21	Mathematical thinking (sort)
	22	Mathematical thinking (order)
	23	Mathematical thinking (relationship)
	24	Science (observe)
	25	Science (classifies)
	26	Science (life science)
Approaches to Learning		
	27	Eagerness to learn
	28	Adaptable
	29	Persistence
	30	Attention
	31	Creativity

Language and literacy development. There were six indicators of language and literacy development. Reading achievement scores were Item Response Theory (IRT)-scaled scores from an individually-administered standardized reading assessment. The assessment measured language use and literacy skills and was developed specifically for the ECLS-K study. Five items from the teacher-reported Academic Rating Scale–Language and literacy were also included. The scale included assessments of children's story comprehension, letter identification, reading, early writing behaviors, and print knowledge. Items were assessed using a scale from 1 (Not yet) to (Proficient).

Cognition and general knowledge. There were seven indicators of cognition and general knowledge. Mathematics achievement scores were IRT-scaled scores from an individually administered standardized mathematics assessment developed for the ECLS-K study. The assessment measured skills in conceptual and procedural knowledge and problem-solving in specific content areas (e.g., number sense, properties, and operations). Three items from the teacher-reported Academic Rating Scale–Mathematical were also included: sorting, ordering, and quantity relationships. Also, three items from the teacher-reported Academic Rating Scale–Science scale were used: observation skills, living and non-living things classification, and understanding of life science concepts. Items were measured on a scale from 1 (Not yet) to (Proficient).

Approaches to learning. There were five indicators of children's approaches to learning. Four teacher-report items measured how often children were: eager to learn, adaptable, persistent, and paid attention. One parent-report item measured children's creativity in work or play. All items were measured on a scale from 1 (Never) to 4 (Very often).

(Note: Correlations among all items used in the analyses can be found in Appendix A in **Supplementary Material**).

Statistical analysis

Diagnostic Classification Models (DCM; Rupp et al., 2010) was applied to classify children into different school readiness profiles, which provide some advantages over LCA used in the previous studies. Compare to LCA as an exploratory analysis model, DCMs are confirmatory latent class models, which can (1) provide an individual mastery status of each latent variable (mastery/non-mastery), often called the latent attribute in the DCM literature, and then (2) classify individuals into pre-determined latent profiles.

Each latent profile illustrates a distinct pattern of mastery status for all latent attributes. For example, suppose a total of 2 binary attributes (A_1, A_2) are measured, then, each individual will have two mastery statuses for each measured attribute (1 = mastery; 0 = non-mastery) and four possible latent profiles: A_1 = (0,0), A_2 = (1,0), A_3 = (0,1), and A_4 = (1,1). A_1 represents non-mastery for all attributes; A_2 represents mastery on Attribute 1 and non-mastery on Attribute 2; A_3 represents mastery on Attribute 2 and non-mastery on Attribute 1, and A_4 represents mastery on both Attribute 1 and Attribute 2. In general, when A binary attributes are estimated, a total of 2^A possible latent profiles could be possible. Therefore, researchers know the number of latent profiles and the meaning of each latent profiles *a priori*. In contrast, the number and meaning of latent profiles provided by LCA were decided after conducting data analyses. Furthermore, model-based classifications in DCMs are more objective and relatively independently; results and interpretation could be compared across studies. Therefore, we chose to use DCM over LCA for these advantages.

A total of six school readiness skills were evaluated in the current study to either be on-track (mastery status) or be off-track (non-mastery status). Thus, there were up to $\mathbf{2^6 = 64}$ distinct school readiness profiles for six binary attributes, and we would

know the pattern and meaning of each school readiness profile. Let 0 and 1 represent off-track and on-track for each subdomain. For example, pattern $\mathbf{A_r} = (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0})$ indicates that child $\mathbf{r}$ is off-track for all attributes, and pattern $\mathbf{A_{r'}} = (\mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0})$ indicates child $\mathbf{r}'$ is on-track for the first two attributes but off-track for other attributes.

Data analyses proceed in two steps: First, the alignment between assessment items and school readiness attributes, also called Q-matrix, was specified by the second and third author of the current study by reviewing previous studies and items provided in the ECLS-K:2011 data set, which can be found in Appendix A in **Supplementary Material**.

Second, the Log-linear Cognitive Diagnosis Model (LCDM; Henson et al., 2009), the most general DCM, was applied to (1) determine the mastery status of attributes of individuals, and (2) classify individuals into different school readiness profiles. To achieve these two objectives, first, the item responsibility given by the school readiness profile membership was estimated through Equation (1). Second, the school readiness profile membership probabilities for all possible profiles were estimated through Equation (2) for individuals.

Last, the final school readiness profile for individuals was determined by using the maximum a posteriori (MAP) estimate, which was the largest probability of school readiness profile membership.

$$P(X_r = x_r) = \sum_{c=1}^{C} \nu_c \prod_{i=1}^{I} \pi_{ic}^{x_{ir}} (1 - \pi_{ic})^{1-x_{ir}} \quad (1)$$

where, x_r represents a vector of item responses from individual r, π_{ic} represents the certain item response probability for item i given school readiness profile c. So, Equation (1) expresses the probability of observing a vector of item response X_r of an individual is a function of the probability of observing a certain item response and the probability of being in the school readiness profile.

$$\sum_{c=1}^{C} \nu_c = 1 \quad (2)$$

where, ν_c represents the probability of being school readiness c. Since each child is a member of one and only one school readiness profile. Such that, all school readiness profile probabilities are sum up to 1.

Both the item-level fit and test level-fit were evaluated in the current study. Posterior predictive model checking (PPMC, e.g., Rubin, 1984; Meng, 1994; Gelman et al., 1996) was used to assess the item fit. Results from the LCDM were used to simulate a new data set and then generate model-implied correlation coefficients between paired items. Then, the model fit was evaluated by inspecting the discrepancy between model-implied and data-implied correlation coefficients between paired items. Smaller discrepancy indicated better model data fit. In the current study, 0.15 was set as the cut-off value. Therefore, the absolute discrepancy $\leq$0.15 indicated acceptable model-data fit for a pair of items. The mean absolute difference for the item-pair correlations statistic (MADcor, DiBello et al., 2006) was the difference between the data-implied and the model-implied item correlation. For the sake of page limits, more details of estimation and model fit information of the LCDM can be found in Appendix B in **Supplementary Material**.

The data analysis was carried out using Mplus version 7.4 (Muthén and Muthén, 1998–2015) via maximum likelihood estimation.

Results

Model fit

Table 2 presents the proportion of fit and unfit pairs of items for each type of correlation coefficient. The results found 72% of pairs of items showed acceptable model-data fit based on our criteria. Regarding the test fit, MADcor is 0.053. Previous researchers suggested a MADcor value of 0.06 acceptable for the LCDM (e.g., Henson et al., 2009; Lei and Li, 2016). Therefore, the LCDM achieved acceptable model fit and it was plausible to interpret the results from the current LCDM (see Appendix C for more details of item parameter estimates and Appendix D for more details of model fit results in **Supplementary Material**).

Attribute classification

Results showed that the majority of the sample (85.17%) were classified into six school readiness profiles; 23 profiles had zero children, indicating no child showed these patterns of attributes; and 35 profiles had <3% of the sample, indicating these school readiness profiles were less likely to occur. A full description of all possible 64 attribute classifications can be found in Appendix E in **Supplementary Material**. Details of attribute reliability could be found in Appendix F in **Supplementary Material**.

Table 3 shows the proportion of the sample assigned to the top six profiles: (1) Positive Development profile included 28% of children who were on-track for all attributes, except health; (2) Comprehensive At-Risk profile included 24% of children who were off-track for all attributes; (3) Personal and Social Strengths profile included 20% of children who were off-track for health, language development, and cognitive development and on-track for self-regulation, social and emotional development, and approaches to learning; (4) Cognitive and Language Strengths profile included 5% of children who were on-track for language development and cognitive development but off-track for other attributes; (5) Health Strength profile included 5% of children who were only on-track for health; and (6) Cognitive, Personal and Social Strengths profile included 3% of children who were on-track for self-regulation, social and emotional development, cognitive development and approaches to learning but off-track for health and language development.

Discussion

Findings showed that three profiles represented 71.60% of the sample and the top six profiles represented 85.17% of the sample. Other than the inclusion of self-regulation, the top six profiles were conceptually similar to those found in previous studies that classified children as being on- or off-track for school readiness domains (e.g., Hair et al., 2006; Halle et al., 2012). These results indicate that personal and social skills

TABLE 2 | Summary for LCDM model fit statistics.

	N of pairs of categorical items (%)	N of pairs of continuous items	N of pairs of categorical and continuous items	Total pairs of items
Fit[a]	207(75.00%)	10(47.62%)	119(7.83%)	336 (72.26%)
Unfit[b]	69(25.00%)	11(52.38%)	49(29.17%)	129(27.74%)

[a]*Fit: the absolute discrepancy between model-implied and data-implied correlation coefficients between paired items ≤0.15.*
[b]*Unfit: the absolute discrepancy between model-implied and data-implied correlation coefficients between paired items >0.15.*

TABLE 3 | Attribute classification results for the top six profiles.

Profile	HEA	SR	SED	LAN	COG	APL	N	Proportion
Positive development	0	1	1	1	1	1	4214	28.18%
Comprehensive at-risk	0	0	0	0	0	0	3566	23.85%
Personal and social strengths	0	1	1	0	0	1	2927	19.57%
Cognitive and language strengths	0	0	0	1	1	0	793	5.30%
Health strength	1	0	0	0	0	0	722	4.83%
Cognitive, personal, and social strengths	0	1	1	0	1	1	514	3.44%

HEA, health; SR, self-regulation; SED, social and emotional development; LAN, language development; COG, cognitive development; APL, approaches to learning. An entry of 0 means off-track on the attribute; an entry of 1 means on-track on the attribute.

appeared to cluster together. That is, children who were on-track for self-regulation were on-track for other personal and social skills while children who were off-track for self-regulation were off-track for other personal and social skills. Cognitive abilities were similarly clustered together. This clustering pattern provides some evidence that self-regulation could be considered a personal and social dimension of school readiness and reflects skills that operate similarly compared to other personal and social skills. Though, it can be considered its own dimension because there were profiles that only included self-regulation skills and those that included self-regulation skills with different combinations of school readiness skills (e.g., a profile with self-regulation mastery and cognitive mastery) (see Appendix E in **Supplementary Material** for a list of all possible profiles from the study).

Study 2

Study 2 addressed the following research question: How does school readiness profile membership predict growth in reading and math achievement from kindergarten to grade 2, after controlling for child demographic and background variables?

Method

Sample

Study 2 used a total of four waves of data[1] from the sample used in Study 1. **Table 4** presents the descriptions of samples of Study 2.

Measures

IRT scores of reading and math achievement assessments were used as outcomes in Study 2. In the ECLS-K:2011 data set, assessments were vertically linked to make it a longitudinal measure of growth in achievement. However, scores for different subject areas were not comparable to each other because of different numbers of questions and content. A set of children and family background variables were used as control variables, including ethnicity, children gender, children disability status, family poverty status, parent education level, single parent household, and mom's age at first birth. More details of these measures could be found in the user's manual for the ECLS-K:2011 (Tourangeau et al., 2009).

Statistical analysis

The Growth Curve Model (GCM) was used to analyze children's academic growth in reading and math achievement across time. After inspecting the growth trajectory for each subject across time (see **Figure 1**), both reading and math achievement showed linear growth trajectories across time, on average[2]. Therefore, a two-level linear growth model was adopted in the current study. A total of four waves of data, including kindergarten–fall, kindergarten–spring, grade 1–spring and grade 2–spring[3] were used in the data analyses. At Level 1, individual's test scores were predicted by the length of his/her receiving formal education (in months). Also, a random intercept and a random slope of the time variable (the length of time in formal education) were assumed, meaning that each child could have his/her own initial

[1]kindergarten–fall, kindergarten–spring, grade 1–spring and grade 2–spring. Due to the sample design, only one-third of the original samples were selected in grade 1 – fall and grade 2 – fall, so, these two waves of data were not included in the GCMs.

[2]The average test scores were the saturated means, which was calculate by using Full Information Maximum Likelihood (FIML) to reduce the impacts of missing data at grade 1 – fall and grade 2 – fall.

[3]A total of six waves of data were available in the ECLS: K-2011 database, including kindergarten – fall, kindergarten – spring, grade 1 – fall, grade 1 – spring, grade 2 – fall, and grade 2 – spring. However, only a subsample (40%) of the total sample was selected for grade 1 – fall and grade 2 – fall data collection by design. So, only four waves of data were used in the current study to avoid the impacts of missing data.

TABLE 4 | Descriptive statistics of children demographic and background variables (N = 14,954).

Background variables[a]		
Categorical variables	***N***	**%**
Ethnicity		
White	7,331	49.02
Black	1,907	12.75
Hispanic	3,659	24.47
Asian	1,115	7.46
Others	908	6.07
Missing	34	0.23
Gender		
Male	7,591	5.76
Female	7,330	49.02
Missing	33	0.22
Disability Status		
Students without disability	8,613	57.60
Students with disability	2,111	14.12
Missing	4,230	28.29
Poverty Status		
Not poverty	8,597	57.49
Poverty	2,716	18.16
Missing	3,641	24.35
Parent Education Level		
High school	2,940	19.66
Middle school or lower	1,736	11.61
College	4,447	29.74
Bachelor degree	2,731	18.26
Master degree or higher	1,748	11.69
Single Parent Household		
Not a single parent household	9,650	64.53
Single parent household	2,953	19.75
Missing	2,351	15.72
Teenage Mom		
Not a teenage mom	9,105	6.89
Teenage mom	3,041	2.34
Missing	2,808	18.78
Continuous variables	**Mean**	**SD**
Age at entering the kindergarten	67.17	4.16
Household income	1.71	5.57
Parent occupational prestige	44.84	12.03
Outcome variables[b]		
Reading—Kindergarten fall	46.91	11.54
Reading—Kindergarten spring	61.40	13.44
Reading—Grade 1 spring	84.59	15.50
Reading—Grate 2 spring	96.63	12.03
Math—Kindergarten fall	31.71	11.40
Math—Kindergarten spring	45.31	12.18
Math—Grade 1 spring	67.15	15.26
Math—Grate 2 spring	81.40	13.58

[a] *Background variables were control variables in GCMs, including both categorical and continuous variables.*
[b] *Outcome variables were outcome variables in GCMs.*

level of the achievement at kindergarten entry as well as his/her own growth rates (Equation 3). At Level 2, a set of child and family background variables, as well as the school readiness profiles, were used as the predictors of Level 1 intercept and slope to investigate if school readiness profiles were associated with individual's academic growth above and beyond background variables (Equation 4 and 5). More details of GCMs can be found in Appendix G in **Supplementary Material**.

$$Y_{it} = \beta_{0i} + \beta_{1i} Time_{ti} + \epsilon_{it} \tag{3}$$

$$\beta_{0i} = \gamma_{00} + \sum_{k=1}^{K} \gamma_{0k} X_{ki} + u_{0i} \tag{4}$$

$$\beta_{1i} = \gamma_{10} + \sum_{k=1}^{K} \gamma_{1k} X_{ki} + u_{1i} \tag{5}$$

In Equation (3), Y_{it}, represents the test score for child i at time t, which can be expressed as a linear combination of a random intercept, β_{0i}, which represents each child had his/her own initial starting point, and the product of a random slope, β_{1i}, which represents each child had his/her own growth rates, and timing variable of child i, $Time_{ti}$, which is the length of receiving formal education in months, and the time specific error, ϵ_{it}.

In Equations (4,5), a set of K time-invariant variables predicted the random intercept and random slope, which included child background and demographic variables and school readiness profile membership. γ_{00} is the Level 2 intercept for Level 1 intercept; $\gamma_{01} - \gamma_{0k}$ are coefficients for time-invariant variables for the Level 1 intercept, representing the effects the time-invariant variables on the between-person variation in the intercept; and u_{oi} is the Level 2 residual for Level 1 intercept. γ_{10} is the Level 2 intercept for the Level 1 slope; $\gamma_{11} - \gamma_{1k}$ are coefficients for time-invariant variables for the Level 1 slope, representing the effects the time-invariant variables on the between-person variation in the slope; and u_{1i} is the Level 2 residual for Level 1 slope.

The total R^2, the squared correlation between the observed test scores and the test scores predicted by the model fixed effects was calculated to represent the variance explained by the time-invariant predictors. The pseudo-R^2 value for the proportion reduction in each random effect variance was calculated to evaluate the effect size of adding school readiness profile membership into the model. Additionally, Log-likelihood values were used to evaluate the relative model fit. Smaller values indicate better model fit. Residual maximum likelihood (REML) was used in estimating and reporting all model parameters. Denominator degrees of freedom was estimated by using the Satterthwaite method (Satterthwaite, 1946). The significance of fixed effects was evaluated with univariate Wald tests. The GCM analysis was generated using SAS Studio via PROC MIXED (SAS Institution, 2016).

Unrepresentative samples and missing data

To provide national-level estimates, the current study used one sampling weight variable (W6C6P_20) provided by ECLS-K:2011 to account for the unrepresentativeness and missing

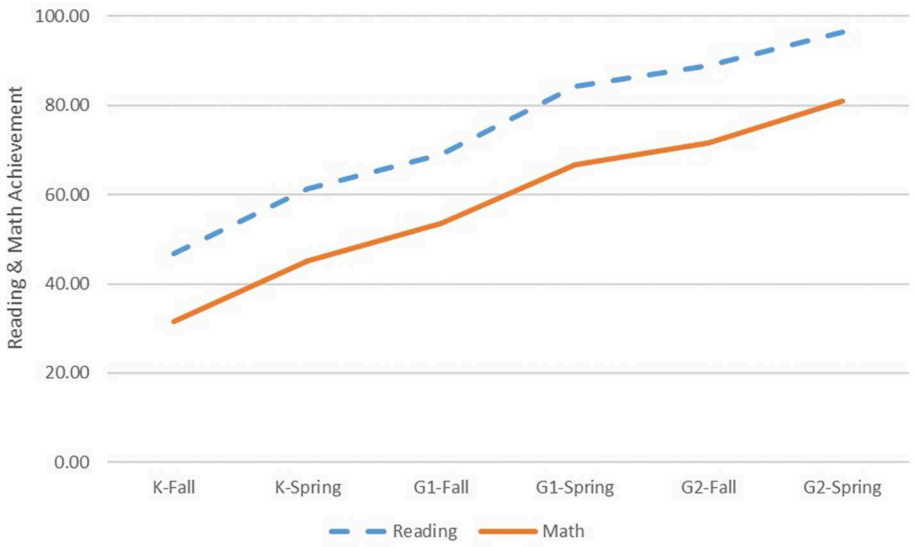


FIGURE 1 | Data-implied average reading and math growth trajectory across time.

data (Bernstein et al., 2014). This sampling weight variable adjusted for nonresponse associated with child assessment data from both kindergarten rounds, spring first grade and spring second grade, as well as parent data from fall kindergarten or spring kindergarten.

Results

Descriptive statistics

The descriptive statistics for child test scores, background, and demographic variables are presented in **Table 4**. For categorical variables, the first category was treated as the reference group in the GCMs. Continuous background and demographic variables were centered at the mean before entering into the GCMs.

Growth model results

Reading achievement growth Children's reading achievement scores across four measurement occasions were predicted from a set of child background and demographic variables as the baseline model (Reading Model 1). As shown in **Table 6**, 73.76% of the total variance in the reading achievement was explained by including the length of education, and child background and demographic variables. **Table 6** presents the estimated coefficients, where the intercept and slope parameters represent the reading achievement and growth rates for the average child who was in all reference groups. Moreover, there was a negative covariance between the intercept and the slope, which indicates that a higher initial level of reading achievement was associated with a slower growth rate.

For Reading Model 1, in terms of time-invariant predictors, some factors were related with higher initial level of reading achievement, including being an Asian child, chronologically older children at kindergarten entry, females, higher family income, higher parent occupation prestige, and higher parental education level. Also, the following variables were associated with lower initial levels of achievement: disability status, poverty status, having a teenage mother, and living in a single-parent household. Results revealed that there was a negative relationship between intercept variance and slope variance, indicating that a higher initial level of reading achievement was associated with a lower growth rate. In addition, results found two types of factors that were related to the lower growth rate: (1) sociodemographic factors, including race (i.e., Black and Hispanic children had a lower growth rate), having a disability, and family income under the poverty threshold; and (2) factors associated with high initial level of reading achievement (i.e., Asian children, chronologically older children at kindergarten entry, and having parents with a higher education level).

In Reading Model 2, as shown in **Table 5**, school readiness profile membership was added as an additional predictor of the intercept and the linear slope. Profile membership was treated a dummy variable and the Comprehensive At-Risk profile was the reference group. Results found that the total cumulative R^2 from Model 2 is $R^2 = 77.64\%$, approximately a 3.9% increase due to the addition of school readiness profile membership. In terms of the pseudo-R^2, school readiness profile membership accounted for 26.00% of random intercept variance, 6.57% of random slope variance, and 0.84% of the residual variance. In addition, smaller negative Log-likelihood values in the Model 2 indicate that Model 2 fit the data better than Model 1. Intercept, slope, and coefficients of other time-invariant predictor estimates are presented in **Table 6**. Results showed all other profiles except for Health Strength profile had significantly higher initial reading level compared to Comprehensive At-Risk profile. Regarding to the growth rates, comparing to the reference profile, Positive Development profile, Cognitive and Language Strengths, and Health Strength profile had significantly lower growth rates, however, Personal and Social Strengths profile and Cognitive,

Personal and Social Strengths profile had significantly higher growth rates.

Figure 2 shows the data-implied growth trajectories for the top six school readiness profiles[4], which illustrates the growth trajectories of reading achievement of the top six school readiness profiles. Children who had a Positive Development profile had higher initial reading levels and maintained that status over time. Children who were off-track for some cognitive attributes but on-track for personal and social attributes (Personal and Social Strengths) had lower initial reading achievement levels but eventually caught up to the reading achievement of children who were on-track for cognitive skills and language and literacy by the end of second grade (Cognitive and Language Strengths). Children who had less on-track attributes (Comprehensive At-Risk and Health Strength profiles) had a lower initial level and maintained that status over time.

Math achievement growth

Similar to the analysis conducted for reading achievement, math achievement across four measurement occasions were first predicted from a set of child background and demographic variables as the baseline model (Math Model 1). **Table 5** shows that 76.00% of the total variance in the math achievement was explained by including the education time and child background and demographic variables. Compare to the development of reading, there was a positive relationship between intercept variance and slope variance, indicating that a higher initial level of math achievement was associated with a higher growth rate.

For Math Model 1, regarding time-invariant predictors, results found several factors that were associated with a lower initial level of math achievement compared to the reference group: race (i.e., Black and Hispanic children had a lower initial level of math achievement), having a disability, having a parent with lower than a high school education, being from a single-parent household, and having a teenage mother. Other factors were found to be related with the higher initial level of math achievement: being an Asian child, high family income, having a parent with a college education and above, and having a parent with higher occupational prestige. Some factors were associated with a lower growth rate, including being Black or Hispanic, having a disability, family income under the poverty threshold, and single-parent household status. Also, females had lower growth rates even though females and males had the same initial level upon kindergarten entry. Similar to the results from the reading growth models, older age at kindergarten entry was also related to a higher initial math achievement level and a lower growth rate of math achievement.

In Math Model 2, school readiness profile membership was added as an additional predictor for intercept and slope. Profile membership was treated a dummy variable, and the Comprehensive At-Risk profile was the reference group. As shown in **Table 5**, a total of 79.45% of the total variance was explained, indicating a 3.50% increase in variance explained due to the addition of school readiness profile membership. Regarding the pseudo-R^2, school readiness profile membership

[4]The average observed test scores.

TABLE 5 | Summary of variance components estimates and model fit statistics for growth models.

	Reading model 1				Reading model 2					Math model 1				Math Model 2				
Variance Components	**Est.**	**SE**	***z*** [a]	***p*** [b]	**Est.**	**SE**	***z*** [a]	***p*** [b]	**R^2** [c]	**Est.**	**SE**	***z*** [a]	***p*** [b]	**Est.**	**SE**	***z*** [a]	***p*** [b]	**R^2** [c]
Intercept	86.45	2.50	34.55	<0.01	63.97	2.04	31.42	<0.01	0.26	67.65	1.89	35.83	<0.01	5.47	1.54	32.88	<0.01	0.25
Slope	0.03	<0.01	12.74	<0.01	0.03	<0.01	12.22	<0.01	0.07	0.03	<0.01	15.91	<0.01	0.03	<0.01	15.90	<0.01	0.01
Covariance [d]	−0.41	0.06	−6.75	<0.01	−0.28	0.05	−5.25	<0.01		0.30	0.04	7.20	<0.01	0.30	0.04	7.89	<0.01	
Residual	19321	243.04	79.49	<0.01	19219	24.55	79.90	<0.01	0.01	14034	175.22	8.09	<0.01	13963	173.48	8.49	<0.01	0.01
Total R^2							0.78			0.76				0.79				
Total R^2 change							Total R^2 change											
								0.04									0.03	
Model Fit Statistics																		
−2LL	161246				159842					155955		154657						

[a]*z is plausible z-value.* [b]*p is the plausible p-value.* [c]*R^2 is pseudo R^2.* [d]*Covariance: covariance between the intercept and the slope.*

TABLE 6 | Summary of fixed effect estimates in growth models.

	Reading model 1				Reading model 2				Math model 1				Math model 2			
	Est.	**SE**	***t*** [e]	***p*** [f]	**Est.**	**SE**	***t*** [e]	***p*** [f]	**Est.**	**SE**	***t*** [e]	***p*** [f]	**Est.**	**SE**	***t*** [e]	***p*** [f]
Fixed Effects																
Intercept	42.08	0.58	72.97	<0.01	38.76	0.58	67.25	<0.01	26.80	0.50	53.23	<0.01	23.37	0.50	46.45	<0.01
Ethnicity																
Black	1.03	0.54	1.92	0.06	0.72	0.49	1.48	0.14	−2.37	0.47	−5.04	<0.01	−2.54	0.42	−5.97	<0.01
Hispanic	−0.59	0.45	−1.31	0.19	−0.27	0.41	−0.65	0.51	−2.30	0.39	−5.83	<0.01	−1.90	0.36	−5.29	<0.01
Asian	5.08	0.77	6.57	<0.01	4.55	0.72	6.34	<0.01	2.55	0.67	3.80	<0.01	2.24	0.62	3.59	<0.01
Other	1.31	0.67	1.96	0.05	1.03	0.61	1.68	0.09	0.68	0.59	1.16	0.25	0.48	0.53	0.90	0.37
K_age [a]	0.52	0.04	13.61	<0.01	0.28	0.04	7.92	<0.01	0.70	0.03	21.05	<0.01	0.49	0.03	15.85	<0.01
Female	1.26	0.31	4.00	<0.01	0.30	0.29	1.01	0.31	−0.36	0.27	−1.33	0.18	−1.34	0.25	−5.27	<0.01
SWD [b]	−2.54	0.40	−6.33	<0.01	−1.21	0.37	−3.30	<0.01	−3.13	0.35	−8.94	<0.01	−1.89	0.32	−5.87	<0.01
Income [c]	0.16	0.05	3.15	<0.01	0.07	0.05	1.45	0.15	0.25	0.04	5.64	<0.01	0.16	0.04	4.09	<0.01
Poverty	−1.23	0.58	−2.13	0.03	−0.97	0.53	−1.85	0.06	−0.97	0.51	−1.92	0.05	−0.71	0.46	−1.55	0.12
Parent Education Level																
Middle School or Lower	−2.09	0.71	−2.94	<0.01	−1.64	0.65	−2.53	0.01	−1.94	0.62	−3.13	<0.01	−1.58	0.57	−2.79	0.01
College	1.27	0.47	2.73	0.01	0.91	0.42	2.15	0.03	1.15	0.41	2.81	<0.01	0.84	0.37	2.28	0.02
Bachelor	4.87	0.56	8.66	<0.01	3.58	0.51	6.97	<0.01	4.41	0.49	8.99	<0.01	3.29	0.45	7.35	<0.01
Master or higher	6.32	0.66	9.58	<0.01	4.70	0.60	7.81	<0.01	5.90	0.58	1.26	<0.01	4.57	0.53	8.69	<0.01
Occupation Prestige [d]	0.07	0.02	4.18	<0.01	0.05	0.01	3.29	<0.01	0.05	0.01	3.48	<0.01	0.03	0.01	2.50	0.01
Single Parent Household	−1.78	0.44	−4.07	<0.01	−1.33	0.40	−3.36	<0.01	−1.43	0.38	−3.73	<0.01	−0.98	0.35	−2.81	<0.01
Teenage Mom	−1.35	0.44	−3.07	<0.01	−0.93	0.40	−2.33	0.02	−1.06	0.38	−2.78	0.01	−0.77	0.35	−2.20	0.03
School readiness profile [g]																
Positive development					11.93	0.42	28.14	<0.01					1.83	0.37	29.25	<0.01
Personal and social strengths					3.38	0.46	7.36	<0.01					4.34	0.4	1.82	<0.01
Cognitive and language strengths					1.2	0.67	15.15	<0.01					8.15	0.59	13.88	<0.01
Health strength					−1.48	0.76	−1.95	0.05					−1.95	0.66	−2.95	<0.01
Cognitive, personal and social strengths					2.75	0.82	3.37	<0.01					5.29	0.71	7.42	<0.01
Slope	1.74	0.02	94.11	<0.01	1.73	0.02	85.59	<0.01	1.77	0.02	98.41	<0.01	1.77	0.02	98.41	<0.01
Ethnicity																

(Continued)

TABLE 6 | Continued

	Reading model 1				Reading model 2				Math model 1				Math model 2			
	Est.	**SE**	***t*** [e]	***p*** [f]	**Est.**	**SE**	***t*** [e]	***p*** [f]	**Est.**	**SE**	***t*** [e]	***p*** [f]	**Est.**	**SE**	***t*** [e]	***p*** [f]
Black	−0.09	0.02	−5.45	<0.01	−0.09	0.02	−5.42	<0.01	−0.18	0.01	−12.00	<0.01	−0.18	0.01	−12.07	<0.01
Hispanic	−0.03	0.01	−2.36	0.02	−0.03	0.01	−2.32	0.02	−0.06	0.01	−4.32	<0.01	−0.06	0.01	−4.32	<0.01
Asian	−0.10	0.03	−3.53	<0.01	−0.09	0.03	−3.28	<0.01	0.02	0.02	0.88	0.31	0.02	0.02	1.01	0.31
Other	−0.03	0.02	−1.51	0.13	−0.03	0.02	−1.40	0.16	−0.04	0.02	−2.05	0.05	−0.04	0.02	−1.96	0.05
K_age [a]	−0.01	0.00	−5.41	<0.01	−0.01	0.00	−4.15	<0.01	−0.01	0.00	−7.70	<0.01	−0.01	0.00	−7.31	<0.01
Female	0.04	0.01	3.55	<0.01	0.03	0.01	3.12	<0.01	−0.06	0.01	−6.11	<0.01	−0.06	0.01	−6.32	<0.01
SWD [b]	−0.05	0.01	−3.76	<0.01	−0.05	0.01	−3.73	<0.01	−0.05	0.01	−4.27	<0.01	−0.04	0.01	−3.74	<0.01
Income [c]	0.00	0.00	0.15	0.88	0.00	0.00	0.17	0.87	0.00	0.00	−0.65	0.51	0.00	0.00	−0.81	0.42
Poverty	−0.08	0.02	−4.48	<0.01	−0.08	0.02	−4.43	<0.01	−0.05	0.02	−3.15	<0.01	−0.05	0.02	−3.01	<0.01
Parent Education Level																
Middle School or Lower	−0.02	0.02	−0.90	0.37	−0.03	0.02	−1.23	0.22	0.01	0.02	0.37	0.71	0.01	0.02	0.26	0.79
College	0.01	0.01	0.82	0.41	0.01	0.01	1.00	0.32	0.02	0.01	1.67	0.10	0.02	0.01	1.59	0.11
Bachelor	−0.02	0.02	−1.04	0.30	−0.01	0.02	−0.56	0.58	−0.01	0.02	−0.86	0.39	−0.01	0.02	−0.79	0.43
Master or higher	−0.04	0.02	−2.01	0.04	−0.03	0.02	−1.40	0.16	−0.02	0.02	−1.14	0.26	−0.02	0.02	−1.04	0.30
Occupation prestige [d]	0.00	0.00	−0.73	0.47	0.00	0.00	−0.58	0.56	<0.01	<0.01	−0.49	0.62	0.00	0.00	−0.39	0.70
Single parent household	0.00	0.01	0.14	0.89	0.01	0.01	0.45	0.65	−0.03	0.01	−2.24	0.03	−0.02	0.01	−1.79	0.07
Teenage mom	0.01	0.01	0.68	0.50	0.01	0.01	0.42	0.68	−0.03	0.01	−2.24	0.03	0.00	0.01	0.00	1.00
School readiness profiles [g]																
Positive development					−0.05	0.02	−3.13	<0.01					<0.01	0.01	0.36	0.72
Personal and social strengths					0.08	0.02	5.1	<0.01					0.07	0.01	4.74	<0.01
Cognitive and language strengths					−0.09	0.02	−3.83	<0.01					<0.01	0.02	0.04	0.97
Health strength					−0.07	0.03	−2.55	0.01					−0.05	0.02	−2.26	0.02
Cognitive, personal and social strengths					0.06	0.03	2.18	0.03					0.07	0.03	2.64	0.01

*[a]K_age is Kindergarten Entry Age and centered at 60 months. [b]SWD is student with disability. [c]Income is family income and centered at 10. [d] Occupation Prestige is parent occupation prestige and centered at 45. [e]t is the plausible t value. [f]p is the plausible p value. [g]Profile of comprehensive at-risk was the reference group. All none-zero school readiness profiles were included in the Model 2, because of the page limits, regression coefficients of top 6 school readiness profiles were reported. A full description of regression coefficients for all none-zero school readiness profiles could be found in Appendix E (**Supplementary Material**).*

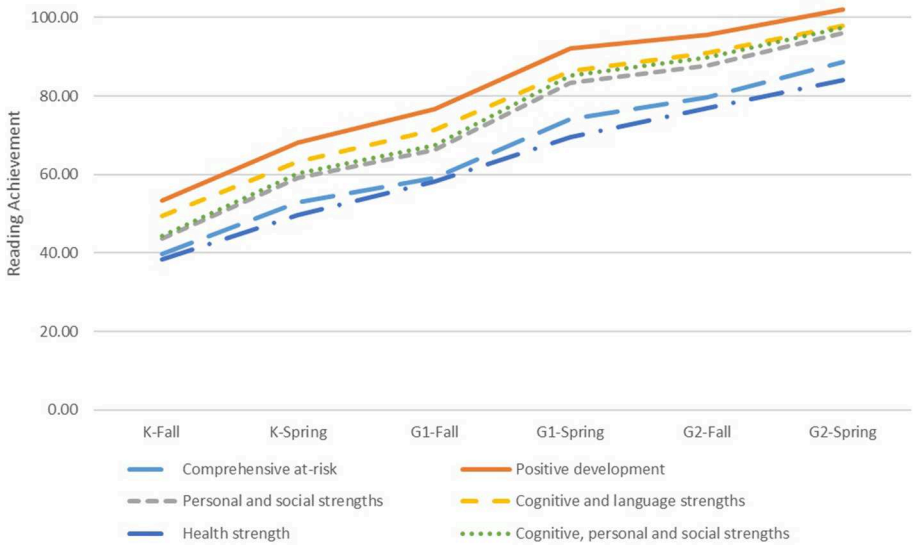


FIGURE 2 | Data implied reading growth trajectories across top 6 profiles.

accounted for 25.39% of random intercept variance, 1.11% of random slope variance, and 0.51% of the residual variance. Furthermore, smaller negative Log-likelihood values obtained from Math Model 2 indicate that Math Model 2 fit the data better than Math Model 1. **Table 6** shows the intercept, slope, and coefficients of other time-invariant predictor estimates of Math Model 2. Results showed Health Strength profile had significantly lower initial math level and all other profiles had significantly higher initial math level, compared to the Comprehensive At-Risk profile. Regarding the growth rate, comparing to the rerefence profile, Health Strength profile had significantly lower growth rates; in contrast, Personal and Social Strengths profile and Cognitive, Personal and Social Strengths profile had significantly higher growth rates.

As shown in **Figure 3**, similar results were found for the impact of school readiness profiles on the development of math achievement. Children in the Positive Development profile membership had higher initial achievement and maintained that status over time. At-risk profile membership (Comprehensive At-Risk and Health Strength profiles) was associated with lower initial achievement and maintained that status over time. However, children in the Personal and Social Strengths and Cognitive, Personal and Social Strengths profile caught up to their peers in the Cognitive and Language Strengths profile by the end of second grade, even though these children started behind their peers at kindergarten entry. The gap between children who were on-track for personal and social skills and children who were on-track for cognitive and language skills were closed over time as the former children demonstrated a higher growth rate compared to their peers starting around the spring semester of kindergarten.

Discussion

Results from Study 2 showed that school readiness profile membership could uniquely predict children's academic growth trajectories in both reading and math achievement, above and beyond child demographic and background variables. In other words, children's membership in different school readiness profiles could impact their academic growth. Furthermore, based on the data-implied growth trajectories for the top six profiles, children with the Positive Development profile entered kindergarten ahead and continued to perform higher than their peers, indicating the importance of children starting school with necessary school readiness skills. These findings were consistent with previous studies that show preschool cognitive skills could predict later academic achievement (e.g., Duncan et al., 2007; Pagani et al., 2010; Romano et al., 2010). However, it is notable that children who were on-track for personal and social attributes (e.g., self-regulation, social and emotional development, and approaches to learning) but off-track for cognitive attributes (e.g., language development, cognitive development) were able to catch up their peers by second grade in both reading and math achievement. Overall, these results provide evidence for the importance of personal and social skills in children's academic growth, as suggested by previous studies (e.g., Rimm-Kaufman et al., 2002; Bodovski and Farkas, 2007; DiPerna et al., 2007; Li-Grining et al., 2010). Additionally, findings support the inclusion of personal and social skills, including self-regulation, as components of school readiness that are important for children's continued academic achievement (Blair and Raver, 2015).

GENERAL DISCUSSION

The current study viewed school readiness as a six-dimensional construct, comprised of health, social and emotional development, language and literacy development, cognitive development, approaches to learning, and self-regulation. The inclusion of self-regulation in our conceptualization of

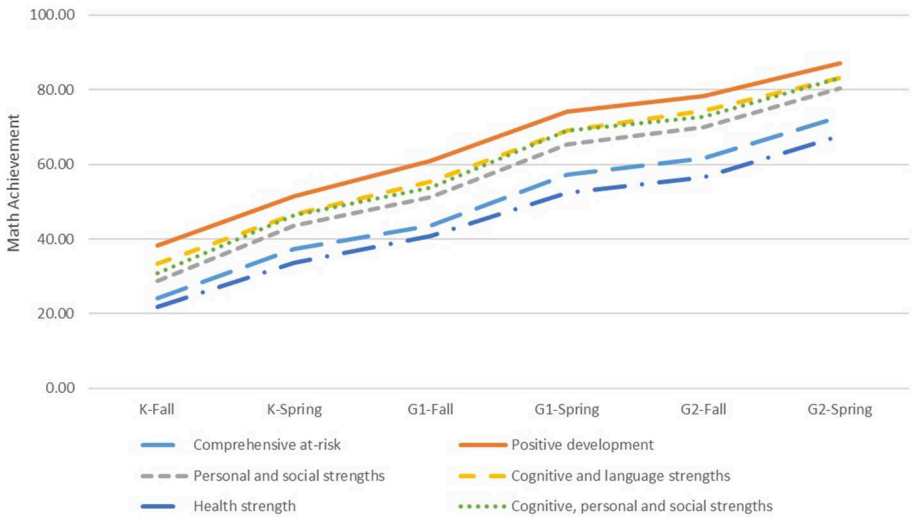


FIGURE 3 | Data implied math growth trajectories across top 6 profiles.

self-regulation was based on theory (e.g., Blair and Raver, 2015) and empirical evidence of its contributions to later achievement (e.g., Trentacosta and Izard, 2007; Welsh et al., 2010). This research extended previous work (Hair et al., 2006; Halle et al., 2012) by including self-regulation as a distinct domain in assessing the relations between children's school readiness and later academic achievement and using more recent, advanced statistical methods.

Study 1 applied Log-linear Cognitive Diagnostic Model (LCDM) to classify children into up to 64 possible pre-defined school readiness profiles. This approach overcame the limitations of general Latent Class Models used in previous studies because the LCDM classified children in a confirmatory model, such that the number of school readiness profiles and the label of each school readiness profile were known before the data analysis. Results showed that 85% of children were classified into the top six profiles: Positive Development (28%), Comprehensive At-Risk (24%), Personal and Social Strengths (20%), Cognitive and Language Strengths (5%), Health Strength (5%), and Cognitive, Personal and Social Strengths (3%). Other than the inclusion of self-regulation, the top six profiles were conceptually similar to those found in previous studies that classified children as being on- or off-track for school readiness domains (e.g., Hair et al., 2006; Halle et al., 2012).

Identifying school readiness profiles is important because they indicate the key developmental areas that children need support before entering kindergarten. Further, the profiles can be used to understand how school readiness attributes may group or interact with each other to inform practice. Importantly, these profiles allow a conceptualization of school readiness strengths and "risk" based on children's knowledge and skills, rather than purely based on familial and social backgrounds. In other words, these profiles allow a conceptualization of school readiness risk that focuses on attributes early childhood educators and interventionists can intervene on.

Based on this conceptualization, school readiness intervention programs and programs that supplement general high-quality ECE could be tailored to the school readiness profile that children belong to. That is, rather than delivering a general school readiness intervention or program to a group of children deemed "at-risk" due to their socioeconomic status or other family characteristics, educators and interventionists could focus on tailoring school readiness-focused instruction and intervention to distinct domains and/or profiles of school readiness. If high-quality ECE is conceived as a Tier 1, universal support for children's school readiness, interventions and embedded instruction tailored to children's school readiness profile could be added as a Tier 2 support for children identified as belonging to a profile other than the global Positive Development profile. This reflects a tiered system of support that has already shown benefits for children identified to have developmental delays or disabilities (e.g., Greenwood et al., 2011). Future research could explore the feasibility and effectiveness of such tailored intervention based on school readiness profiles.

Study 2 identified the unique contribution of school readiness profiles to academic growth, above and beyond demographic and family variables. In general, children who were well prepared in the cognitive attributes had the highest performance over time. It is noteworthy that children who were not well prepared in cognitive attributes but well prepared in personal and social attributes (e.g., self-regulation, social and emotional development, and approaches to learning) started off with lower reading and math achievement compared to children on-track for cognitive skills, but they closed the gap with their peers by second grade. Moreover, children who were not well prepared in both cognitive and personal and social attributes had the lowest initial levels and maintained that status through second grade. These findings suggest that personal and social skills, such as self-regulation, social and emotional development, and

approaches to learning, could help children with lower cognitive preparation catch up to their peers over time. These skills may be important because they help children attend to learning, regulate their emotions and behavior, learn and play with peers, and appropriately attend to and use new information. Previous studies suggested that self-regulation was positively related to the motivation and engagement for learning activities (e.g., Blair, 2002). Also, social and emotional competence could impact children's opportunities for learning by influencing the ways they interact with classroom adults and peers. Valiente et al. (2007, 2008) found that children with greater emotional regulation challenges were less likely to participate in class, were absent from school more often, and reported liking school less than their peers with greater emotional regulation. Thus, these growth trajectories indicate the importance of personal and social skills as contributors to school readiness and academic achievement.

Study 2 also indicated that self-regulation operates similarly compared to other personal and social attributes. We separated self-regulation as its own domain based on theory, and previous research indicates it does operate differently from other skills often grouped as approaches to learning. Specifically, research indicates that self-regulation supports academic achievement by reducing challenging behaviors that interfere with learning and improving interactions with other children (e.g., Montroy et al., 2014). Similarly, other research indicates that self-regulation may uniquely support attention and reasoning abilities (Blair et al., 2015). Thus, the growth trajectories identified in the present study in conjunction with the previous research indicate the significance of self-regulation as a distinct attribute that complements other personal and social skills to contribute to school readiness and academic achievement. Given the importance of personal and social skills for children's academic growth, particularly if a child is off-track on cognitive and language development, such skills should be treated as distinct abilities requiring specific instructional strategies, similar to the ways academic knowledge is divided into content areas. The present investigation represents one step toward further parsing out the specific personal and social skills that early childhood educators and interventionists can target to support children's school readiness.

In sum, results from Study 1 and Study 2 suggest that children can generally be characterized according to a six-dimensional conceptualization of school readiness that includes health, social and emotional development, language and literacy development, cognitive development, approaches to learning, and self-regulation. The present investigation also found that the particular combination of school readiness skills children possess upon kindergarten entry can impact their future growth and development, with personal and social skills allowing children to catch up if they start kindergarten behind their peers in cognitive and language development. The six-dimensional conceptualization of school readiness put forth by this study advances a more nuanced view of school readiness that accounts for the needs of the whole child rather than only academic or cognitive knowledge (Diamond, 2010). It is important that early childhood educators and other professionals are intentional in providing opportunities to develop children's school readiness skills, and defining these skills with more specificity can allow more targeted instruction and intervention. Early education programs should emphasize both cognitive, and personal and social skills as they prepare children for kindergarten as that dual focus could have lasting effects on children's academic achievement.

LIMITATIONS AND FUTURE STUDY

Although the ECLS-K:2011 data set is nationally representative, only a portion of the data was publicly available. The public data set provides scale scores for measures rather than the original item responses. Even though most measures had high reliabilities, we cannot exclude measurement error when applying the LCDM to classify children into different profiles. Future studies should analyze the item responses directly to get measurement-error-free estimates. Also, future studies can apply the same model to other cohorts of ECLS dataset to cross-validate the findings from the current study. The present study contributes to a growing body of literature arguing for the importance of self-regulation as a nuanced skill that significantly impacts children's academic achievement. Future research can build on these findings by continuing to explore the unique contributions of self-regulation to school readiness, including the specific mechanisms through which it impacts children's academic achievement. For example, future studies can look at different aspects of self-regulation (e.g., emotional and cognitive), and how they contribute to children's school readiness and academic achievement. Finally, future research could explore the feasibility and effectiveness of tailoring interventions to children's school readiness profiles. This might be done within the context of a response-to-intervention framework in which high-quality ECE is supplemented by targeted instruction based on children's school readiness profile. Considering the importance of both cognitive, and personal and social skills in children's academic achievement, early childhood educators should treat each skill as worthy of targeted support.

DATA AVAILABILITY STATEMENT

The datasets analyzed for this study can be found in the Early Childhood Longitudinal Program (ECLS-K:2011) (https://nces.ed.gov/ecls/datainformation2011.asp).

AUTHOR CONTRIBUTIONS

QP drafted the manuscript, conducted and interpreted the statistical analyses. KT drafted the literature review. KT and HL formed the measures and wrote the literature review and discussion. JT provided expertise on data analysis and feedback to the manuscript.

ACKNOWLEDGMENTS

We thank Dr. Lesa Hoffman for sharing her wisdom with us in specifying and interoperating the growth curve models, which improved the manuscript greatly.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2019.00127/full#supplementary-material

REFERENCES

Abenavoli, R. M., Greenberg, M. T., and Bierman, K. L. (2017). Identification and validation of school readiness profiles among high-risk kindergartners. *Early Childhood Res. Q.* 38, 33–43. doi: 10.1016/j.ecresq.2016.09.001

Altun, D. (2018). A paradigm shift in school readiness: a comparison of parents', pre-service and in-service preschool teachers' views. *Int. J. Prog. Edu.* 14, 37–56. doi: 10.29329/ijpe.2018.139.4

Anderson, L. M., Shinn, C., Fullilove, M. T., Scrimshaw, S. C., Fielding, J. E., Normand, J., et al. (2003). The effectiveness of early childhood development programs: a systematic review. *Am. J. Prevent. Med.* 24, 32–46. doi: 10.1016/S0749-3797(02)00655-4

Bernstein, S., West, J., Newsham, R., and Reid, M. (2014). *Kindergartners' Skills at School Entry: An Analysis of the ECLS-K*. Mathematica Policy Research.

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *Am. Psychol.* 57:111. doi: 10.1037/0003-066X.57.2.111

Blair, C., and Raver, C. C. (2015). School readiness and self-regulation: a developmental psychobiological approach. *Annu. Rev. Psychol.* 66, 711–731. doi: 10.1146/annurev-psych-010814-015221

Blair, C., Ursache, A., Greenberg, M., Vernon-Feagans, L., and Family Life Project, I. (2015). Multiple aspects of self-regulation uniquely predict mathematics but not letter-word knowledge in the early elementary grades. *Dev. Psychol.* 51, 459–472. doi: 10.1037/a0038813

Bodovski, K., and Farkas, G. (2007). Mathematics growth in early elementary school: the roles of beginning knowledge, student engagement, and instruction. *Element. Schl. J.* 108, 115–130. doi: 10.1086/525550

Carlton, M. P., and Winsler, A. (1999). School readiness: the need for a paradigm shift. *Schl. Psychol. Rev.* 28:338.

Claessens, A., Duncan, G., and Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Econ. Educ. Rev.* 28, 415–427. doi: 10.1016/j.econedurev.2008.09.003

Cooper, B. R., Moore, J. E., Powers, C., Cleveland, M., and Greenberg, M. T. (2014). Patterns of early reading and social skills associated with academic success in elementary school. *Early Educ. Dev.* 25, 1248–1264. doi: 10.1080/10409289.2014.932236

Curby, T. W., Brown, C. A., Bassett, H. H., and Denham, S. A. (2015). Associations between preschoolers' social–emotional competence and preliteracy skills. *Infant Child Dev.* 24, 549–570. doi: 10.1002/icd.1899

Diamond, A. (2010). The evidence base for improving school outcomes by addressing the whole child and by addressing skills and attitudes, not just content. *Early Educ. Dev.* 21, 780–793. doi: 10.1080/10409289.2010.514522

DiBello, L. V., Roussos, L. A., and Stout, W. (2006). "A review of cognitively diagnostic assessment and a summary of psychometric models," in *Handbook of Statistics*, Vol. 26, eds C. R. Rao and S. Sinharay (Elsevier), 979–1030. doi: 10.1016/S0169-7161(06)26031-0

DiPerna, J. C., Lei, P.-W., and Reid, E. E. (2007). Kindergarten predictors of mathematical growth in the primary grades: An investigation using the Early Childhood Longitudinal Study–Kindergarten cohort. *J. Educ. Psychol.* 99:369. doi: 10.1037/0022-0663.99.2.369

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Dev. Psychol.* 43:1428. doi: 10.1037/0012-1649.43.6.1428

Eiesnberg, N., Guthrie, I. K., Fabes, R. A., Reiser, M., Murphy, B., Holgren, R., et al. (1997). The relations of regulation and emotionality to resiliency and competent social functioning in elementary school children. *Child Dev.* 68, 295–311. doi: 10.1111/j.1467-8624.1997.tb01941.x

Elizabeth Graue, M. (1992). Social interpretations of readiness for kindergarten. *Early Childhood Res. Q.* 7, 225–243. doi: 10.1016/0885-2006(92)90006-K

Gelman, A., Meng, X.-L., and Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Stat. Sin.* 6, 733–760.

Gilliom, M., Shaw, D. S., Beck, J. E., Schonberg, M. A., and Lukon, J. L. (2002). Anger regulation in disadvantaged preschool boys: strategies, antecedents, and the development of self-control. *Dev. Psychol.* 38:222. doi: 10.1037/0012-1649.38.2.222

Grant, H., and Dweck, C. S. (2003). Clarifying achievement goals and their impact. *J. Pers. Soc. Psychol.* 85, 541–553. doi: 10.1037/0022-3514.85.3.541

Greenwood, C. R., Bradfield, T., Kaminski, R., Linas, M., Carta, J. J., and Nylander, D. (2011). The response to intervention (RTI) approach in early childhood. *Focus Except. Child.* 43, 1–22.

Gresham, F. M., and Elliott, S. N. (1990). *Social Skills Rating System (SSRS)*. American Guidance Service.

Hair, E., Halle, T., Terry-Humen, E., Lavelle, B., and Calkins, J. (2006). Children's school readiness in the ECLS-K: Predictions to academic, health, and social outcomes in first grade. *Early Childhood Res. Q.* 21, 431–454. doi: 10.1016/j.ecresq.2006.09.005

Halle, T. G., Hair, E. C., Wandner, L. D., and Chien, N. C. (2012). Profiles of school readiness among four-year-old Head Start children. *Early Childhood Res. Q.* 27, 613–626. doi: 10.1016/j.ecresq.2012.04.001

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science* 312, 1900–1902. doi: 10.1126/science.1128898

Henson, R. A., Templin, J. L., and Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika* 74, 191–210. doi: 10.1007/s11336-008-9089-5

Hoffman, L. (2015). *Longitudinal Analysis: Modeling Within-Person Fluctuation and Change*. New York, NY: Routledge.

Jones, D. E., Greenberg, M., and Crowley, M. (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. *J. Inform.* 105, 2283–2290. doi: 10.2105/AJPH.2015.302630

Kagan, S. L. (1992). Readiness past, present, and future: Shaping the agenda. *Young Child.* 48, 48–53.

La Paro, K. M., and Pianta, R. C. (2000). Predicting children's competence in the early school years: a meta-analytic review. *Rev. Educ. Res.* 70, 443–484. doi: 10.3102/00346543070004443

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educ. Res.* 31, 3–12. doi: 10.3102/0013189X031001003

Lei, P.-W., and Li, H. (2016). Performance of fit indices in choosing correct cognitive diagnostic models and Q-matrices. *Appl. Psychol. Meas.* 40, 405–417. doi: 10.1177/0146621616647954

Lerner, R. M. (2015). Promoting positive human development and social justice: Integrating theory, research and application in contemporary developmental science. *Int. J. Psychol.* 50, 165–173. doi: 10.1002/ijop.12162

Liew, J. (2012). Effortful control, executive functions, and education: bringing self-regulatory and social-emotional competencies to the table. *Child Dev. Perspect.* 6, 105–111. doi: 10.1111/j.1750-8606.2011.00196.x

Li-Grining, C. P., Votruba-Drzal, E., Maldonado-Carreño, C., and Haas, K. (2010). Children's early approaches to learning and academic trajectories through fifth grade. *Dev. Psychol.* 46:1062. doi: 10.1037/a0020066

Magnuson, K., and Waldfogel, J. (2016). Trends in income-related gaps in enrollment in early childhood education: 1968 to 2013. *AERA Open* 2, 1–13. doi: 10.1177/2332858416648933

McClelland, M. M., and Cameron, C. E. (2011). Self-regulation and academic achievement in elementary school children. *N. Direct. Child Adolesc. Dev.* 2011, 29–44. doi: 10.1002/cd.302

McCoach, D. B., O'connell, A. A., Reis, S. M., and Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. *J. Educ. Psychol.* 98:14. doi: 10.1037/0022-0663.98.1.14

McCutcheon, A. L. (1987). *Latent Class Analysis*. Thousand Oaks, CA: Sage. doi: 10.4135/9781412984713

Meng, X.-L. (1994). Posterior predictive p-values. *Ann. Stat.* 22, 1142–1160. doi: 10.1214/aos/1176325622

Montroy, J. J., Bowles, R. P., Skibbe, L. E., and Foster, T. D. (2014). Social skills and problem behaviors as mediators of the relationship between behavioral self-regulation and academic achievement. *Early Childhood Res. Q.* 29, 298–309. doi: 10.1016/j.ecresq.2014.03.002

Morgan, P. L., Farkas, G., Hillemeier, M. M., and Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educ. Res.* 45, 18–35. doi: 10.3102/0013189X16633182

Mulligan, G. M., Hastedt, S., and McCarroll, J. C. (2012). *First-Time Kindergartners in 2010-11: First Findings from the Kindergarten Rounds of the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K: 2011).* NCES 2012-049. National Center for Education Statistics.

Muthén, L. K., and Muthén, B. O. (1998–2015). Mplus User's Guide, 8th Edn. Los Angeles, CA: Muthén & Muthén.

National Center for Health Statistics (2000). *Clinical Growth Charts*. Retrieved from: https://www.cdc.gov/growthcharts/clinical_charts.htm

National Education Goals Panel (1991). *Goal 1 Technical Planning Group Report on School Readiness*. Washington, DC: NEGP. Retrieved from https://www2.ed.gov/pubs/AchGoal1/goal1.html

Pagani, L. S., Fitzpatrick, C., Archambault, I., and Janosz, M. (2010). School readiness and later achievement: A French Canadian replication and extension. *Dev. Psychol.* 46, 984–994. doi: 10.1037/a0018881

Pretti-Frontczak, K. (2014). Stop trying to make kids "ready" for kindergarten. *Young Except. Child.* 17, 51–53. doi: 10.1177/1096250614523346

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., et al. (2010). *Head Start Impact Study*. Final Report. Administration for Children & Families.

Putnam, S. P., and Rothbart, M. K. (2006). Development of short and very short forms of the children's behavior questionnaire. *J. Pers. Assess.* 87, 102–112. doi: 10.1207/s15327752jpa8701_09

Rimm-Kaufman, S. E., Early, D. M., Cox, M. J., Saluja, G., Pianta, R. C., Bradley, R. H., et al. (2002). Early behavioral attributes and teachers' sensitivity as predictors of competent behavior in the kindergarten classroom. *J. Appl. Dev. Psychol.* 23, 451–470. doi: 10.1016/S0193-3973(02)00128-4

Romano, E., Babchishin, L., Pagani, L. S., and Kohen, D. (2010). School readiness and later achievement: replication and extension using a nationwide Canadian survey. *Dev. Psychol.* 46:995. doi: 10.1037/a0018880

Rubin, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Stat.* 12, 1151–1172. doi: 10.1214/aos/1176346785

Rupp, A. A., Templin, J., and Henson, R. A. (2010). *Diagnostic Measurement: Theory, Methods, and Applications*. New York, NY: Guilford Press.

SAS Institution (2016). *SAS Studio 3.5 User's Guide*. Retrieved from http://support.sas.com/documentation/cdl/en/webeditorug/68828/PDF/default/webeditorug.pdf

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometr. Bull.* 2, 110–114. doi: 10.2307/3002019

Snow, K. L. (2006). Measuring school readiness: conceptual and practical considerations. *Early Educ. Dev.* 17, 7–41. doi: 10.1207/s15566935eed1701_2

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., and Najarian, M. (2009). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K): Combined User's Manual for the ECLS-K Eighth-Grade and K-8 Full Sample Data Files and Electronic Codebooks*. NCES 2009-004. National Center for Education Statistics.

Trentacosta, C. J., and Izard, C. E. (2007). Kindergarten children's emotion competence as a predictor of their academic competence in first grade. *Emotion* 7, 77–88. doi: 10.1037/1528-3542.7.1.77

U.S. Department of Health and Human Services (2015). *Supporting the School Readiness and Success of Young African American Boys Project: Reflections on a Culturally Responsive Strength-Based Approach*. Retrieved from: https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/young-african-american-boys-project-guide.pdf

Valiente, C., Lemery-Chalfant, K., and Castro, K. S. (2007). Children's effortful control and academic competence: mediation through school liking. *Merrill Palmer Q.* 53, 1–25. doi: 10.1353/mpq.2007.0006

Valiente, C., Lemery-Chalfant, K., Swanson, J., and Reiser, M. (2008). Prediction of children's academic competence from their effortful control, relationships, and classroom participation. *J. Educ. Psychol.* 100, 67–77. doi: 10.1037/0022-0663.100.1.67

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., and Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *J. Educ. Psychol.* 102, 43–53. doi: 10.1037/a0016738

Wertheimer, R. F., Moore, K. A., Hair, E. C., and Croan, T. (2003). *Attending Kindergarten and Already Behind: A Statistical Portrait of Vulnerable Young Children*. Washington, DC: Child Trends. doi: 10.1037/e317232004-001

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

Winter, S. M., and Kelley, M. F. (2008). Forty years of school readiness research: What have we learned? *Childhood Educ.* 84, 260–266. doi: 10.1080/00094056.2008.10523022

Woodcock, R. W., Mather, N., McGrew, K. S., and Wendling, B. J. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing Company.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. *Nat. Protoc.* 1:297. doi: 10.1038/nprot.2006.46

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.

Copyright © 2019 Pan, Trang, Love and Templin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 07 November 2019
doi: 10.3389/fpsyg.2019.02487

Relationship Between Digital Creativity, Parenting Style, and Adolescent Performance

María del Carmen Pérez-Fuentes[1,2], María del Mar Molero Jurado[1], Nieves Fátima Oropesa Ruiz[1], María del Mar Simón Márquez[1] and José Jesús Gázquez Linares[3]*

[1]Department of Psychology, University of Almería, Almería, Spain, [2]Universidad Politécnica y Artística del Paraguay, Asunción, Paraguay, [3]Department of Psychology, Universidad Autónoma de Chile, Santiago, Chile

Introduction: Today's adolescents live immersed in the digital world and are much more familiarized with the use of electronic devices. At the same time, the new technologies have become established as a powerful resource in teaching and learning, providing new texts where the limits of time and space are overcome. Digital creativity is part of people's daily lives and must be developed from the school and family context.

Objective: The objective of this study was to analyze the relationship among digital creativity, parenting style, and academic performance.

Method: This analysis was carried out in a sample of 742 adolescents in Middle School and High School aged 13–19. Digital creativity was evaluated using the Creative Behavior Questionnaire: Digital (CBQD). The Parenting Style Scale was used to evaluate the perception teenage boys and girls who have the various dimensions of their parents' educational style.

Results: Parenting styles were established as a mediating variable in the relationship between digital creativity and academic performance.

Conclusion: The roles of digital creativity, which is proposed as a facilitating tool in teaching, and parenting styles in academic performance for improving the family-school relationship are discussed.

Keywords: academic performance, adolescence, digital creativity, family-school relationship, parenting style

OPEN ACCESS

Edited by:
Raquel Gilar, University of Alicante, Spain

Reviewed by:
Francisco Manuel Morales Rodríguez, University of Granada, Spain
Pedro Rosário, University of Minho, Portugal

****Correspondence:***
María del Mar Molero Jurado
mmj130@ual.es

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

Received: *17 August 2019*
Accepted: *22 October 2019*
Published: *07 November 2019*

Citation:
Pérez-Fuentes MC, Molero Jurado MM, Oropesa Ruiz NF, Simón Márquez MM and Gázquez Linares JJ (2019) Relationship Between Digital Creativity, Parenting Style, and Adolescent Performance. Front. Psychol. 10:2487. doi: 10.3389/fpsyg.2019.02487

INTRODUCTION

The new technologies are considered a powerful resource in teaching and learning because they are providing new contexts where the limits of time and space are overcome. They facilitate collaboration, innovation, and creativity in individuals and organizations (Ala-Mutka et al., 2008). The quality of teaching today goes through recognition of the wide variety of stimuli and possibilities opened through the use of the new technologies. Youths, although they have beliefs about their level of competence in managing different virtual tools (García-Martín et al., 2014), have more knowledge in the use of social networks (Cabero-Almenara and Díaz, 2014).

What Is Digital Creativity?

Two different focuses on the study of creativity can be found in the literature. First, the more artistic or scientific side of creativity from the perspective of skills linked to music, dancing, painting, sculpture, literature, mathematics, and physics, where authors such as Amabile (1985), Gardner (1993), and Csikszentmihalyi (1996) and others have made enormous contributions. Second, studies on creativity from another perspective, a more everyday approach to creativity, do not require predominance of scientific skills or artistic abilities and are mainly the one in people's daily lives. The studies by Richards et al. (1988) and Runco (2004) should be mentioned here. In this approach, everyday creativity has been defined as self-expression in daily activities, interpersonal style, professional activities, and daily problem solving (Richards et al., 1988; Torrance, 1988). Both types of creativity are moderately related to each other, since similar psychological processes are involved in not only both, but also different, as mentioned by Ivcevic (2007). Thus, everyone is considered to have a potential for creativity, which can be manifested in many ways, due partly to the plasticity of our brains (Sun et al., 2016); creative thought having been empirically found to fluctuate at different times, increasing in adolescence (Claxton et al., 2005; Kim, 2011).

At the present time, we are witnessing huge digital transformations in a globalized world where technological changes occur in a practically inappreciable time. The new adolescents grow up immerged in the digital world and are much more familiarized with the use of electronic devices (cell phones, computers, tablets, and consoles). Following this reasoning, digital creativity is part of the daily life of today's youth and has been defined based on three main components: digital creative achievement, school-based everyday creativity, and self-expressive digital creativity (Hoffmann et al., 2016). Creative achievement actually refers to elements in which a person has achieved something with credibility. School-based everyday creativity usually includes elements of digital creativity, which the students may have done at school or for homework. Finally, self-expressive creativity is other creative efforts, which the students have made but have not usually received the same external validation as the creative achievement component. Thus, anyone can experience digital creativity. This study concentrates on school-based everyday creativity.

Some authors have suggested that the benefits of using the new technologies for teaching and learning depend on the learning approach used, teacher skills and coexistence in support settings for students and faculty (Ala-Mutka et al., 2008; Fernández-Batanero and Rodríguez-Martín, 2017), and being able to avoid maladjusted behavior that has been related to low academic performance (Pérez-Fuentes et al., 2011; Estévez et al., 2018). Concerning the focus on learning, some studies have demonstrated that if methodologies developing academic performance of the student (fluidity, flexibility, originality, and elaboration) are used from early ages, their academic performance improves (Guilford, 1950; Renzulli et al., 1986). Thus, we believe that creativity applied to the digital world can be a resource for teaching-learning that could increase academic performance of high school students. Based on this hypothesis, and keeping in mind that digital creativity is a novel subject, and therefore, little studied at the present time, one of the objectives of this study was to analyze school-based everyday creativity, as defined in the paragraph above, and its relationship with academic performance.

Dimensional Approach to Parenting Styles in Adolescent Development

Parenting style refers to how parents and their children act, behave, and relate to each other in any everyday situation. Since Baumrind (1968) distinguished among the authoritarian, permissive, and democratic parenting styles, much research has been done on its influence in the psychological adjustment of children and adolescents. Later, Maccoby and Martin (1983) differentiated a fourth parenting style, negligent parenting, and since then, numerous studies have tried to corroborate these findings from different perspectives and approaches.

The difference between dimensional and typological approaches in the study of parenting styles should also be mentioned. The dimensional approach, which is the subject of this study, and according to Oliva et al. (2007), is made up of six dimensions, which can explain the relationships of parents with their adolescent children and of children toward their parents: *affect* and communication (emotional nearness, support, harmony, and cohesion), *promotion of autonomy* (respect for the decisions of the minor through conversations where agreements are tolerated), *behavioral control* (set rules, limits, penalties, responsibilities, supervision, and monitoring of behavior), *psychological control* (lack of respect for individuality, intrusive control, manipulation, induction of guilt, emotional blackmail, or withdrawing affect), *self-disclosure* (spontaneous communication by the adolescent about activities, friends, and partner), and *humor* (relaxed aptitude, cheerful, and optimistic). Oliva et al. (2008) described three parenting styles according to the dimensions above: democratic parents (with high levels in all the dimensions mentioned except psychological control, medium levels in behavioral control), strict (strong behavioral, psychological, and affective control), strict (strong behavioral, psychological, and affective control), and indifferent (strong psychological control and weak in rest of dimensions).

There is quite good agreement among researchers on parenting styles that affect and communication have a primary role in adolescent adjustment (Oliva et al., 2008; Fuentes et al., 2015; Riquelme et al., 2018). Pelegrina et al. (2002) stated that affect and communication by parents are an indispensable condition for achieving adequate behavior in their children, optimum self-esteem and self-confidence. Psychological control, on the other hand, is considered independent of promoting autonomy (Silk et al., 2003), although related (Hauser-Kunz and Grych, 2013), especially when promotion of autonomy is understood as promotion of volitional functioning (Soenens and Vansteenkiste, 2010). Meanwhile, psychological control is associated negatively with psychological adjustment (Barber et al., 2012; Cui et al., 2014), in particular, among adolescents who have difficulties in regulating their emotions (Cui et al., 2014) and need psychological intervention (Ho et al., 2018). Although behavioral control by parents in adolescence usually prevents externalizing problems (Silk et al., 2003; Parra and Oliva, 2006; Ruiz-Hernández et al., 2018), the

many possible combinations between the dimensions characterizing parenting styles mentioned above and the sensitivity of parents toward the adolescent's own characteristics must also be kept in mind. In this direction, Oliva et al. (2008) found that when behavior control combines with self-disclosure, adolescent adjustment is more positive. Furthermore, other studies have found that the most effective monitoring favors spontaneous communication (self-disclosure) by adolescents toward their parents (Stattin and Kerr, 2000; Gracia et al., 2012; Jiménez-Iglesias et al., 2013). Other family functioning variables (among which are affect or fondness of members of the family) have also been associated with fewer externalizing problems (in particular with adolescent aggressive behavior) (Pérez-Fuentes et al., 2019).

Digital Creativity, Parenting Styles, and Academic Performance

In the review of previous research on the relationship between digital creativity and parenting styles, some studies showed significant positive correlations between a permissive style and creativity (Miller et al., 2012) and negative correlations between the authoritarian parenting style and creativity (Fearon et al., 2013). However, in the study by Jackson et al. (2011), affect and control by parents were not significantly related to the use of the new technologies or their children's creativity.

Moreover, the relationship between parenting style and academic performance has been extensively studied from a psychoeducational perspective. Here we briefly comment on some of the most relevant results of studies reviewed. One previous study done by Pelegrina et al. (2002), on the typologies of parenting styles (democratic, permissive, authoritarian, and indifferent) and academic performance, with a sample of adolescents aged 11–15, revealed that children who perceived their parents as democratic or permissive had higher scores in academic areas. Later, Hernando et al. (2012) found results in the same direction, but in this case with regard to the different dimensions of parenting styles, such that the dimensions of behavior control and disclosure were significantly positively associated with academic performance, while psychological control and humor were negatively correlated. More recently, Rodríguez et al. (2018) found that the permissive style of both parents had a more positive influence on getting better grades and that both the permissive and democratic styles of father and mother were associated with stronger involvement at school. Other studies on how school and family can improve academic performance with digital methodologies and tools (Pérez-Fuentes et al., 2015) found that they improve academic motivation as a major part of learning and achievement behavior (González, 2018), to increase constructive thinking and psychological wellbeing (Quevedo-Aguado and Benavente, 2018).

Based on these empirical findings, the main objective of this study focused on the relationship between school-based everyday creativity and academic performance and the mediating role of parenting styles in this relationship. In addition to exploring the behavior of these variables in this sense, it also analyzed the predictive value of school-based everyday creativity and parenting styles for adolescent academic performance. In this respect, it was expected that parenting styles would exert a mediating effect on the relationship between school-based everyday creativity and academic performance in adolescence and that school-based everyday creativity and parenting styles would partially explain the academic performance of adolescents.

METHOD

Participants

The participants were selected by random sampling. Inclusion criteria were that the participants must be in high school, and exclusion criteria were not speaking the language well or having learning problems that caused them to be unable to answer the questionnaires on their own. This analysis was carried out in a total sample of 742 adolescents from five public middle schools and high schools (55.8 and 44.2%, respectively), in Almeria with aged 13–19 (M = 15.63; SD = 1.24). The sample of adolescents had a similar distribution of boys (46.7%) and girls (53.3%).

Instruments

Academic performance was measured based on *ad hoc* dichotomous yes/no questions on whether they had ever failed a subject or had ever repeated a year. Some authors use grade average as an optimum measure of academic performance (Lamas, 2015).

Everyday School-Based Activity

The Everyday School-Based Activity subscale of the Creative Behavior Questionnaire: Digital (CBQD) designed by Hoffmann et al. (2016) was used. The subscale is comprised of 10 items expressed on a Likert-type scale from 1 (*never*) to 5 (*4 or more times*). It is a self-reported measure of creative behavior at school. The CBQD has adequate reliability and validity (Hoffmann et al., 2016). In our sample for the School-Based Everyday Creativity Subscale, the Cronbach's α was 0.625. Some examples of the items on this scale are: *How many times in the last 3 months … Have you made presentations using PowerPoint, Prezi, KeyNote or others? How many times in the last 3 months … Have you made videos or movies using an app (a video app, for example)? How often in the last year, did you develop a blog or website for a class or a school project?*

Parenting Style

The Parenting Style Scale (Álvarez-García et al., 2016) was used. This is an adaptation of the instrument designed by Oliva et al. (2007). It has 24 items referring to the adolescent's perception of the educational style of their parents grouped in six dimensions: affect and communication, promotion of autonomy, behavioral control, psychological control, self-disclosure, and humor. In this study, the Cronbach's α for each of the subscales was 0.843, 0.814, 0.687, 0.710, 0.800, and 0.817, respectively.

Procedure

The study was approved by the Bioethics Committee of the University of Almeria (Ref: UALBIO2018/015). In all cases,

the ethical standards of research were compiled by using an informed consent sheet, and the ethical principles of the declaration of Helsinki were respected. To acquire the data, the management teams at the schools were contacted, and dates, schedules, and groups the instruments would be applied to were agreed upon.

Data Analysis

The study is a descriptive and correlational cross-sectional design. First, the correlation analyses were performed to explore the relationship between the variables, and the descriptive statistics were presented. For comparison of the fail/no fail subject and repeat/no repeat year groups, a Student's *t* test was performed with the Cohen's *d* for effect size estimation.

Then binary logistic regression models were estimated using the enter method. For this, the dependent variables in each case were Fail subject and Repeat year, with the dichotomous answer (yes/no). The predictor variables included were digital creativity (i.e., the School-Based Everyday Creativity Subscale) and parenting styles (Affect and communication, Promotion of autonomy, Behavior control, Psychological control, Self-disclosure, and Humor). The SPSS Statistical Package ver. 23.0 for Windows was used for data processing and analysis.

Finally, to perform the simple mediation analysis, the predictor variable was, in each case, having failed a subject or not and having repeated a year or not, respectively. In each case, as possible mediators, parenting styles that had resulted in involving the logistic equation were entered. For computation of the mediation models, the PROCESS macro for SSPS (Hayes, 2013) was used, applying bootstrapping with coefficients estimated from 5,000 bootstraps.

RESULTS

Digital Creativity and Parenting Style: Correlations and Descriptive Statistics

As shown in **Table 1**, school-based everyday creativity correlated positively with parenting styles: Affect and communication ($r = 0.12$, $p < 0.01$), Promotion of autonomy ($r = 0.09$, $p < 0.05$), Behavioral control ($r = 0.10$, $p < 0.01$), Self-disclosure ($r = 0.19$, $p < 0.001$), and Humor ($r = 0.10$, $p < 0.01$).

Table 2 presents the means in school-based everyday creativity, and each of the parenting styles, when groups who had failed a subject and those who had never failed, was compared. As observed in the table, the students who had failed a subject ($M = 2.28$, SD = 0.53) scored significantly lower in creativity ($t = -3.36$, $p < 0.01$; $d = 0.28$) than the group who had not failed ($M = 2.12$, SD = 0.58). The differences between the two groups with regard to parenting styles were in favor of those who had never failed a subject, scoring significantly higher in Affect and communication ($t = -4.27$, $p < 0.001$, $d = 0.35$), Promotion of autonomy ($t = -3.52$, $p < 0.001$, $d = 0.29$), Behavioral control ($t = -3.25$, $p < 0.01$, $d = 0.27$), Self-disclosure ($t = -5.14$, $p < 0.001$, $d = 0.42$), and Humor ($t = -2.41$, $p < 0.05$, $d = 0.20$). The Psychological control style had significantly higher scores ($t = 3.92$, $p < 0.001$, $d = 0.32$) in the group who had suspended a subject ($M = 9.45$, SD = 3.08) than those who had not ($M = 8.58$, SD = 2.45).

TABLE 1 | Digital creativity and parenting styles. Correlations and descriptive statistics (N=742).

	SEC	A_C	P_A	B_C	P_C	S_D	HU
SEC	—						
A_C	0.12**	—					
P_A	0.09*	0.67***	—				
B_C	0.10**	0.14***	0.02	—			
P_C	−0.04	−0.31***	−0.42***	0.29***	—		
S_D	0.19***	0.55***	0.46***	0.20***	−0.17***	—	
HU	0.10**	0.65***	0.57***	0.14***	−0.28***	0.51***	—
M	2.17	13.06	12.80	12.85	9.20	10.50	12.69
SD	0.58	2.82	2.77	2.62	2.92	3.33	2.69

*SEC, School-based everyday creativity; A_C, Affect and communication; P_A, Promotion of autonomy; B_C, Behavioral control; P_C, Psychological control; S_D, Self-disclosure; HU, Humor. Correlations and descriptive statistics (N = 742). *p < 0.05; **p < 0.01; ***p < 0.001.*

The comparison of repeaters/not repeaters with respect to school-based everyday activity, as observed in the table, showed that repeaters ($M = 2.01$, SD = 0.58) scored significantly lower ($t = -3.93$, $p < 0.001$, $d = 0.36$) than the group of non-repeaters ($M = 2.22$, SD = 0.56). In addition, by parenting styles, significant statistical differences were observed in which the repeaters ($M = 9.69$, SD = 3.21) had higher scores in the Psychological control style ($t = 2.10$, $p < 0.05$, $d = 0.19$) than non-repeaters ($M = 9.06$, SD = 2.84), whereas non-repeaters had higher mean scores in the Affect and communication ($t = -3.48$, $p < 0.01$, $d = 0.32$), Promotion of autonomy ($t = -2.15$, $p < 0.05$, $d = 0.20$), Behavioral control ($t = -2.86$, $p < 0.01$, $d = 0.26$), and Self-disclosure ($t = -2.84$, $p < 0.01$, $d = 0.26$) styles.

Logistic Regression Model: Fail a Subject

The dependent variable for logistic regression was having failed a subject or not. The predictor variables entered in the equation were school-based everyday creativity and parenting styles (Affect and communication, Promotion of autonomy, Psychological control, Self-disclosure, and Humor). These variables, the regression coefficients, standard error of estimation, Wald statistic, degrees of freedom, and associated probability, the partial correlation coefficient, and the cross-product ratio are shown in **Table 3**.

The odds ratio or cross-product ratio found for each variable shows that (1) the risk of failing a subject is higher in adolescents whose parents' educational styles are based on psychological control and humor and (2) school-based everyday creativity and behavior control and self-disclosure parenting styles are protective factors against the probability of failing a subject.

The overall fit of the model was ($\chi^2 = 45.68$; df = 7; $p < 0.001$) confirmed by the Hosmer-Lemeshow test ($\chi^2 = 7.77$;

TABLE 2 | Digital creativity and parenting style. Descriptive statistics and *t* test by failed subject/repeated year.

	Failed subject						*t*	*p*
	Yes			No				
	N	Mean	SD	*N*	Mean	SD		
School-based everyday creativity	481	2.12	0.58	215	2.28	0.53	−3.36**	0.001
Affect and communication	481	12.78	2.90	215	13.71	2.51	−4.27***	0.000
Promotion of autonomy	481	12.57	2.81	215	13.35	2.58	−3.52***	0.000
Behavioral control	481	12.64	2.67	215	13.32	2.45	−3.25**	0.001
Psychological control	481	9.45	3.08	215	8.58	2.45	3.92***	0.000
Self-disclosure	481	10.06	3.30	215	11.45	3.22	−5.14***	0.000
Humor	481	12.52	2.76	215	13.03	2.48	−2.41*	0.016
	Repeated year						***t***	***p***
	Yes			**No**				
	N	**Mean**	**SD**	***N***	**Mean**	**SD**		
School-based everyday creativity	149	2.01	0.58	552	2.22	0.56	−3.93***	0.000
Affect and communication	149	12.24	3.25	552	13.28	2.65	−3.48**	0.001
Promotion of autonomy	149	12.36	2.80	552	12.92	2.74	−2.15*	0.031
Behavioral control	149	12.25	2.83	552	13.01	2.55	−2.86**	0.005
Psychological control	149	9.69	3.21	552	9.06	2.84	2.10*	0.036
Self-disclosure	149	9.79	3.23	552	10.68	3.34	−2.84**	0.005
Humor	149	12.48	3.14	552	12.74	2.55	−0.94	0.346

Descriptive statistics and t test by failed subject/repeated year.

TABLE 3 | Results derived from the logistic regression for the probability of failing a subject, by school-based everyday creativity and parenting style.

Variables	β	SE	Wald	df	Sig.	Exp (β)	95% CI
School-based everyday creativity	−0.322	0.161	3.992	1	0.046	0.725	0.528–0.994
Affect and communication	−0.087	0.053	2.673	1	0.102	0.917	0.826–1.017
Promotion of autonomy	−0.001	0.050	0.000	1	0.986	0.999	0.906–1.102
Behavioral control	−0.120	0.042	8.288	1	0.004	0.887	0.818–0.962
Psychological control	0.113	0.039	8.387	1	0.004	1.119	1.037–1.208
Self-disclosure	−0.081	0.035	5.392	1	0.020	0.922	0.861–0.987
Humor	0.101	0.049	4.152	1	0.042	1.106	1.004–1.218
Constant	2.765	0.881	9.839	1	0.002	15.871	

df = 8; p = 0.455). The Nagelkerke R^2 showed that 10.5% of the variability in the response variable was explained by the logistic regression model. Based on the classification table, a probability of the logistic function being correct was 68.3%, with a false positive rate of 0.843 and a false negative rate of 0.92.

Logistic Regression Model: Repeat Year

In this case, to perform the logistic regression, the dependent variable was having repeated a year or not, while the predictor variables entered in the equation were, as in the previous model, school-based everyday creativity and the parenting styles. It may be observed in **Table 4** that the odds ratio found for each variable revealed that (1) adolescents who scored higher in school-based everyday creativity and whose parents/guardians had a parenting style based on affect and communication have a lower risk of repeating a year, or in other words, these two variables would be acting as protective factors against the probability of repeating and (2) concerning risk factors, control psychological as a parenting style would be significantly involved in the logistic equation.

Overall fit (χ^2 = 29.20; df = 7; p < 0.001) was confirmed by the Hosmer-Lemeshow test (χ^2 = 11.12; df = 8; p = 0.195), while the Nagelkerke R^2 coefficient showed that 7.8% of the variance was explained by the logistic regression model. Based on the classification table, probability of the logistic function being correct was 80.7%, with a false positive rate of 0.004 and a false negative rate of 0.

Mediation Models

Based on these results, we felt the need to find out whether certain parenting styles could be mediating in the relationship between failing a subject/repeating a year and the level of school-based everyday creativity. Therefore, simple mediation

TABLE 4 | Results derived from the logistic regression for the probability of repeat year, according to everyday creativity and parental styles.

Variables	β	SE	Wald	df	Sig.	Exp (β)	95% CI
School-based everyday creativity	−0.546	0.198	7.590	1	0.006	0.579	0.393–0.854
Affect and communication	−0.166	0.057	8.303	1	0.004	0.847	0.757–0.948
Promotion of autonomy	0.061	0.055	1.224	1	0.269	1.063	0.954–1.185
Behavioral control	−0.078	0.045	3.080	1	0.079	0.925	0.847–1.009
Psychological control	0.091	0.044	4.345	1	0.037	1.095	1.005–1.193
Self-disclosure	−0.010	0.040	0.063	1	0.801	0.990	0.915–1.071
Humor	0.109	0.057	3.628	1	0.057	1.115	0.997–1.247
Constant	−0.073	0.944	0.006	1	0.939	0.930	

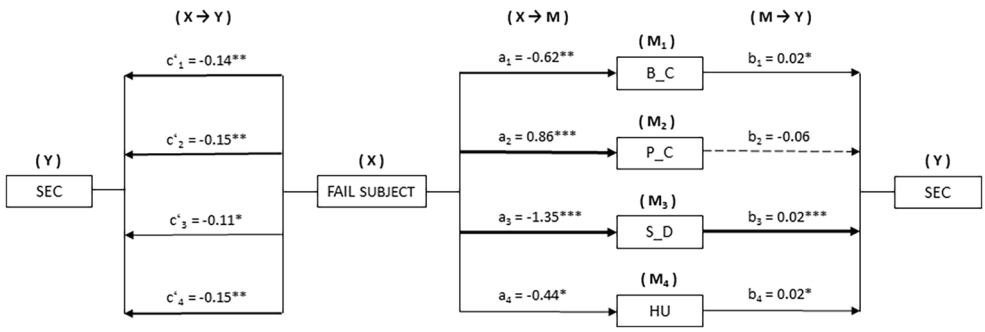


FIGURE 1 | Mediation model for parenting styles on the relationship between failing a subject and school-based everyday creativity. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

models were computed, including the parenting styles involved in the corresponding logistic equation as mediators in each case.

Figure 1 shows the mediation models taking fail a subject or not as the independent variable (X). In this case, the behavioral control (M_1), psychological control (M_2), self-disclosure (M_3), and humor (M_4) parenting styles were entered as possible mediators on the effect in school-based everyday creativity (Y).

In the first place, a significant relationship was observed between failing a subject (X) and parenting styles (M): B_C [$B = -0.62$, $p < 0.01$], P_C [$B = 0.86$, $p < 0.001$], S_D [$B = -1.35$, $p < 0.001$], and HU [$B = -0.44$, $p < 0.05$]. Estimation of direct effects X → Y showed significance of failing a subject on school-based everyday creativity (Y) in each of the models computed: B_C [$B = -0.14$, $p < 0.01$], P_C [$B = -0.15$, $p < 0.01$], S_D [$B = -0.11$, $p < 0.05$], and HU [$B = -0.15$, $p < 0.01$]. Furthermore, estimation of M → Y found significant effects on school-based everyday creativity (Y) in three parenting styles (M): B_C [$B = 0.02$, $p < 0.05$], S_D [$B = 0.02$, $p < 0.001$], and HU [$B = 0.02$, $p < 0.05$].

With the analysis of indirect effects (X → M → Y) with bootstrapping, significance was found in three of four models computed: B_C [$B = -0.01$, SE = 0.007, 95% CI (−0.031, −0.002)], S_D [$B = -0.04$, SE = 0.012, 95% CI (−0.069, −0.018)], and HU [B = −0.009, SE = 0.006, 95% CI (−0.027, −0.000)].

Figure 2 presents the mediation models with repeat year or not (X) as the independent variable. In this case, the Affect and communication (M_1) and Psychological control (M_2) parenting styles were entered as possible mediators on the effect in school-based everyday creativity (Y).

In the first place, significant relationships were observed between repeating a year (X) and parenting styles (M): A_C [$B = -1.01$, $p < 0.001$] and P_C [$B = 0.68$, $p < 0.05$]. The estimation results of the direct effects (X → Y) revealed the significant relationships of repeating the year (X) on school-based everyday creativity (Y), in computation of both models: A_C [$B = -0.18$, $p < 0.01$] and P_C [$B = -0.19$, $p < 0.001$]. The estimation of the M → Y effects found a significant effect of the A_C parenting style [$B = 0.02$, $p < 0.05$] on school-based everyday creativity.

Finally, with the analysis of the indirect effects (X → M → Y) with bootstrapping, significance was found in the model that took the A_C [$B = -0.02$, SE = 0.010, 95% CI (−0.047, −0.003)] parenting style as the mediator.

DISCUSSION

Based on the pioneering study by Hoffmann et al. (2016) on digital creativity and its conceptualization in the school, in this study we analyzed the relationship among school-based everyday creativity, parenting styles, and academic performance. Our first conclusion is that parenting styles have a relevant role in school-based everyday creativity and academic performance. These main results are discussed below.

First, with respect to the relationship between *parenting style and digital creativity*, our results showed that adolescents who perceived that their parents had a more democratic relational style, although with higher scores in behavior control,

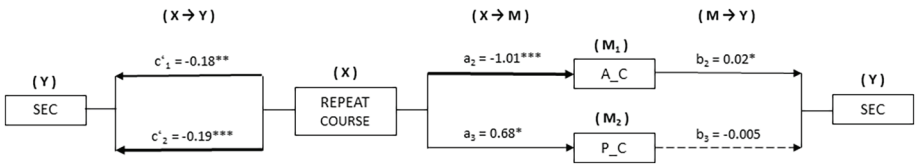


FIGURE 2 | Mediation model of parenting styles on the relationship between repeating a year and school-based everyday creativity. *$^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.*

had better results in school-based everyday creativity. In this sense, other researchers, analyzing creativity from a perspective other than digital, have also found significant correlations between young people's perception of high affect in parenting style and the perception of possessing personal characteristics and creative thinking styles (Miller et al., 2012), pointing in the same direction as our data. In another direction, a study by Fearon et al. (2013) with Jamaican students and their parents found the high level of control, which characterized the authoritarian parenting style to be associated with lower levels of creativity. On the other hand, in the study by Jackson et al. (2011), parental behavior in the dimensions of affect and control (measured through the perception of parents and adolescents) did not significantly correlate with the use of new technologies, or with their child's creativity, which was measured using the Torrance Creative Thinking Test. With respect to the dimension of parental control, these results could be due to the lack of differentiation by the researchers between the dimensions of behavioral and psychological control in measuring parenting styles, so that in other studies, psychological control, which has been associated with adolescent adjustment problems (Parra and Oliva, 2006; Barber et al., 2012), would have been measured instead of behavioral control. Second, concerning the comparison of *academic performance and digital creativity*, our data showed that those students who did not fail a subject or who had not repeated a year scored higher on school-based everyday creativity. Other researchers have reported similar findings, emphasizing the positive role of the use of creative methodologies as facilitators of more significant student learning (Guilford, 1950; Renzulli et al., 1986; Limiñana et al., 2010). Third, with regard to the *parenting style and academic performance*, we found that the perception of psychological control by adolescents was a risk factor, for both failing a subject and repeating a year. Furthermore, perceived humor in the figure of the parents also contributed to failing a subject, and along with the perception of psychological control may allow ways of parental relating to be seen that do not respect the individualism of the child, forms of relating that have also been associated with low self-esteem (Silk et al., 2003). Other studies have reported similar results (Hernando et al., 2012). Apart from that, in the study by Pelegrina et al. (2002), the perception of democratic and permissive styles by adolescents correlated positively with academic performance, but the perception of parental control did not relate negatively with academic performance, perhaps again because behavioral and psychological control were not differentiated. The results of the study by Rodríguez et al. (2018) support those results.

Fourth, the objective of developing an explanatory model for academic performance based on variables such as digital creativity and parenting styles was met in a logistic regression analysis. The results of that analysis reflected school-based everyday creativity as a factor that can promote adequate academic performance, with other parental variables relevant for not failing a subject, such as the behavioral control and self-disclosure parenting style dimensions, and the affect and communication parenting style dimension for not repeating the year. The model explained 10.5 and 7.8% of adolescent academic performance, respectively. Therefore, both school-based everyday creativity and behavioral control and self-disclosure, on one hand, and the affect and communication parenting style, on the other, contributed to explaining part of adolescents' academic performance. These variables could probably have explained an even larger proportion of academic performance if the teachers had employed more active teaching learning methodologies. However, most of the schools still use a traditional teaching-learning methodology involving psychological processes, which do not promote creativity, but, on the contrary, limit its development. Other studies have emphasized the preponderant role of creative digital methodologies as a significant facilitating tool for classroom teaching and learning, leading to improved digital competence from an innovative perspective (Guilford, 1950; Renzulli et al., 1986; Ala-Mutka et al., 2008; Limiñana et al., 2010; Fernández-Batanero and Rodríguez-Martín, 2017), making use this way of the high creative potential of the adolescent, which is greater than in other stages of development (Claxton et al., 2005; Kim, 2011).

Fifth, the new analyses performed in this study on the mediating role of parenting style in the relationship between school-based everyday creativity and academic performance revealed that the association between school-based everyday creativity and failing a subject was mediated by the perception of parenting styles (high behavioral control, high self-disclosure, and low psychological control). Furthermore, the relationship between school-based everyday creativity and repeating a year was mediated by the perception of parenting styles with high doses of communication and affect and low psychological control in parent-child interaction.

Among the limitations of this study is its cross-sectional design, which could be completed with longitudinal studies contributing to generalization of the results. Another possible limitation has to do with the fact that the parenting styles were evaluated based only on the perceptions of the adolescents,

and other contributions could be added, such as those of teachers and parents. Based on all of the above, more studies are required to enlarge these first results on digital creativity in the school and that include new variables referring to the adolescent, the family, and teachers.

CONCLUSIONS

In view of the above, the results of this study enable two broad conclusions to be arrived at for a possible intervention proposal. It was concluded that parenting styles have a relevant role in developing digital creativity in the school and its relationship with academic performance in adolescence. Therefore, the data suggest that the family continues to play a relevant role as an educating agent in this stage of development, and above all, the parents' educational style. Thus, the perception of adolescents of a parenting style characterized by high affect and communication, high behavioral control, low psychological control, and high self-disclosure predicts better academic performance in adolescence. This result is important from the point of view of educational intervention to the extent that it demonstrates that intervention with families must be a priority objective of secondary education teachers. Thus, the school could take action with the family in participative, individualized counseling sessions, where teachers can advise the families based on their real needs or interests, with simple, plausible educational patterns contributing to the adolescent's positive development. Educational projects could also be carried out that include parenting workshops and activities, thereby favoring and improving the family-school relationship.

Considering that the more traditional teaching methodologies and evaluation systems cannot faithfully include adolescent digital creativity, it is concluded that digital creative methodologies must be developed as a significant facilitating tool of classroom teaching and learning for the main goal of improving adolescent academic performance, by applying digital creativity to the school environment. This type of methodology could also be widened to the family environment.

As a future line of research, the educational implications of digital creativity in adolescent development could continue to be studied, including new variables in the design, such as self-concept or self-esteem, agreeableness, and self-efficacy, and in any case, involved directly with personal development. The study of digital creativity could also be widened by adapting the instrument to other contexts and populations.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Bioethics Committee of the University of Almeria. Written informed consent to participate in this study was provided by the participants and their legal guardian/next of kin where appropriate. The study was approved by the Bioethics Committee of the University of Almeria (Ref: UALBIO2018/015).

AUTHOR CONTRIBUTIONS

MP-F, MM, NO, and MS contributed to the conception and design of the review. JG applied the search strategy. All authors applied the selection criteria, completed the assessment of risk of bias, and analyzed the data and interpreted data. MM, MP-F, MS, and NO wrote this manuscript. MP-F and JG edited this manuscript. MP-F is responsible for the overall project.

FUNDING

The present study was undertaken in collaboration with the Peer violence and alcohol and tobacco use in Secondary Education program: an augmented reality program for detection and intervention (Reference: EDU2017-88139-R), funded by the State Research Program, Development and Innovation Oriented to the Challenges of Society, within the framework of the State Plan for Scientific and Technical Research and Innovation, and co-financing with Structural Funds of the European Union. Part of this work has been developed thanks to the financing of University Teaching Training in Deficit Areas, Gerty Cori Aid, for the help in hiring research staff in predoctoral training, granted by María del Mar Simón Márquez.

REFERENCES

Ala-Mutka, K., Punie, Y., and Redecker, C. (2008). *ICT for learning, innovation and creativity.* Luxembourg: Institute for Prospective Technological Studies (IPTS), European Commission, Joint Research Center.

Álvarez-García, D., García, T., Barreiro-Collazo, A., Dobarro, A., and Antúnez, Á. (2016). Parenting style dimensions as predictors of adolescent antisocial behavior. *Front. Psychol.* 7:1383. doi: 10.3389/fpsyg.2016.01383

Amabile, T. M. (1985). Motivation and creativity: effects of motivational orientation creative writers. *J. Pers. Soc. Psychol.* 48, 393–399.

Barber, B. K., Xia, M., Olsen, J. A., McNeely, C. A., and Bose, K. (2012). Feeling disrespected by parents: refining the measurement and understanding of psychological control. *J. Adolesc.* 35, 273–287. doi: 10.1016/j.adolescence.2011.10.010

Baumrind, D. (1968). Authoritarian vs. authoritative parental control. *Adolescence* 3, 255–272.

Cabero-Almenara, J., and Díaz, V. M. (2014). Posibilidades educativas de las redes sociales y el trabajo en grupo: Percepciones de los alumnos universitarios [educational possibilities of social networks and group work. University students' perceptions]. *Comun.: Rev. Científica Iberoam. Comun. Educ.* 42, 165–172. doi: 10.3916/C42-2014-16

Claxton, A. F., Pannells, T. C., and Rhoads, P. A. (2005). Developmental trends in the creativity of school-age children. *Creat. Res. J.* 17, 327–335. doi: 10.1207/s15326934crj1704_4

Csikszentmihalyi, M. (1996). *Creativity. Flow and the psychology of discovery and invention*. New York: Harper Collins Publishers.

Cui, L., Morris, A. S., Criss, M. M., Houltberg, B. J., and Silk, J. S. (2014). Parental psychological control and adolescent adjustment: the role of adolescent emotion regulation. *Parenting: Sci. Prac.* 14, 47–67. doi: 10.1080/15295192.2014.880018

Estévez, E., Jiménez, T. I., and Moreno, D. (2018). Aggressive behavior in adolescence as a predictor of personal, family, and school adjustment problems. *Psicothema* 30, 66–73. Available at: http://www.redalyc.org/articulo.oa?id=72754594011

Fearon, D. D., Copeland, D., and Saxon, T. F. (2013). The relationship between parenting styles and creativity in a sample of Jamaican children. *Creat. Res. J.* 25, 119–128. doi: 10.1080/10400419.2013.752287

Fernández-Batanero, J. M., and Rodríguez-Martín, A. (2017). TIC y diversidad funcional: conocimiento del profesorado [ICT and functional diversity: knowledge of the teaching staff]. *Eur. J. Invest. Health Psychol. Educ.* 7, 157–175. doi: 10.30552/ejihpe.v7i3.204

Fuentes, M. C., García, F., Gracia, E., and Alarcón, A. (2015). Los estilos parentales de socialización y el ajuste psicológico. Un estudio con adolescentes españoles [Parental styles of socialization and psychological adjustment. A study with Spanish adolescents]. *Rev. Psicodidáctica* 20, 117–138. doi: 10.1387/RevPsicodidact.10876

García-Martín, J., García-Sánchez, J. N., Álvarez-Fernández, M. L., and Díez-Caso, H. (2014). Efectos en la competencia digital tras la aplicación de un programa de competencias ocupacionales [Effects on digital competence after application of a program of occupational skills]. *Eur. J. Educ. Psychol.* 7, 73–81. doi: 10.1989/ejep.v7i2.180

Gardner, H. (1993). *Creating minds*. New York: Basic Books.

González, M. L. G. (2018). Incidencia del sexo, número de hermanos y orden de Nacimiento en las metas académicas de estudiantes universitarios [incidence of sex, number of siblings and birth order in the academic goals of university students]. *Eur. J. Child Dev., Educ. Psychopathology* 6, 57–66. doi: 10.30552/ejpad.v6i1.62

Gracia, E., Fuentes, M. C., Garcia, F., and Lila, M. (2012). Perceived neighborhood violence, parenting styles, and developmental outcomes among Spanish adolescents. *J. Community Psychol.* 40, 1004–1021. doi: 10.1002/jcop.21512

Guilford, J. P. (1950). Creativity. *Am. Psychol.* 5, 444–454.

Hauser-Kunz, J., and Grych, J. H. (2013). Parental psychological control and autonomy granting: distinctions and associations with child and family functioning. *Parenting* 13, 77–94. doi: 10.1080/15295192.2012.709147

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional Process analysis: A regression-based approach*. New York, EE.UU: The Guilford Press.

Hernando, Á., Oliva, A., and Pertegal, M. Á. (2012). Variables familiares y rendimiento académico en la adolescencia [Family variables and academic achievement in adolescence]. *Estud. Psicol.* 33, 51–65. doi: 10.1174/021093912799803791

Ho, S. M., Dai, D. W. T., Mak, C., and Liu, K. W. K. (2018). Cognitive factors associated with depression and anxiety in adolescents: a two-year longitudinal study. *Int. J. Clin. Health Psychol.* 18, 227–234. doi: 10.1016/j.ijchp.2018.04.001

Hoffmann, J., Ivcevic, Z., and Brackett, M. (2016). Creativity in the age of technology: measuring the digital creativity of millennials. *Creat. Res. J.* 28, 149–153. doi: 10.1080/10400419.2016.1162515

Ivcevic, Z. (2007). Artistic and everyday creativity: an act-frequency approach. *J. Creat. Behav.* 41, 271–290. doi: 10.1002/j.2162-6057.2007.tb01074.x

Jackson, L. A., Witt, E. A., Fitzgerald, H. E., Von Eye, A., and Zhao, Y. (2011). "Parent behavior, children's technology use and creativity: videogames count but parents don't" in *Proceedings of the 5th WSEAS international conference on Communications and information technology*. eds. N. E. Mastorakis, and Z. Bojkovicc (Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS)), 19–24.

Jiménez-Iglesias, A., Moreno, C., García-Moya, I., and Ramos, P. (2013). How can parents obtain knowledge about their adolescent children? *Infanc. Aprendizaje* 36, 181–197. doi: 10.1174/021037013806196256

Kim, K. H. (2011). The creativity crisis: the decrease in creative thinking scores on the Torrance tests of creative thinking. *Creat. Res. J.* 23, 285–295. doi: 10.1080/10400419.2011.627805

Lamas, H. A. (2015). Sobre el rendimiento escolar [School Performance]. *Propós. represent.* 3, 313–386. doi: 10.20511/pyr2015.v3n1.74

Limiñana, R. M., Bordoy, M., Juste, G., and Corbalán, J. (2010). Creatividad, aptitudes intelectuales y estilos de respuesta: implicaciones para el rendimiento académico en secundaria [Creativity, intelectual abilities and response styles: Implications for academic performance in the secondary school]. *Ann. Psychol.* 26, 212–219. Retrieved from: https://revistas.um.es/analesps/article/view/109081

Maccoby, E. E., and Martin, J. A. (1983). "Socialization in the context of the family: parent-child interaction" in *Handbook of child psychology, Vol. IV: Socialization, personality and social development. 4th Edn.* eds. E. M. Hetherington, and P. H. Mussen (Nueva York: Wiley), 1–101.

Miller, A. L., Lambert, A. D., and Speirs Neumeister, K. L. (2012). Parenting style, perfectionism, and creativity in high-ability and high-achieving young adults. *J. Educ. Gift.* 35, 344–365. doi: 10.1177/0162353212459257

Oliva, A., Parra, Á., and Arranz, E. (2008). Estilos relacionales parentales y ajuste adolescente [Parenting styles and adolescent adjustment]. *Infanc. Aprendizaje* 31, 93–106. doi: 10.1174/021037008783487093

Oliva, A., Parra, Á., Sánchez-Queija, I., and López, F. (2007). Estilos educativos materno y paterno: evaluación y relación con el ajuste adolescente [maternal and paternal parenting styles: assessment and relationship with adolescent adjustment]. *Ann. Psychol.* 23, 49–56. Retrieved from: https://www.redalyc.org/articulo.oa?id=167/16723107

Parra, Á., and Oliva, A. (2006). Un análisis longitudinal sobre las dimensiones relevantes del estilo parental durante la adolescencia [Relevant dimensions of parenting style during adolescence: A longitudinal study]. *Infanc. Aprendizaje* 29, 453–470. doi: 10.1174/021037006778849594

Pelegrina, S., García-Linares, M. C. G., and Casanova, P. F. (2002). Los estilos educativos de los padres y la competencia académica de los adolescentes [Parenting styles and adolescents' academic performance]. *Infanc. Aprendizaje* 25, 147–168. doi: 10.1174/021037002317417796

Pérez-Fuentes, M. D. C., Álvarez-Bermejo, J. A., Molero, M. D. M., Gázquez, J. J., and López, M. A. L. (2015). Violencia escolar y rendimiento académico (VERA): Aplicación de realidad aumentada [School violence and academic achievement (vera): Augmented reality application]. *Eur. J. Invest. Health, Psychol. Educ.* 1, 71–84. doi: 10.1989/ejihpe.v1i2.6

Pérez-Fuentes, M. D. C., Gázquez, J. J., Mercader, I., Molero, M. D. M., and García, M. D. M. (2011). Rendimiento académico y conductas antisociales y delictivas en alumnos de Educación Secundaria Obligatoria [Academic achievement and antisocial behavior in public secondary education students]. *Int. J. Psychol. Psychol. Ther.* 11, 401–412. Retrieved from: https://www.ijpsy.com/volumen11/num3/307.html

Pérez-Fuentes, M. D. C., Molero, M. D. M., Barragán, A. B., and Gázquez, J. J. (2019). Family functioning, emotional intelligence, and values: analysis of the relationship with aggressive behavior in adolescents. *Int. J. Environ. Res. Public Health* 16, pii: 478. doi: 10.3390/ijerph16030478

Quevedo-Aguado, M. P., and Benavente, M. H. (2018). Análisis de variables de personalidad, bienestar psicológico y pensamiento constructivo en estudiantes de Ciencias de la Salud [analysis of personality variables, psychological well-being and constructive thinking in students of sciences of health]. *Eur. J. Health Res.* 4, 5–18. doi: 10.30552/ejhr.v4i1.86

Renzulli, M. J., Gay-Ford, B., Smith, L., and Renzulli, J. (1986). *New directions in creativity*. Connecticut: Creative Learning Pres, Inc.

Richards, R., Kinney, D. K., Benet, M., and Merzel, A. P. C. (1988). Assessing everyday creativity: characteristics of the everyday creativity scales and validation with three large samples. *J. Pers. Soc. Psychol.* 54, 467–485.

Riquelme, M., Garcia, O. F., and Serra, E. (2018). Psychosocial maladjustment in adolescence: parental socialization, self-esteem, and substance use. *An. Psicol.* 34, 536–544. doi: 10.6018/analesps.34.3.315201

Rodríguez, L., Revuelta, M., Sarasa, M., and Fernández, O. (2018). The role of parental socialization styles in school engagement and academic performance. *Eur. J. Educ. Psychol.* 11, 123–139. doi: 10.30552/ejep.v11i2.226

Ruiz-Hernández, J. A., Moral-Zafra, E., Llor-Esteban, B., and Jiménez-Barbero, J. A. (2018). Influence of parental styles and other psychosocial variables on the development of externalizing behaviors in adolescents: a systematic review. *Eur. J. Psychol. Appl. Leg. Context* 11, 9–21. doi: 10.5093/ejpalc2018a11

Runco, M. A. (2004). "Everyone has creative potential" in *Creativity: From potential to realization*. eds. R. J. Sternberg, E. L. Grigorenko, and J. L. Singer (Washington, DC: American Psychological Association), 21–30.

Silk, J. S., Morris, A. S., Kanaya, T., and Steinberg, L. (2003). Psychological control and autonomy granting: opposite ends of a continuum or distinct constructs? *J. Res. Adolesc.* 13, 113–128. doi: 10.1111/1532-7795.1301004

Soenens, B., and Vansteenkiste, M. (2010). A theoretical upgrade of the concept of parental psychological control: proposing new insights on the

basis of self-determination theory. *Dev. Rev.* 30, 74–99. doi: 10.1016/j.dr.2009.11.001

Stattin, H., and Kerr, M. (2000). Parental monitoring: A reinterpretation. *Child Dev.* 71, 1072–1085. doi: 10.1111/1467-8624.00210

Sun, J., Chen, Q., Zhang, Q., Li, Y., Li, H., Wei, D., et al. (2016). Training your brain to be more creative: brain functional and structural changes induced by divergent thinking training. *Hum. Brain Mapp.* 37, 3375–3387. doi: 10.1002/hbm.23246

Torrance, E. P. (1988). "The nature of creativity as manifest in its testing" in *The nature of creativity: Contemporary psychological perspectives*. ed. R. J. Sternberg (New York, NY, US: Cambridge University Press), 43–75.

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.

Copyright © 2019 Pérez-Fuentes, Molero Jurado, Oropesa Ruiz, Simón Márquez and Gázquez Linares. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 08 November 2019
doi: 10.3389/fpsyg.2019.02497

Academic Performance in Adolescent Students: The Role of Parenting Styles and Socio-Demographic Factors – A Cross Sectional Study From Peshawar, Pakistan

Sarwat Masud[1], Syed Hamza Mufarrih[2], Nada Qaisar Qureshi[2], Fahad Khan[1], Saad Khan[1] and Muhammad Naseem Khan[1]**

[1] Institute of Public Health & Social Sciences, Khyber Medical University, Peshawar, Pakistan, [2] Department of Medicine, Aga Khan University Hospital, Karachi, Pakistan

OPEN ACCESS

Edited by:
Raquel Gilar,
University of Alicante, Spain

Reviewed by:
Carolina Pérez-Dueñas,
Universidad de Córdoba, Spain
Olivia López Martínez,
University of Murcia, Spain

***Correspondence:**
Sarwat Masud
sarwatmasud@gmail.com
Muhammad Naseem Khan
drnasim@kmu.edu.pk;
drnasim@msn.com

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

Received: 16 May 2019
Accepted: 22 October 2019
Published: 08 November 2019

Citation:
Masud S, Mufarrih SH, Qureshi NQ, Khan F, Khan S and Khan MN (2019) Academic Performance in Adolescent Students: The Role of Parenting Styles and Socio-Demographic Factors – A Cross Sectional Study From Peshawar, Pakistan. Front. Psychol. 10:2497. doi: 10.3389/fpsyg.2019.02497

Academic performance is among the several components of academic success. Many factors, including socioeconomic status, student temperament and motivation, peer, and parental support influence academic performance. Our study aims to investigate the determinants of academic performance with emphasis on the role of parental styles in adolescent students in Peshawar, Pakistan. A total of 456 students from 4 public and 4 private schools were interviewed. Academic performance was assessed based on self-reported grades in the latest internal examinations. Parenting styles were assessed through the administration of the Parental Bonding Instrument (PBI). Regression analysis was conducted to assess the influence of socio-demographic factors and parenting styles on academic performance. Factors associated with and differences between "care" and "overprotection" scores of fathers and mothers were analyzed. Higher socio-economic status, father's education level, and higher care scores were independently associated with better academic performance in adolescent students. Affectionless control was the most common parenting style for fathers and mothers. When adapted by the father, it was also the only parenting style independently improving academic performance. Overall, mean "care" scores were higher for mothers and mean "overprotection" scores were higher for fathers. Parenting workshops and school activities emphasizing the involvement of mothers and fathers in the parenting of adolescent students might have a positive influence on their academic performance. Affectionless control may be associated with improved academics but the emotional and psychosocial effects of this style of parenting need to be investigated before recommendations are made.

Keywords: parenting styles, academic performance, adolescent students, Pakistan, care, overprotection, parental bonding instrument

INTRODUCTION

Despite residual ambiguity in the term, definitions over time have identified several elements of "academic success" (Kuh et al., 2006; York et al., 2015). Used interchangeably with "student success," it encompasses academic achievement, attainment of learning objectives, acquisition of desired skills and competencies, satisfaction, persistence, and post-college performance (Kuh et al., 2006; York et al., 2015). Linked to happiness in undergraduate students (Flynn and MacLeod, 2015) and low health risk behavior in adolescents (Hawkins, 1997), a vast amount of literature is available on the determinants of academic success. Studies have shown socioeconomic characteristics (Vacha and McLaughlin, 1992; Ginsburg and Bronstein, 1993; Chow, 2000; McClelland et al., 2000; Tomul and Savasci, 2012), student characteristics including temperament, motivation and resilience (Ginsburg and Bronstein, 1993; Linnenbrink and Pintrich, 2002; Farsides and Woodfield, 2003; Valiente et al., 2007; Beauvais et al., 2014) and peer (Dennis et al., 2005), and parental support (Cutrona et al., 1994; Sanders, 1998; Dennis et al., 2005; Bean et al., 2006) to have a bearing on academic performance in students.

The influence of parenting styles and parental involvement is particularly in focus when assessing determinants of academic success in adolescent children (Shute et al., 2011; Rahimpour et al., 2015; Weis et al., 2016; Checa and Abundis-Gutierrez, 2017; Zhang et al., 2019). The influence may be of significance from infancy through adulthood (Steinberg et al., 1989; Weiss and Schwarz, 1996; Zahedani et al., 2016) and can be appreciated across a range of ethnicities (Desimone, 1999; Battle, 2002; Jeynes, 2007). Previously, the authoritative parenting style has been most frequently associated with better academic performance among adolescent students (Steinberg et al., 1989, 1992; Deslandes et al., 1997, 1998; Aunola et al., 2000; Adeyemo, 2005; Checa et al., 2019), while purely restrictive and negligent styles have shown to have a negative influence on academic performance (Hillstrom, 2009; Parsasirat et al., 2013; Osorio and González-Cámara, 2016). Parenting styles have also been linked to academic performance indirectly through regulation of emotion, self-expression (Deslandes et al., 1997; Weis et al., 2016), and self-esteem (Zakeri and Karimpour, 2011).

Significant efforts have been made to explore and integrate factors which influence parenting stress and behaviors (Belsky, 1984; Abidin, 1992; Östberg and Hagekull, 2000). A number of factors, including parent personality and psychopathology (in terms of extraversion, neuroticism, agreeableness, depression and emotional stability), parenting beliefs, parent-child relationship, marital satisfaction, parenting style of spouse, work stress, child characteristics, education level, and socioeconomic status have been highlighted for their role in determining parenting styles (Belsky, 1984; Simons et al., 1990, 1993; Bluestone and Tamis-LeMonda, 1999; Huver et al., 2010; Smith, 2010; McCabe, 2014). Studies have also highlighted differences between fathers and mothers in how these factors influence them (Simons et al., 1990; Ponnet et al., 2013).

Insight into determinants of academic success and the role of parenting styles can have significant impact on policy recommendations. However, most existing data comes from western cultures where individualistic themes predominate. While some studies highlight differences between the two (Wang and Leichtman, 2000), evidence from eastern collectivist cultures, including Pakistan, is scarce (Masud et al., 2015; Khalid et al., 2018).

The aim of this study is to identify the determinants of academic performance, including the influence of parenting styles, in adolescent students in Peshawar, Pakistan. We also aim to investigate the factors affecting parenting styles and the differences between parenting behaviors of father and mothers.

MATERIALS AND METHODS

The manuscript has been reported in concordance with the STROBE checklist (Vandenbroucke et al., 2014).

Study Design

A cross sectional study was conducted by interviewing school-going students (grades 8, 9, and 10) to assess determinants of academic grades including the influence of parenting styles.

Setting

The study took place in the city of Peshawar in Pakistan at eight schools, four from the public sector and four from the private sector. The data collection process began in January 2017 concluded in December 2017.

Study Size

The prevalence of high grades (A and A plus) among adolescent students was between 42.6 and 57.4% in a previous study (Cohen and Rice, 1997 #248). Based on this, a sample size of 376 students was calculated to study the determinants of high grades in adolescent students with a confidence level of 95%. Assuming a non-response rate of approximately 20%, we decided to target 500 students from four public and four private schools. A total of 456 students participated in our study.

Participants

Inclusion Criteria

From the eight schools which provided admin consent to conduct the study, students enrolled in grade 8, 9, or 10 were invited to take part in the study. Following consent from the parents and assent from the student, he or she was included in the study.

Exclusion Criteria

Any student unable to understand or fill out the interview *pro forma* or questionnaire independently.

Data Sources and Measurement

Data was collected through a one on one interaction between each student and the data collector individually. The following tools were used.

Demographic *pro forma* (Supplementary Datasheet 1)

A brief and simple *pro forma* was structured to address all demographic related variables needed for the study.

Parental Bonding Instrument (PBI) (Supplementary Datasheet 2)

The original version of the Parental Bonding Instrument (Parker et al., 1979), previously validated for internal consistency, convergent validity, satisfactory construct, and independence from mood effects in several different populations, including Turkish and Chinese (Parker et al., 1979; Parker, 1983, 1990; Cavedo and Parker, 1994; Dudley and Wisbey, 2000; Wilhelm et al., 2005; Murphy et al., 2010; Liu et al., 2011; Behzadi and Parker, 2015), was employed in our study. This tool, composed of 25 questions, assesses parenting styles as two independent measures of "care" and "control" as perceived by the child. It is filled out separately for the father and the mother. It is available online for use without copyright. The use of PBI has been validated for British Pakistanis (Mujtaba and Furnham, 2001) and Pakistani women (Qadir et al., 2005). A paper by Qadir et al. on the validity of PBI for Pakistani women, reports the Cronbach alpha scores to be 0.91 and 0.80 for the "care" and "overprotection" scales, respectively (Qadir et al., 2005).

The demographic *pro forma* and the parental bonding index were translated into Urdu by an individual fluent in both languages and validated with the help of an epidemiologist and two experts in the field (**Supplementary Datasheet 3**). Pilot testing of translated versions was done with 20 students to ensure clarity and assess understanding and comprehension by the students. Both versions for the two tools were provided in hard copy to each student to fill out whichever one he/she preferred. The data collector first verbally explained the items on the demographic *pro forma* and the PBI to the student following which the student was allowed to fill it out independently.

Variables

Using the data sources mentioned above, data was collected for the following variables.

Student Related

Gender, type of school (public or private), class grade (8th, 9th, and 10th) and academic performance.

In Pakistan, public and private schools may differ in several aspects including fee structures, class strength and difficulty levels of internal examinations, with private schools being more expensive, with fewer students per classroom, and subjectively tougher internal examinations.

The academic performance was judged as the overall grade (a combination of all subjects including English, Mathematics and Science) in the latest internal examinations sat by the student as A+, A, B, C, or D.

Family Related

Family structure and type of accommodation (rented or owned).

Parent Related

Information on living status, education level, employment status, employment type and parenting styles was obtained from the student separately for the father and mother.

TABLE 1 | Calculation of an estimated socioeconomic status.

Variables	Points	
School type	Public = 0	Private = 1
Family structure	Joint = 0	Nuclear = 1
Father's employment status	No = 0	Yes = 1
Accommodation status	Rented = 0	Owned = 2
Father's job type	Unemployed = 0	Government sector = 1 Private sector = 2
Mother's employment status	Unemployed = 0	Employed = 1

High SES, score ≥ 6; Low SES, score < 6; SES, socioeconomic status.

Quantitative Variables

Academic Performance

The grades A+, A were categorized as "high" grades and grades B, C, and D were categorized as "low" grades.

Socio-Economic Status

We used variables which adolescent students are expected to have knowledge of to calculate a score which categorized students as belonging to either a high or low socioeconomic status. The points assigned to each variable are show in **Table 1**.

Parenting Styles

The PBI is a 25 item questionnaire, with 12 items measuring "care" and 13 items measuring "overprotection." All responses have a 4 point Likert scale ranging from 0 (very unlikely) to 3 (very likely). The responses are summed up to categorize each parent to exhibit low or high "care" and low or high "overprotection." Based on these findings, each parent can then be put into one of the 4 quadrants representing parenting styles including "affectionate constraint," "affectionless control," "optimal parenting," and "neglectful parenting." This computation is explained in **Figure 1** obtained from the information provided with the PBI (Parker et al., 1979).

Bias

Students were allowed to fill in the *pro forma* and questionnaire independently to avoid bias during the data collection process. However, self-reporting of grades in latest examination may be subject to recall bias.

Statistical Methods

Statistical analysis was performed using SPSS v.23 (IBM Corp., Armonk, NY, United States). Descriptive analyses were conducted on all study variables including socio-demographic factors and parenting styles. Categorical variables were reported as proportions and continuous variables as measures of central tendency. All continuous variables were subjected to a normality test. Mean and median values were reported for variables with normally distributed and skewed data, respectively.

The summary *t*-test was used to study the differences between mean "care" and "overprotection" scores of fathers and mothers. The independent sample *t*-test was used to study the factors associated with "care" and "overprotection" scores of fathers and mothers. Threshold for significance was $p = 0.05$.

Parental bonding quadrants
In addition to generating care and protection scores for each scale, parents can be effectively "assigned" to one of four quadrants:

"affectionate constraint" = high care and high protection	"affectionless control" = high protection and low care
"optimal parenting" = high care and low protection	"neglectful parenting" = low care and low protection

Assignment to "high" or "low" categories is based on the following cut-off scores:

- For **mothers**, a *care* score of 27.0 and a *protection* score of 13.5.
- For **fathers**, a *care* score of 24.0 and a *protection* score of 12.5.

FIGURE 1 | Assigining parenting styles using the PBI (Parker, 1979 #192).

The determinants of high grades including the influence of parenting styles were assessed using regression analysis. The outcome variable, student grades, was treated as binary (high grades and low grades). The threshold for statistical significance was $p = 0.05$. Crude Odds Ratios were adjusted for gender, school type, socioeconomic status, family structure, class grade, parents' employments and education status.

Ethics Statement

The study was approved by the Ethical Committee of the Khyber Medical University, Advance Studies and Research Board (KMU-AS&RB) in August 2016. Identifying information of students was not obtained. Permissions were obtained from the relevant authorities in the school administration before approaching the students and their parents. Written consent was obtained from the parents through the home-work diary of the students and verbal assent of each student was obtained.

RESULTS

Participants and Descriptive Data

A total of 456 students were interviewed, with 249 (54.6%) males and 207 (45.4%) females. The majority (52.5%) were students of grade 8. Despite including an equal number of public and private schools, 63.6% of the students belonged to a public sector school. The reason may be due to the larger class strength in public schools in comparison to private schools. The nuclear family structure was dominant (64.3%), with most students living in rented accommodation (70.4%) with 42.8% reporting to have obtained high grades (A plus or A) in their latest internal examinations (**Table 2**).

Majority of the students had both parents alive at the time of the interview. While all students' mothers were alive, 14 students reported their father to have passed away. Surprisingly, only 46% of the students were able to report their father's level of education compared to 99.5% for their mother. 9.2% of students reported their father to have an education level of grade 12 or above compared to 26% regarding their mother's qualification. This was in contrast to 90% of the fathers being employed compared to only 11% of the mothers (**Table 2**).

A Total of 257 (56%) students reported their mother to exhibit a high level of "care" vs. only 9 (2%) students reporting the same for their father. In terms of "overprotection," 343 (75%) and 296 (65%) students reported a high level for their father and mother, respectively. Based on combinations of these measures, the most common parenting style for both fathers (73%) and mothers (35%) was affectionless control and the least common for fathers was optimal parenting (0%) and neglectful parenting for mothers (9%). 121 (26%) students had both parents with the same parenting style, with 23% students having both parents show affectionless control and not a single student with both parents showing optimal parenting (**Figure 2**).

Determinants of High Grades

Our results show that high socioeconomic status [adjusted OR 2.78 (1.03, 7.52)], father's education level till undergrad or above [adjusted OR 4.58 (1.49, 14.09)], father's high "care" [adjusted OR 1.09 (1.01, 1.18)] and father's affectionless control style of parenting [adjusted OR 3.23 (1.30, 8.03)] are significant factors contributing to high grades (**Table 3**).

Differences in "Care" and "Overprotection" Between Fathers and Mothers

Care

The mean "care" score for mothers were significantly higher than fathers overall. The difference remained significant for male and female students, public and private schools, joint and nuclear family structures and low and high socioeconomic statuses (**Table 4**).

Overprotection

The mean "overprotection" score was significantly higher for fathers overall. The difference remained significant for female students, private schools, nuclear family structure, and low socioeconomic status. However, there was no significant difference in mean "overprotection" scores

TABLE 2 | Participant and descriptive data.

Participant characteristics (*n* = 456)			*N* (%)	Response rate (%)
Gender		Male	249 (54.6%)	100%
		Female	207 (45.4%)	
Class grade		Grade 8th	238 (52.2%)	100%
		Grade 9th	110 (24.1%)	
		Grade 10th	108 (23.7%)	
Type of school		Public	290 (63.6%)	100%
		Private	166 (36.4%)	
Father	Living status	Alive	442 (96.9%)	100%
		Deceased	14 (3.1%)	
	Education level	Masters (Post-graduate)	8 (1.8%)	45.6%
		Bachelors (Undergraduate)	7 (1.5%)	
		FSC/FA (Grade 12)	27 (5.9%)	
		Metric (Grade 10)	46 (10.1%)	
		Middle school (Grade 7)	18 (3.9%)	
		Below middle school or no education	102 (22.4%)	
	Employment status	Employed	418 (91.7%)	100%
		Unemployed*	38 (8.3%)	
	Type of employment	Government job	174 (38.2%)	99%
		Private	176 (38.6%)	
		Business	57 (12.5%)	
		Don't know	49 (10.7%)	
Mother	Living status	Alive	456 (100%)	100%
		Deceased	0 (0%)	
	Education level	Masters (Post-graduate)	35 (7.7%)	99.5%
		Bachelors (Undergraduate)	40 (8.8%)	
		FSC/FA (Grade 12)	44 (9.6%)	
		Metric (Grade 10)	56 (12.3%)	
		Middle school (Grade 7)	37 (8.1%)	
		Below middle school or no education	244 (53.5%)	
	Employment status	Employed	51 (11.2%)	100%
		Unemployed	405 (88.8%)	
Family structure		Nuclear family	293 (64.3%)	100%
		Joint family	163 (35.7%)	
Accommodation status		Rented	321 (70.4%)	100%
		Owned	135 (29.6%)	
Academic performance		High grades (A+ and A Grades)	195 (42.8%)	100%
		Low grades (Grades B, C, and D)	261 (57.2%)	

**Includes fathers who are deceased.*

between fathers and mothers for male students, public schools, joint family structures and high socioeconomic status (**Table 4**).

Factors Associated With "Care" and "Overprotection" in Fathers and Mothers

Fathers

The mean "care" score was significantly higher for fathers as reported by children in public schools and with higher grades. There was no significant difference in mean care scores based on student gender, socioeconomic status or family structure (**Table 5**).

For "overprotection" the only factor associated with a significantly higher mean score was "high" grades (**Table 5**).

Mothers

A significantly higher mean "care" score for mothers was reported by female students and students in public schools. No significant differences were observed for the other factors (**Table 5**).

A significantly higher mean "overprotection" score was reported by male students, students in public schools and those with "high" grades for mothers (**Table 5**).

DISCUSSION

Summary of Findings

Results of regression analysis show that socioeconomic status, father's education level and fathers' care scores have a significantly

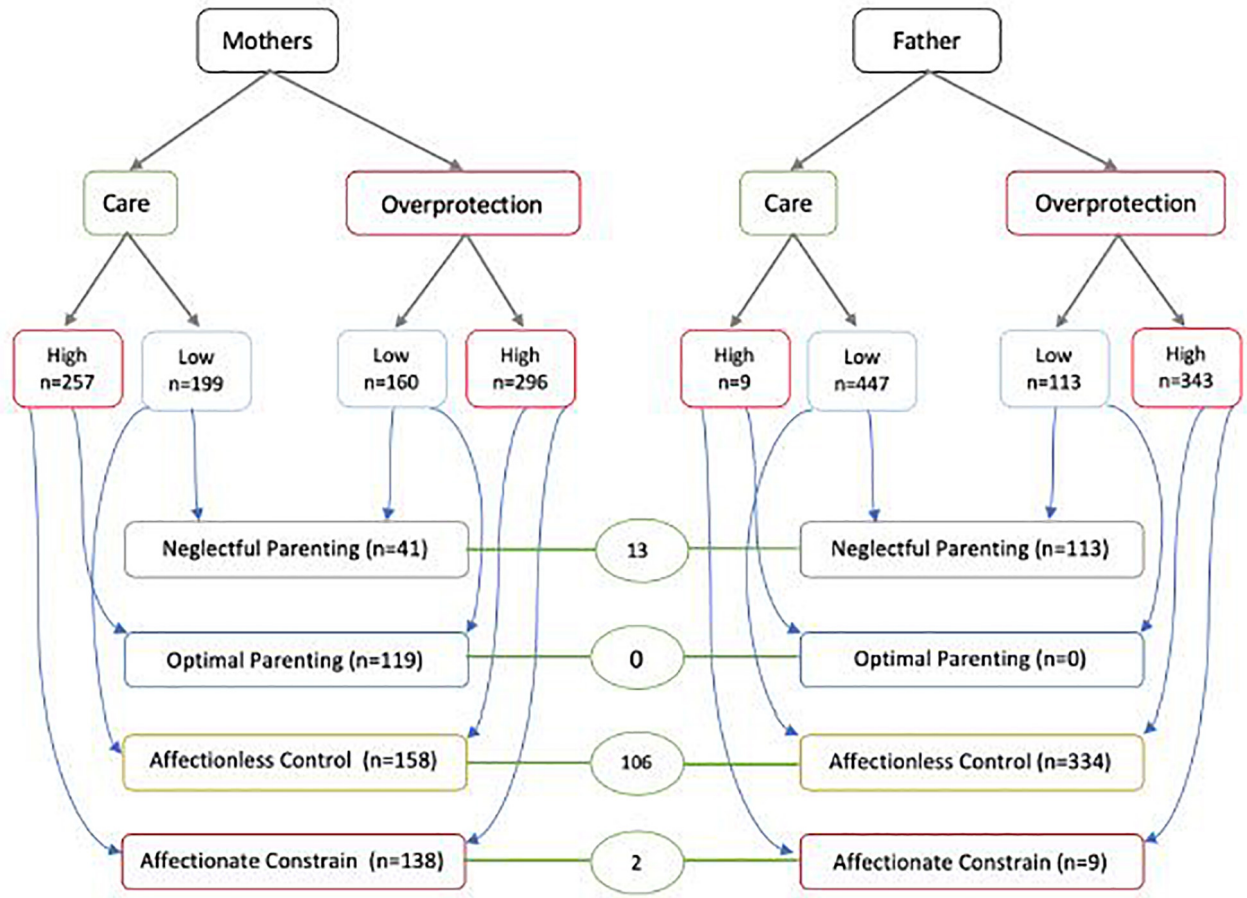


FIGURE 2 | "Care," "overprotection" and parenting styles for fathers and mothers as reported by students (n = 456). Green circles represent students with both parents showing the same parenting style – none of the students received "Optimal parenting" from both parents while 106 students received affectionless control from both parents.

positive influence on the academic performance of adolescent students in Peshawar, Pakistan. The most common parenting style for both fathers and mothers was affectionless control. However, affectionless control exhibited by the father was the only parenting style significantly contributing to improved academic performance.

Overall, the mean "care" score was higher for mothers and the mean "overprotection" score was higher for fathers. However, differences in "overprotection" were eliminated for male students, public schooling, joint family structures and high socioeconomic status.

Public schooling was associated with a significantly higher mean "care" score for both fathers and mothers and a significantly higher mean "overprotection" score for mothers. High grades were associated with a significantly higher mean "overprotection" score for both fathers and mothers and a significantly higher mean "care" score for fathers. For mothers, female students reported a significantly higher mean care score and male students reported a significantly higher mean "overprotection" score.

An additional interesting finding from the results of the study was that only about half the students were able to report their father's level of education compared to almost a 100% for their mother. From amongst those who did report, less than 10% of the father's had an education level equal or above grade 12 compared to a quarter of the mothers. However, only 11% of the mothers were employed in contrast to 90% of the fathers.

Previous Literature and Comparison of Main Findings

The results of our study have identified socioeconomic status, father's education level and high care scores for fathers to be significant predictors of academic success in adolescent students. Previous literature has shown socioeconomic status to be a predictor of academic success (Gamoran, 1996; Sander, 1999; Lubienski and Lubienski, 2006).

Parental education has been frequently associated with improved academic performance (Dumka et al., 2008; Dubow et al., 2009; Masud et al., 2015). In 2011, a study by Farooq et al. described the factors affecting academic performance in 600 students at the secondary school level in a public school in Lahore, Pakistan. Results of their study also associate parental

TABLE 3 | Academic performance: Determinants of "high" grades in the latest internal examinations.

Variables		Crude Odds Ratio (95% CI)	*p-value*	Adjusted* Odds Ratio (95% CI)	*p-value*
Gender	Male	1	0.215	1	0.156
	Female	0.79 (0.53, 1.15)		1.77 (0.80, 3.91)	
School type	Private	1	<0.001	1	0.062
	Public	5.02 (3.22, 7.85)		2.15 (0.96, 4.80)	
Class grade	Grade 8	1	0.019	1	0.381
	Grade 9	1.61 (1.02, 2.55)		1.93 (0.75, 4.94)	
	Grade 10	1.80 (1.14, 2.85)		1.03 (0.37, 2.84)	
SES	Low	1	0.723	1	**0.045**
	High	1.10 (0.65, 1.87)		2.78 (1.03, 7.52)	
Family structure	Joint family	1	0.650	1	0.473
	Nuclear family	0.91 (0.62, 1.35)		0.78 (0.39, 1.55)	
Father's education level	Completed middle school or below	1	<0.001	1	**0.023**
	Completed metric or FSC/FA	2.77 (1.40, 5.45)		1.95 (0.90, 4.22)	
	Completed undergrad or Post-grad	10.00 (4.20, 23.83)		4.58 (1.49, 14.09)	
Mother's education level	Completed middle school or below	1	<0.001	1	0.066
	Completed metric or FSC/FA	3.03 (1.89, 4.86)		2.98 (0.97, 9.17)	
	Completed undergrad or Post-grad	2.19 (1.31, 3.66)		4.25 (0.94, 19.29)	
Father's care	Low care	1	<0.001	1	**0.024**
	High care	1.09 (1.04, 1.14)		1.09 (1.01, 1.18)	
Father's overprotection	Low overprotection	1	0.028	1	0.420
	High overprotection	1.04 (1.01, 1.07)		1.02 (0.97, 1.08)	
Mother's care	Low care	1	0.774	1	0.920
	High care	0.99 (0.96, 1.03)		1.01 (0.93, 1.08)	
Mother's overprotection	Low overprotection	1	<0.001	1	0.098
	High overprotection	1.07 (1.03, 1.11)		1.06 (0.99, 1.13)	
Father's employment status	Unemployed	1	0.442	1	0.308
	Employed	1.31 (0.66, 2.60)		0.55 (0.17, 1.74)	
Mother's employment status	Unemployed	1	0.370	1	0.790
	Employed	1.33 (0.74, 2.39)		0.85 (0.27, 2.72)	
Father's parenting style	Neglectful parenting	1		1	
	Optimal parenting	–	–	–	–
	Affectionless control	1.74 (1.11, 2.72)	0.016	3.23 (1.30, 8.03)	**0.012**
	Affectionate constrain	2.57 (0.65, 10.13)	0.178	1.07 (0.16, 7.04)	0.941
Mother's parenting style	Neglectful parenting	1	0.007	1	
	Optimal parenting	1.43 (0.65, 3,15)	0.370	1.04 (0.24, 4.54)	0.957
	Affectionless control	2.28 (1.07, 4.88)	0.033	0.99 (0.21, 4.62)	0.990
	Affectionate constrain	2.89 (1.34, 6.23)	0.007	2.93 (0.68, 12.62)	0.150

*SES, socioeconomic status. *Adjusted for gender, class grade, school type, socioeconomic status, family structure, father, and mother education and employment status. The significant values are indicated in bold.*

education level with academic success in students. However, their results are significant for the education level of the mother as well as the father. Additionally, they also reported significantly higher academic performance in females and in students belonging to a higher socioeconomic status, factors not significant in our study (Farooq et al., 2011). Differences may be explained by cultural variations in Lahore and Peshawar within Pakistan, which should be explored further.

The description of parenting styles and behaviors has evolved over the years. With some variation in terminologies, the essence lies in a few common principles. Diana Baumrind initially described three main parenting styles based on variations in normal parenting behaviors: authoritative, authoritarian and permissive (Baumrind, 1966, 1967). Building on the concepts put forth by Baumrind, Maccoby and Martin identified two dimensions, "responsiveness" and "demandingness," which could classify parenting styles into 4 types, three of those described by Baumrind with the addition of neglectful parenting (Maccoby et al., 1983). The two dimensions, "responsiveness" and "demandingness," often referred to as "warmth" and "control" in literature (Lamborn et al., 1991; Tagliabue et al., 2014), are similar to the two measures, "care" and "overprotection" assessed by the parental bonding instrument (Parker et al., 1979; Parker, 1989; Dudley and Wisbey, 2000). Based on this, the authoritative, authoritarian, permissive and neglectful parenting styles described by Baumrind and Maccoby are similar to

TABLE 4 | Differences between mean "care" and "overprotection" scores between fathers and mothers.

			Father	Mother (*n* = 456)	*p*-value
Care	**Overall** (*n* = 456)	High	9 (2%)	257 (56.4%)	
		Low	447 (98%)	199 (43.6%)	
		Mean score	**15.06 ± 4.42**	**26.42 ± 4.84**	**<0.001**
	Male students (*n* = 249)	High	6 (2.4%)	116 (46.6%)	
		Low	243 (97.6%)	133 (53.4%)	
		Mean score	**14.84 ± 4.58**	**25.35 ± 4.88**	**<0.001**
	Female students (*n* = 207)	High	3 (1.4%)	141 (68.1%)	
		Low	204 (98.6%)	66 (31.9%)	
		Mean score	**15.32 ± 4.22**	**27.70 ± 4.48**	**<0.001**
	Private school (*n* = 166)	High	1 (0.6%)	110 (66.3%)	
		Low	165 (99.4%)	56 (33.7%)	
		Mean score	**14.06 ± 4.60**	**27.45 ± 4.48**	**<0.001**
	Public school (*n* = 290)	High	8 (2.8%)	147 (50.7%)	
		Low	282 (97.2%)	143 (49.3%)	
		Mean score	**15.63 ± 4.22**	**25.83 ± 4.95**	**<0.001**
	Nuclear family (*n* = 293)	High	6 (2%)	160 (54.6%)	
		Low	287 (98%)	133 (45.4%)	
		Mean score	**15.25 ± 4.30**	**26.45 ± 4.62**	**<0.001**
	Joint family (*n* = 163)	High	3 (1.8%)	97 59.5%)	
		Low	160 (98.2%)	66 (40.5%)	
		Mean score	**14.73 ± 4.63**	**26.37 ± 5.24**	**<0.001**
	Low SES (*n* = 388)	High	8 (2.1%)	220 (56.7%)	
		Low	380 (97.9%)	168 (43.3%)	
		Mean score	**14.99 ± 4.40**	**26.47 ± 4.90**	**<0.001**
	High SES (*n* = 65)	High	1 (1.5%)	36 (55.4%)	
		Low	64 (98.5%)	29 (44.6%)	
		Mean score	**15.57 ± 4.61**	**26.14 ± 4.53**	**<0.001**
Overprotection	**Overall** (*n* = 457)	High	343 (75.2%)	296 (64.9%)	
		Low	113 (24.8%)	160 (35.1%)	
		Mean score	**16.79**	**15.41**	**<0.001**
	Male students (*n* = 249)	High	181 (72.7%)	188 (75.5%)	
		Low	68 (27.3%)	61 (24.5%)	
		Mean score	**16.84 ± 6.39**	**16.87 ± 5.30**	0.957
	Female students (*n* = 207)	High	162 (78.3%)	108 (52.2%)	
		Low	45 (21.7%)	99 (47.8%)	
		Mean score	**16.72 ± 5.22**	**13.67 ± 5.59**	**<0.001**
	Private school (*n* = 166)	High	124 (74.7%)	75 (45.2%)	
		Low	42 (25.3%)	91 (54.8%)	
		Mean score	**16.62 ± 5.93**	**13.12 ± 5.42**	**<0.001**
	Public school (*n* = 290)	High	219 (75.5%)	221 (76.2%)	
		Low	71 (24.5%)	69 23.8%)	
		Mean score	**16.88 ± 5.86**	**16.73 ± 5.37**	0.748
	Nuclear family (*n* = 293)	High	228 (77.8%)	188 (64.2%)	
		Low	65 (22.2%)	105 (35.8%)	
		Mean score	**16.87 ± 5.71**	**15.45 ± 5.53**	**0.002**
	Joint family (*n* = 163)	High	115 (70.6%)	108 (66.3%)	
		Low	48 (29.4%)	55 (33.7%)	
		Mean score	**16.64 ± 6.19**	**15.35 ± 5.89**	0.055
	Low SES (*n* =)	High	289 (74.5%)	249 (64.2%)	
		Low	99 (25.5%)	139 (35.8%)	
		Mean score	**16.82 ± 5.85**	**15.36 ± 5.77**	**<0.001**
	High SES (*n* =)	High	51 (78.5%)	44 (67.7%)	
		Low	14 (21.5%)	21 (32.3%)	
		Mean score	**16.69 ± 6.22**	**15.66 ± 5.06**	0.302

SES: socioeconomic status. The significant values are indicated in bold.

TABLE 5 | Factors associated with "care" and "overprotection" for mothers and fathers.

	Variables	Fathers	*p*-value	Mothers	*p*-value
Care	Male students	14.84 ± 4.58	0.244	25.35 ± 4.88	**<0.001**
	Female students	15.32 ± 4.22		27.70 ± 4.48	
	Private school	14.06 ± 4.60	**0.001**	27.45 ± 4.48	**<0.001**
	Public school	15.63 ± 4.22		25.83 ± 4.95	
	Nuclear family	15.25 ± 4.30	0.230	26.45 ± 4.62	0.872
	Joint family	14.73 ± 4.63		26.37 ± 5.24	
	Low SES	14.99 ± 4.40	0.334	26.47 ± 4.90	0.609
	High SES	15.57 ± 4.61		26.14 ± 4.53	
	Low grades	14.36 ± 4.45	**<0.001**	26.48 ± 4.63	0.745
	High grades	16.00 ± 4.21		26.33 ± 5.12	
Overprotection	Male students	16.84 ± 6.39	0.833	16.87 ± 5.30	**<0.001**
	Female students	16.72 ± 5.22		13.67 ± 5.59	
	Private school	16.62 ± 5.93	0.647	13.12 ± 5.42	**<0.001**
	Public school	16.88 ± 5.86		16.73 ± 5.37	
	Nuclear family	16.87 ± 5.71	0.699	15.45 ± 5.53	0.855
	Joint family	16.64 ± 6.19		15.35 ± 5.89	
	Low SES	16.82 ± 5.85	0.872	15.36 ± 5.77	0.693
	High SES	16.69. ± 6.22		15.66 ± 5.06	
	Low grades	16.26 ± 5.95	**<0.027**	14.55 ± 5.61	**<0.001**
	High grades	17.49 ± 5.79		16.57 ± 5.52	

SES, socioeconomic status. The significant values are indicated in bold.

the affectionate constraint, affectionless control, optimal, and neglectful styles as classified by the parental bonding instrument, respectively (Baumrind, 1991; Cavedo and Parker, 1994).

Results of our study show that affectionless control, similar to the authoritarian style of parenting, adapted by the father is significantly associated with improved academic performance. This differs from the popularity of the authoritative parenting style, similar to affectionate constraint, in determining academic success in literature from western cultures (Steinberg et al., 1989, 1992; Deslandes et al., 1998; Aunola et al., 2000; Adeyemo, 2005; Masud et al., 2015; Pinquart, 2016; Checa et al., 2019). Evidence from societies with cultural similarities with Pakistan presents varied findings. A study from Iran shows support for the authoritarian parenting style similar to our study (Rahimpour et al., 2015). A review of 39 studies published by Masud et al. (2015) in 2015 assesses the effect of parenting styles on academic performance (Masud et al., 2015 #205). The review very aptly described how the authoritative parenting style is the dominant and most effective style in terms of determining academic performance in the West and European countries while Asian cultures show more promising results for academic success for the authoritarian style (Dornbusch et al., 1987; Lin and Fu, 1990; Masud et al., 2015). The results of our study are in synchrony with these findings. However, our results also show that high father's "care" scores are significant contributors to higher academic grades. Since no father showed optimal parenting and only 9 fathers had affectionate constraint, both parenting styles with high care scores, these results may be a reflection of the importance of father's role in determining academic performance in Asian cultures. Findings supporting the authoritarian/affectionless control style may be due to the abundance of this parenting style. Perhaps a fairer comparison may be possible with a larger sample population with fathers showing all types of parenting styles equally.

Interpretation and Explanation of Other Findings

Observations of factors associated with and differences in "care" and "overprotection" between fathers and mothers may be attributed to reverse causality and should be used as hypothesis generating.

Our results show that mothers have higher mean "care" score and fathers have a higher mean "overprotection" score. Since these scores are based on perceptions of the child, part of these observations may be explained by the cultural norms of expression of love and concern by fathers and mothers. With the difference in "overprotection" being eliminated for male and female children, it is possible that mothers are more overprotective of their sons. Male gender preference in Pakistan may be an explanation for this (Qadir et al., 2011).

Our results show lower employment rates for women despite higher education levels. The finding of higher education levels for females compared to males does not agree with national data, which reports findings from rural areas as well where education opportunities are limited for females (Hussain, 2005; Chaudhry and Rahman, 2009). Our results provide a zoomed in look at an urban population, which may have progressed enough to improve women's education but cultural norms, gender discrimination and lack of opportunity still prevent women from stepping into the workface (Chaudhry, 2007; Begum and Sheikh, 2011).

Implications and Future Direction

The findings of our study may have implications for future research and policy making.

Affectionless control is associated with improved academic performance but further research investigating the effects of this style on other aspects of child development, particularly emotional and psychological health, is needed. Factors affecting care and overprotection need to be studied in more detail so that parenting workshops and interventions are tailored to our population. Results also suggest that fathers should play a stronger role in parenting of adolescent students. School policies should make it mandatory for both parents to attend parent-teacher meetings and assigned home activities should include both parents.

Limitations

Since the study is based on the urban population of Peshawar, results may not be generalizable to the adolescent students of the country which includes large rural populations. Academic performance was judged on latest internal examinations, the marking criteria for which may vary across schools. The use of external examinations would have standardized grades across schools but limited the sample to students of grade 9 and 10.

Conclusion

Our study concludes that socioeconomic status, father's level of education and high care scores for fathers are associated with improved academic outcomes in adolescent students in Peshawar, Pakistan. Affectionless control is the most common parenting style as perceived by the students and when adapted by the father, contributes to better grades. Further research investigating the effects of demonstrating affectionless control on the emotional and psychological health of students needs to be conducted. Parenting workshops and school policies should include recommendations to increase involvement of fathers in the parenting of adolescent children.

DATA AVAILABILITY STATEMENT

Data collected and stored as part of this study is available upon reasonable request.

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

SM contributed in conceiving, designing, data acquisition, grant submission, and manuscript review. SHM involved in data analysis and manuscript writing. NQ involved in manuscript writing. MK was the principal investigator and supervisor for the project. FK and SK contributed in literature review and data management. All authors proofread and agreed on the final draft and accept responsibility for the work.

FUNDING

This project was graciously funded by the Research Promotion and Development World Health Organization Regional Office for the Eastern Mediterranean (RPPH Grant 2016-2017, TSA reference: 2017/719467-0).

ACKNOWLEDGMENTS

The authors thank Dr. Nazish Masud (King Saud bin Abdulaziz University), and Dr. Khabir Ahmad and Dr. Bilal Ahmad (The Aga Khan University) for their contributions to the project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02497/full#supplementary-material

REFERENCES

Abidin, R. R. (1992). The determinants of parenting behavior. *J. Clin. Child Psychol.* 21, 407–412. doi: 10.1207/s15374424jccp2104_12

Adeyemo, D. A. (2005). Parental involvement, interest in schooling and school environment as predictors of academic self-efficacy among fresh secondary school students in Oyo State, Nigeria. *Electron. J. Res. Educ. Psychol.* 3, 163–180.

Aunola, K., Stattin, H., and Nurmi, J.-E. (2000). Parenting styles and adolescents' achievement strategies. *J. Adolesc.* 23, 205–222. doi: 10.1006/jado.2000.0308

Battle, J. (2002). Longitudinal analysis of academic achievement amonga nationwide sample of hispanic students in one-versus dual-parent households. *Hisp. J. Behav. Sci.* 24, 430–447. doi: 10.1177/0739986302238213

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Dev.* 37, 887–907. doi: 10.1111/j.1467-8624.1966.tb05416.x

Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behavior. *Genet. Psychol. Monogr.* 75, 43–88.

Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance use. *J. Early Adolesc.* 11, 56–95. doi: 10.1177/0272431691111004

Bean, R. A., Barber, B. K., and Crane, D. R. (2006). Parental support, behavioral control, and psychological control among African American youth: the relationships to academic grades, delinquency, and depression. *J. Fam. Issues* 27, 1335–1355. doi: 10.1177/0192513X06289649

Beauvais, A. M., Stewart, J. G., DeNisco, S., and Beauvais, J. E. (2014). Factors related to academic success among nursing students: a descriptive correlational research study. *Nurse Educ. Today* 34, 918–923. doi: 10.1016/j.nedt.2013.12.005

Begum, M. S., and Sheikh, Q. A. (2011). Employment situation of women in Pakistan. *Int. J. Soc. Econ.* 38, 98–113. doi: 10.1108/03068291111091981

Behzadi, B., and Parker, G. (2015). A Persian version of the parental bonding instrument: factor structure and psychometric properties. *Psychiatry Res.* 225, 580–587. doi: 10.1016/j.psychres.2014.11.042

Belsky, J. (1984). The determinants of parenting: a process model. *Child Dev.* 55, 83–96. doi: 10.2307/1129836

Bluestone, C., and Tamis-LeMonda, C. S. (1999). Correlates of parenting styles in predominantly working- and middle-class African American mothers. *J. Marriage Fam.* 61, 881–893. doi: 10.2307/354010

Cavedo, L., and Parker, G. (1994). Parental bonding instrument. *Soc. Psychiatr. Psychiatr. Epidemiol.* 29, 78–82.

Chaudhry, I. S. (2007). Gender inequality in education and economic growth: case study of Pakistan. *Pakistan Horizon* 60, 81–91.

Chaudhry, I. S., and Rahman, S. (2009). The impact of gender inequality in education on rural poverty in Pakistan: an empirical analysis. *Eur. J. Econ. Finance Adm. Sci.* 15, 174–188.

Checa, P., and Abundis-Gutierrez, A. (2017). Parenting and temperament influence on school success in 9–13 year olds. *Front. Psychol.* 8:543. doi: 10.3389/fpsyg.2017.00543

Checa, P., Abundis-Gutierrez, A., Pérez-Dueñas, C., and Fernández-Parra, A. (2019). Influence of maternal and paternal parenting style and behavior problems on academic outcomes in primary school. *Front. Psychol.* 10:378. doi: 10.3389/fpsyg.2019.00378

Chow, H. P. (2000). The determinants of academic performance: Hong Kong immigrant students in Canadian schools. *Can. Ethn. Stud. J.* 32, 105–105.

Cohen, D. A., and Rice, J. (1997). Parenting styles, adolescent substance use, and academic achievement. *J. Drug Educ.* 27, 199–211. doi: 10.2190/QPQQ-6Q1G-UF7D-5UTJ

Cutrona, C. E., Cole, V., Colangelo, N., Assouline, S. G., and Russell, D. W. (1994). Perceived parental social support and academic achievement: an attachment theory perspective. *J. Pers. Soc. Psychol.* 66, 369–378. doi: 10.1037/0022-3514.66.2.369

Dennis, J. M., Phinney, J. S., and Chuateco, L. I. (2005). The role of motivation, parental support, and peer support in the academic success of ethnic minority first-generation college students. *J. Coll. Stud. Dev.* 46, 223–236. doi: 10.1353/csd.2005.0023

Desimone, L. (1999). Linking parent involvement with student achievement: do race and income matter? *J. Educ. Res.* 93, 11–30. doi: 10.1080/00220679909597625

Deslandes, R., Bouchard, P., and St-Amant, J.-C. (1998). Family variables as predictors of school achievement: sex differences in Quebec adolescents. *Can. J. Educ.* 23, 390–404.

Deslandes, R., Royer, E., Turcotte, D., and Bertrand, R. (1997). School achievement at the secondary level: influence of parenting style and parent involvement in schooling. *McGill J. Educ.* 32,

Dornbusch, S. M., Ritter, P. L., Leiderman, P. H., Roberts, D. F., and Fraleigh, M. J. (1987). The relation of parenting style to adolescent school performance. *Child Dev.* 58, 1244–1257. doi: 10.1111/j.1467-8624.1987.tb01455.x

Dubow, E. F., Boxer, P., and Huesmann, L. R. (2009). Long-term effects of parents' education on children's educational and occupational success: mediation by family interactions, child aggression, and teenage aspirations. *Merrill Palmer Q.* 55, 224–249. doi: 10.1353/mpq.0.0030

Dudley, R. L., and Wisbey, R. L. (2000). The relationship of parenting styles to commitment to the church among young adults. *Relig. Educ.* 95, 38–50. doi: 10.1080/0034408000950105

Dumka, L. E., Gonzales, N. A., Bonds, D. D., and Millsap, R. E. (2008). Academic success of Mexican origin adolescent boys and girls: the role of mothers' and fathers' parenting and cultural orientation. *Sex Roles* 60, 588–599. doi: 10.1007/s11199-008-9518-z

Farooq, M. S., Chaudhry, A. H., Shafiq, M., and Berhanu, G. (2011). Factors affecting students' quality of academic performance: a case of secondary school level. *J. Qual. Technol. Manag.* 7, 1–14.

Farsides, T., and Woodfield, R. (2003). Individual differences and undergraduate academic success: the roles of personality, intelligence, and application. *Pers. Individ. Differ.* 34, 1225–1243. doi: 10.1016/S0191-8869(02)00111-3

Flynn, D. M., and MacLeod, S. (2015). Determinants of Happiness in Undergraduate University Students. *Coll. Stud. J.* 49, 452–460.

Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. *Educ. Eval. Policy Anal.* 18, 1–18. doi: 10.3102/01623737018001001

Ginsburg, G. S., and Bronstein, P. (1993). Family factors related to children's intrinsic/extrinsic motivational orientation and academic performance. *Child Dev.* 64, 1461–1474. doi: 10.1111/j.1467-8624.1993.tb02964.x

Hawkins, J. D. (1997). "Academic performance and school success: sources and consequences," in *Healthy Children 2010: Enhancing Children's Wellness*, eds R. P. Weissberg, T. P. Gullotta, R. L. Hampton, B. A. Ryan, and G. R. Adams, (Thousand Oaks, CA: Sage Publications, Inc), 278–305.

Hillstrom, K. A. (2009). *Are Acculturation and Parenting Styles Related to Academic Achievement Among Latino students?* dissertation, University of Southern California, Los Angeles, CA.

Hussain, I. (2005). "Education, employment and economic development in Pakistan," in *Education Reform in Pakistan: Building for the Future*, ed. R. M. Hathaway, (Washington, DC: Woodrow Wilson International Center for Scholars), 33–45.

Huver, R. M. E., Otten, R., de Vries, H., and Engels, R. C. (2010). Personality and parenting style in parents of adolescents. *J. Adolesc.* 33, 395–402. doi: 10.1016/j.adolescence.2009.07.012

Jeynes, W. H. (2007). The relationship between parental involvement and urban secondary school student academic achievement: a meta-analysis. *Urban Educ.* 42, 82–110. doi: 10.1177/0042085906293818

Khalid, A., Qadir, F., Chan, S. W., and Schwannauer, M. (2018). Parental bonding and adolescents' depressive and anxious symptoms in Pakistan. *J. Affect. Disord.* 228, 60–67. doi: 10.1016/j.jad.2017.11.050

Kuh, G. D., Kinzie, J. L., Buckley, J. A., Bridges, B. K., and Hayek, J. C. (2006). *What Matters to Student Success: A Review of the Literature*. Washington, DC: National Postsecondary Education Cooperative.

Lamborn, S. D., Mounts, N. S., Steinberg, L., and Dornbusch, S. M. (1991). Patterns of competence and adjustment among adolescents from authoritative, authoritarian, indulgent, and neglectful families. *Child Dev.* 62, 1049–1065. doi: 10.1111/j.1467-8624.1991.tb01588.x

Lin, C. Y. C., and Fu, V. R. (1990). A comparison of child-rearing practices among Chinese, immigrant Chinese, and Caucasian-American parents. *Child Dev.* 61, 429–433. doi: 10.1111/j.1467-8624.1990.tb02789.x

Linnenbrink, E. A., and Pintrich, P. R. (2002). Motivation as an enabler for academic success. *Sch. Psychol. Rev.* 31, 313–327. doi: 10.1080/17483107.2018.1471169

Liu, J., Li, L., and Fang, F. (2011). Psychometric properties of the Chinese version of the Parental Bonding Instrument. *Int. J. Nurs. Stud.* 48, 582–589. doi: 10.1016/j.ijnurstu.2010.10.008

Lubienski, C., and Lubienski, S. (2006). *Charter, Private, Public Schools and Academic Achievement: New Evidence from NAEP Mathematics Data*. New York, NY: Columbia University.

Maccoby, E., Martin, J., Hetherington, E., and Mussen, P. (1983). "Socialization in the context of the family: Parent-child interaction," in *Handbook of Child Psychology: Socialization, Personality, and Social Development*, 4th Edn, Vol. 4, ed. E. M. Hetherington, (Hoboken. NJ: John Wiley & Sons), 101.

Masud, H., Thurasamy, R., and Ahmad, M. S. (2015). Parenting styles and academic achievement of young adolescents: a systematic literature review. *Qual. Quant.* 49, 2411–2433. doi: 10.1007/s11135-014-0120-x

McCabe, J. E. (2014). Maternal personality and psychopathology as determinants of parenting behavior: a quantitative integration of two parenting literatures. *Psychol. Bull.* 140, 722–750. doi: 10.1037/a0034835

McClelland, M. M., Morrison, F. J., and Holmes, D. L. (2000). Children at risk for early academic problems: the role of learning-related social skills. *Early Child. Res. Q.* 15, 307–329. doi: 10.1016/S0885-2006(00)00069-7

Mujtaba, T., and Furnham, A. (2001). A cross-cultural study of parental conflict and eating disorders in a non-clinical sample. *Int. J. Soc. Psychiatry* 47, 24–35. doi: 10.1177/002076400104700103

Murphy, E., Wickramaratne, P., and Weissman, M. (2010). The stability of parental bonding reports: a 20-year follow-up. *J. Affect. Disord.* 125, 307–315. doi: 10.1016/j.jad.2010.01.003

Osorio, A., and González-Cámara, M. (2016). Testing the alleged superiority of the indulgent parenting style among Spanish adolescents. *Psicothema* 28, 414–420.

Östberg, M., and Hagekull, B. (2000). A structural modeling approach to the understanding of parenting stress. *J. Clin. Child Psychol.* 29, 615–625. doi: 10.1207/S15374424JCCP2904_13

Parker, G. (1979). Reported parental characteristics in relation to trait depression and anxiety levels in a non-clinical group. *Aust. N. Z. J. Psychiatry* 13, 260–264. doi: 10.3109/00048677909159146

Parker, G. (1983). *Parental Overprotection: A Risk Factor in Psychosocial Development*. New York, NY: Grune & Stratton.

Parker, G. (1989). The parental bonding instrument: psychometric properties reviewed. *Psychiatr. Dev.* 7, 317–335.

Parker, G. (1990). The parental bonding instrument. *Soc. Psychiatr. Psychiatr. Epidemiol.* 25, 281–282.

Parker, G., Tupling, H., and Brown, L. B. (1979). A parental bonding instrument. *Br. J. Med. Psychol.* 52, 1–10.

Parsasirat, Z., Montazeri, M., Yusooff, F., Subhi, N., and Nen, S. (2013). The most effective kinds of parents on children's academic achievement. *Asian Soc. Sci.* 9, 229–242.

Pinquart, M. (2016). Associations of parenting styles and dimensions with academic achievement in children and adolescents: a meta-analysis. *Educ. Psychol. Rev.* 28, 475–493. doi: 10.1007/s10648-015-9338-y

Ponnet, K., Mortelmans, D., Wouters, E., Van Leeuwen, K., Bastaits, K., and Pasteels, I. (2013). Parenting stress and marital relationship as determinants of mothers' and fathers' parenting. *Pers. Relat.* 20, 259–276. doi: 10.1111/j.1475-6811.2012.01404.x

Qadir, F., Khan, M. M., Medhin, G., and Prince, M. (2011). Male gender preference, female gender disadvantage as risk factors for psychological morbidity in Pakistani women of childbearing age - a life course perspective. *BMC Public Health* 11:745. doi: 10.1186/1471-2458-11-745

Qadir, F., Stewart, R., Khan, M., and Prince, M. (2005). The validity of the Parental Bonding Instrument as a measure of maternal bonding among young Pakistani women. *Soc. Psychiatr. Psychiatr. Epidemiol.* 40, 276–282. doi: 10.1007/s00127-005-0887-0

Rahimpour, P., Direkvand-Moghadam, A., Direkvand-Moghadam, A., and Hashemian, A. (2015). Relationship between the parenting styles and students' educational performance among Iranian girl high school students, a cross-sectional study. *J. Clin. Diagn. Res.* 9:JC05–JC07. doi: 10.7860/JCDR/2015/15981.6914

Sander, W. (1999). Private schools and public school achievement. *J. Hum. Resour.* 34, 697–709. doi: 10.2307/146413

Sanders, M. G. (1998). The effects of school, family, and community support on the academic achievement of African American adolescents. *Urban Educ.* 33, 385–409. doi: 10.1177/0042085998033003005

Shute, V. J., Hansen, E. G., Underwood, J. S., and Razzouk, R. (2011). A review of the relationship between parental involvement and secondary school students' academic achievement. *Educ. Res. Int.* 2011:915326.

Simons, R. L., Beaman, J., Conger, R. D., and Chao, W. (1993). Childhood experience, conceptions of parenting, and attitudes of spouse as determinants of parental behavior. *J. Marriage Fam.* 55, 91–106. doi: 10.2307/352961

Simons, R. L., Whitbeck, L. B., Conger, R. D., and Melby, J. N. (1990). Husband and wife differences in determinants of parenting: a social learning and exchange model of parental behavior. *J. Marriage Fam.* 52, 375–392. doi: 10.2307/353033

Smith, C. L. (2010). Multiple determinants of parenting: predicting individual differences in maternal parenting behavior with toddlers. *Parenting* 10, 1–17. doi: 10.1080/15295190903014588

Steinberg, L., Elmen, J. D., and Mounts, N. S. (1989). Authoritative parenting, psychosocial maturity, and academic success among adolescents. *Child Dev.* 60, 1424–1436. doi: 10.2307/1130932

Steinberg, L., Lamborn, S. D., Dornbusch, S. M., and Darling, N. (1992). Impact of parenting practices on adolescent achievement: authoritative parenting, school involvement, and encouragement to succeed. *Child Dev.* 63, 1266–1281. doi: 10.1111/j.1467-8624.1992.tb01694.x

Tagliabue, S., Olivari, M. G., Bacchini, D., Affuso, G., and Confalonieri, E. (2014). Measuring adolescents' perceptions of parenting style during childhood: psychometric properties of the parenting styles and dimensions questionnaire. *Psicol. Teoria e Pesquisa* 30, 251–258. doi: 10.1590/s0102-37722014000300002

Tomul, E., and Savasci, H. S. (2012). Socioeconomic determinants of academic achievement. *Educ. Assess. Eval. Account.* 24, 175–187. doi: 10.1007/s11092-012-9149-9143

Vacha, E. F., and McLaughlin, T. F. (1992). The social structural, family, school, and personal characteristics of at-risk students: policy recommendations for school personnel. *J. Educ.* 174, 9–25. doi: 10.1177/002205749217400303

Valiente, C., Lemery-Chalfant, K., and Castro, K. S. (2007). Children's effortful control and academic competence: mediation through school liking. *Merrill Palmer Q.* 53, 1–25. doi: 10.1353/mpq.2007.0006

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., et al. (2014). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Int. J. Surg.* 12, 1500–1524.

Wang, Q., and Leichtman, M. D. (2000). Same beginnings, different stories: a comparison of American and Chinese children's narratives. *Child Dev.* 71, 1329–1346. doi: 10.1111/1467-8624.00231

Weis, M., Trommsdorff, G., and Muñoz, L. (2016). Children's self-regulation and school achievement in cultural contexts: the role of maternal restrictive control. *Front. Psychol.* 7:722. doi: 10.3389/fpsyg.2016.00722

Weiss, L. H., and Schwarz, J. C. (1996). The relationship between parenting types and older adolescents' personality, academic achievement, adjustment, and substance use. *Child Dev.* 67, 2101–2114. doi: 10.1111/j.1467-8624.1996.tb01846.x

Wilhelm, K., Niven, H., Parker, G., and Hadzi-Pavlovic, D. (2005). The stability of the parental bonding instrument over a 20-year period. *Psychol. Med.* 35, 387–393. doi: 10.1017/s0033291704003538

York, T., Gibson, C., and Rankin, S. (2015). Defining and measuring academic success. *Pract. Assess. Res. Eval.* 20, 1–20.

Zahedani, Z. Z., Rezaee, R., Yazdani, Z., Bagheri, S., and Nabeiei, P. (2016). The influence of parenting style on academic achievement and career path. *J. Adv. Med. Educ. Prof.* 4, 130–134.

Zakeri, H., and Karimpour, M. (2011). Parenting styles and self-esteem. *Proc. Soc. Behav. Sci.* 29, 758–761. doi: 10.1016/j.sbspro.2011.11.302

Zhang, X., Hu, B. Y., Ren, L., Huo, S., and Wang, M. (2019). Young Chinese children's academic skill development: identifying child-, family-, and school-level factors. *New Dir. Child Adolesc. Dev.* 2019, 9–37. doi: 10.1002/cad.20271

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.

Copyright © 2019 Masud, Mufarrih, Qureshi, Khan, Khan and Khan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 15 November 2019
doi: 10.3389/fpsyg.2019.02583

Identification of Factors Associated With School Effectiveness With Data Mining Techniques: Testing a New Approach

*Fernando Martínez-Abad**

University Institute of Educational Sciences, University of Salamanca, Salamanca, Spain

The study of school effectiveness and the identification of factors associated with it are growing fields of research in the education sciences. Moreover, from the perspective of data mining, great progress has been made in the development of algorithms for the modeling and identification of non-trivial information from massive databases. This work, which falls within this context, proposes an innovative approach for the identification and characterization of educational and organizational factors associated with high school effectiveness. Under a perspective of basic research, our aim is to study the suitability of decision trees, techniques inherent to data mining, to establish predictive models for school effectiveness. Based on the available Spanish sample of the PISA 2015 assessment, an indicator of the school effectiveness was obtained from the application of multilevel models with predictor variables of a contextual nature. After selecting high- and low-effectiveness schools in this first phase, the second phase of the study was carried out and consisted of the application of decision trees to identify school, teacher, and student factors associated with high and low effectiveness. The C4.5 algorithm was calculated and, as a result, we obtained 120 different decision trees based on five determining factors (database used; stratification in the initial selection of schools; significance of the predictor variables of the models; use of items and/or scales; and use of the training or validated samples). The results show that the use of this kind of technique could be appropriate if mainly used with correctly pre-processed data that include the combined information available from all educational agents. This study represents a major breakthrough in the study of the factors associated with school effectiveness from a quantitative approach, since it proposes and provides a simple and appropriate procedure for modeling and establishing patterns. In doing so, it contributes to the development of knowledge in the field of school effectiveness that can help in educational decision-making.

Keywords: data mining, school effectiveness, academic achievement, large-scale assessment, decision trees

OPEN ACCESS

Edited by:
Iván Manuel Jorrín Abellán, Kennesaw State University, United States

Reviewed by:
Enrique Navarro, Complutense University of Madrid, Spain
Paul Seitlinger, Tallinn University, Estonia

***Correspondence:**
Fernando Martínez-Abad
fma@usal.es

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

Received: *30 August 2019*
Accepted: *31 October 2019*
Published: *15 November 2019*

Citation:
Martínez-Abad F (2019) Identification of Factors Associated With School Effectiveness With Data Mining Techniques: Testing a New Approach. Front. Psychol. 10:2583. doi: 10.3389/fpsyg.2019.02583

INTRODUCTION

Identification of educational factors associated with academic performance is a key aspect in educational research into school effectiveness (Rutter and Maughan, 2002; Murillo, 2007; Muijs et al., 2014; Creemers and Kyriakides, 2015). Within this context, we propose an innovative approach to the analysis of good educational practices associated with school effectiveness.

This study proposes the application of data mining techniques to identify the main factors that characterize and differentiate high- and low-effectiveness schools.

In contrast to traditionally used techniques (inferential and multivariate correlational statistics), data mining is not based on previous assumptions or theoretical distributions to obtain predictive models. In addition, these techniques are applied with minimal intervention by researchers, which, together with the aforementioned, represent a great advantage for the identification of valuable information in massive databases (Xu, 2005). More specifically, the algorithm proposed in this study is the decision tree (classification algorithm), since it simplifies the analysis and interpretation of the predictor variables and their relationships (Martínez-Abad and Chaparro-Caso-López, 2017).

The main aim of this study, therefore, is the analysis of the fit and predictive power of data mining techniques, specifically decision trees, for the identification of factors associated with school effectiveness in secondary education.

Given this main objective, we can set the following specific objectives:

- Analyze and identify school effectiveness based on cross-sectional data from large-scale assessments.
- Promote methodological alternatives for the study of factors associated with school effectiveness based on mass data.
- Analyze the effectiveness of decision trees (algorithm C4.5) in the study of the process factors associated with school effectiveness.
- Present the possibilities of decision trees for the study of good educational practices in effective schools.

Conceptual Framework

The publication of the *Equality of Educational Opportunity Study*, better known as the "Coleman Report" (Coleman et al., 1966), had a major impact on the educational research field. The study's main conclusions were that socio-economic and demographic conditions (contextual factors) provided a decisive explanation of the differences in academic performance between students and schools. It questioned the impact that educational practices carried out in schools could have on student performance. These hard-hitting results fueled an in-depth debate about the contribution of the education system and educational policies to the knowledge and skills acquired by students.

In response to the Coleman Report, the Effective School Movement (ESM) emerged during the 1980s (Lezotte, 1989; Martínez et al., 2009). The ESM began with the aim of identifying and studying the most effective school environments in order to define good educational practices associated with variables over which the education system has control. Since that time, school effectiveness has increased its presence in the educational research field, and today has an important impact on work and scientific dissemination internationally (Rutter and Maughan, 2002; Wrigley, 2013; Muijs et al., 2014; Creemers and Kyriakides, 2015; Chapman et al., 2016; Martínez-Garrido and Murillo, 2016).

In the 1990s, thanks mainly to improvement in the computing capacity of computer systems and to the widespread use of large-scale assessments, research into school effectiveness experienced strong growth and evolution (Chapman et al., 2016). Multivariate statistical analysis based on linear hierarchical models (also called multilevel models) emerged as the fundamental statistical technique for the identification and analysis of school effectiveness (Bryk and Raudenbush, 1992; Creemers and Scheerens, 1994; Goldstein, 1995). These models respect the nested nature of educational data from large-scale assessments and allow the identification of educational groups that show clearly higher or lower performance than expected, taking into account only contextual factors. In this way, based on the study of the residuals of the school in these contextual models, we are able to identify high- and low-effectiveness schools (Gamazo et al., 2018). While some research works propose the carrying out of a qualitative case study of schools identified as high- or low-effectiveness for the analysis of good educational practices (Murillo, 2007; Joaristi et al., 2014), others that are quantitative in nature study these factors by adding to the multilevel models the process variables of interest and analyzing their individual behavior and interaction with other variables (Cordero et al., 2015; Costa and Araújo, 2017; Pitsia et al., 2017; Tan and Hew, 2017). From a methodological point of view, this quantitative perspective faces some problems: the existence of previous assumptions in the analysis that are rarely met or that are not directly taken into account (homocedasticity, normality of distributions, inexistence of non-linear relationships, multicollineality, etc.); difficulties in the estimation of typical errors when dealing with excessively complex models (lack of parsimony) and/or with an excessive number of subjects; and difficulty in studying the multiple interactions between the predictor variables given their high complexity, or the impossibility of doing so when working with fixed-effects models to simplify the computation and interpretation of the data.

That is why this work is interested in proposing a quantitative alternative for the study of process variables associated with school effectiveness that does not have the above-mentioned limitations. Specifically, based on the perspective of educational data mining (EDM), we apply decision trees to establish the predictive models of high- and low-effectiveness schools that have a better fit to the data, and we analyze under which determining factors these techniques achieve better performance.

Literature Review

The current calculation capacity of computers allows the development and application of appropriate statistical techniques for the analysis of massive data. In this regard, data mining emerges as a set of techniques that add value to large-scale data analysis (Martínez-Abad and Chaparro-Caso-López, 2017). These techniques enable the identification of patterns in the data without proposing previous assumptions or starting models, and with minimal intervention by the researcher (Xu, 2005). Thus, the nature of some of the data mining algorithms, compared to other classic multivariate techniques, can promote significant progress in the identification of factors associated with

school effectiveness, guiding decision-making and the operation of the education system at macro, meso, and micro levels (Bronfenbrenner, 1979).

Despite the potential that these statistical techniques may hold, their use in the establishment of performance prediction models in compulsory education is sporadic (Hung et al., 2012; Oskouei and Askari, 2014; Şara et al., 2015), and their use for the exploration of large-scale assessments is extremely limited (Liu and Ruiz, 2008; Liu and Whitford, 2011; Kılıç et al., 2017; Asensio et al., 2018). A significant presence of EDM can, however, be observed in the study of performance in university education (Guruler et al., 2010; Kasih et al., 2013; Romero et al., 2013; Kirby and Dempster, 2014; Akçapinar et al., 2015; Tan and Shao, 2015; Asif et al., 2017; Casey and Azcona, 2017; Costa et al., 2017).

Given the characteristics of the statistical techniques of data mining, many of these works focused on non-university levels propose a dichotomous variable as a criterion variable in their models, referring to whether a student reaches a minimum performance (Costa et al., 2017; Kılıç et al., 2017) or if he or she abandons his or her studies (Şara et al., 2015). In this regard, although there are data mining techniques that allow the use of ordinal or quantitative criterion variables, and that this dichotomization causes a reduction in the information contained in the original variable (Jacobucci, 2018), the inclusion of dichotomous criterion variables promotes the obtaining of decision trees that have simpler structures and are, therefore, more easily interpretable.

As for the statistical technique applied, although numerous works are carried out with the aim of comparing classification algorithms (Jamain and Hand, 2008; Oskouei and Askari, 2014; Yu et al., 2014; Akçapinar et al., 2015; Asif et al., 2017; Costa et al., 2017; Kılıç et al., 2017), there is no agreement in the scientific community on which ones are more appropriate for the prediction of academic performance. "The literature review suggests that in general there is no single classifier that works best in all contexts to provide good prediction" (Kirby and Dempster, 2014).

In this regard, Jamain and Hand (2008) performed a meta-analysis in which they compared 5807 results of research papers in which classification algorithms were applied. In total, the fit of nine classifiers was compared (linear discriminant analysis, logistic discrimination, kernel methods, naive Bayes, k-nearest-neighbors, CART decision tree, C4.5 decision tree, CN2 rule induction, and multilayer perceptrons). The best fit indices in the models with dichotomous criterion variables were obtained in the decision trees using the C4.5 algorithm.

Oskouei and Askari (2014) established predictive models for the performance of secondary school students based on C4.5 classifiers, Naive Bayes, multilayer perceptrons, RBF neural networks, bagging meta-classifiers, and AdaBoost. Incorporating the educational level of parents contextual variable, the best classifier turned out to be C4.5.

Regarding the application of data mining in the prediction of performance at university levels, in a work by Costa et al. (2017), the C4.5 algorithm achieved the best levels of accuracy in online learning and accuracies similar to the best algorithm (support-vector machine) in classroom learning. In all cases, this algorithm is more accurate than neural networks and naive Bayes. A study by Asif et al. (2017) identified the naive Bayes algorithm as the most accurate in predicting the performance of a university graduate from the initial information provided by the student during enrolment. Other works on university education focus on online learning. A work by Romero et al. (2013) established a predictive model for student performance based on their participation in discussion forums. The naive Bayes classifier and the EM clustering algorithm achieved the best levels of fit, above C4.5. Finally, a study by Akçapinar et al. (2015) analyzed the predictive power of interactions in online environments relating to student performance. Again, the naive Bayes classifier achieved the best levels of fit, slightly higher than those obtained with decision tree C4.5.

In an analysis of large-scale assessments, Kılıç et al. (2017), based on TIMSS 2011 data, established predictive models for mathematics performance based on decision trees (random forest and C4.5), Bayesian networks (naive Bayes), neural networks (multilayer perceptron), and logistic regression. The C4.5 decision tree achieved higher levels of fit than random forest, remaining at levels similar to logistic regression and Bayesian neural networks. For their part, Liu and Whitford (2011) used data from the PISA 2006 assessment to establish predictive models for performance in science. They initially categorized performance in science as a dichotomous variable (satisfactory and unsatisfactory performance), and included predictor variables related to opportunities to learn at home. The levels of accuracy in cross-validation of the models applied were around 70%.

Although the use of data mining has been significantly extended in educational research, no applied works that include a comprehensive study of the stability of the models beyond the report of overfit normally from the cross validation have been found. Stability can be defined as the degree to which an algorithm returns constant results from different samples of the same population (Turney, 1995). Since stability is inversely related to the size of the tree obtained (Jacobucci, 2018), it needs to be studied together with the goodness of fit indices in the analysis of the models.

We should point out that all of the works cited in the state of play use the gross performance of the student as a criterion variable for the predictive models, and only include in a few cases, among their predictor variables, some contextual factors. If we define school effectiveness as "the relation between the observed outcomes and the expected outcomes given the socio-economic context of education systems" (Lenkeit and Caro, 2014, p. 147), we can affirm that these studies do not take into account the fundamentals of the ESM. This argument justifies the proposal that we make in our research. Instead of using gross performance as a study criterion variable, which skews the models in favor of students and schools with higher socio-economic levels (Chapman et al., 2016), we use the identification of the level of effectiveness of the school obtained from the residuals of the schools in the multilevel models applied initially (Gamazo et al., 2018).

METHODOLOGY

Based on an analysis of secondary data from the PISA 2015 assessment (OECD, 2019), this research used a cross-sectional non-experimental research design. To avoid bias in the data related to differences in socio-economic level and the structure of educational systems between countries, we decided to select data from a single country. Spain was the sample selected for several reasons:

- Multilevel models with contextual variables applied to OECD countries based on PISA 2009 data show that Spain is one of the countries with the smallest difference between observed and estimated scores in both reading and mathematics (Lenkeit and Caro, 2014). These results suggest that the Spanish educational system, in relation to other countries assessed in PISA, reaches higher levels of equity. Therefore, data mining models obtained using as criterion variable both gross performance and school-level residuals will be more similar to each other than models obtained in other countries. Given the main objective of this study, it is interesting to be able to compare the fit of the obtained decision trees with other models based on gross performance.
- The size of the Spanish sample in PISA 2015 was much larger than that of most of the sampled countries since each of its 16 autonomous communities is taken as a stratum.

Participants

Taking into account the aforementioned, our starting point was the population of Spanish students who at the time of the 2015 PISA assessment were 15 years old, their teachers, and the schools in which they studied. In Spain, students who had undergone standardized schooling were at that time in the final year of compulsory secondary school.

From this population, the initial sample obtained was 32,330 students, 4286 teachers, and 976 schools. However, to obtain more stable estimates of the aggregated variables at the school-level (obtained from the calculation of the average score of the first level variables), and to get better estimates of multilevel model parameters, we removed from the sample all schools with less than 20 students (Hox, 2010; McNeish and Stapleton, 2016). After this filtering, the sample on which the multilevel models of the first phase of the research were applied had 871 schools, 31,105 students, and 3682 teachers.

As will be discussed later, the results of the multilevel models enabled the selection of high- and low-effectiveness schools.

The sample weights proposed in the PISA 2015 data served to weigh the data in both phases.

Variables and Instruments

The instruments included in the 2015 PISA tests, which we used in this study, were obtained from two sources:

- Performance tests in reading, mathematics, and science. The PISA tests used a sampling of items from which the ability of each student in the three areas was estimated using the item response theory (IRT). Thus, PISA assessment includes an estimate of 10 plausible values of the achievement of each student in the three main assessed areas.
- Questionnaires used with management teams (school information), teachers, and students. These questionnaires included abundant information on socio-economic context, educational processes and organizational issues, cognitive and personal aspects of students, etc.

While the reliability and validity of the achievement tests included in PISA are evidenced extensively in the technical reports (OECD, 2017), with there being general agreement in the academic community about their relevance (Hopfenbeck et al., 2018), this same level of agreement does not exist in relation to the context questionnaires obtained at student, teacher, and school level (Rutkowski and Rutkowski, 2010; González et al., 2012; Fernandez-Cano, 2016; González-Such et al., 2016). Although estimation of the dimension scores is obtained from psychometric procedures based on IRT, lower reliability is observed on these estimates, evidenced by the low correlation between the responses of students and families on similar matters (Rutkowski and Rutkowski, 2010). We should not forget that the context questionnaires in PISA include several self-perception scales and self-report measurements. Thus, there are broad criticisms on several fundamental matters:

- Social, cultural, and economic significance of the defined constructs: Cultural differences between countries make it difficult to compare the significance of these constructs and, therefore, to make cross-cultural comparisons (Hopfenbeck et al., 2018).
- Lack of stability in the definition of indicators, items, and constructs: Several items and scales change from one edition to another, others are discarded, and some others are included (Fernandez-Cano, 2016).
- Poor translations of the questionnaires into languages other than English: The versions of these questionnaires (including in this case the achievement measurement tests) in the different languages make their comparability difficult (Huang et al., 2016).
- Missing data: Contrary to what happens with the achievement measurements, which rarely include missing data in the student database, the measurements and constructs of the context questionnaires include missing data on a regular basis (Hopfenbeck et al., 2018).

As a result, although the OECD is making significant efforts in the latest editions of PISA for the improvement of these aspects (Jornet, 2016), we need to be cautious when interpreting the results obtained from these scales in their transfer to educational policies.

Regarding variables, the following were used:

- In the application of the multilevel models, the criterion variables used were gross performance in the three areas assessed at student level (Level 1) and the average performance of the school (Level 2). Unfortunately,

although a teacher database is included in PISA 2015, we could not include classroom-level variables in the models since these data do not allow to associate teachers with students in their classroom.

- The predictive variables used, which were exclusively contextual in nature, were the following:
  - Level 1: Gender; birth month; academic year; index of economic, social, and cultural status (ESCS); migratory status; repetition of academic year; number of school changes; mother tongue.
  - Level 2: Size of the school; classroom size; shortage of resources; shortage of teachers; school ownership; student/teacher ratio; average ESCS; repeater rate; immigrant student rate; proportion of girls.
- The decision trees included as a criterion variable the identification of the school as high or low effectiveness (dichotomous variable). The predictor variables included in the decision trees were all non-contextual items and scales included in the PISA 2015 databases, both in schools and in teachers and students. In total, the decision trees used included 232 variables (39 of teachers, 139 of students, and 54 of school).

Selection of the variables included in the multilevel models draws from the focus of this research, which is based on the context-input-process-output (CIPO) model. This model (Creemers and Scheerens, 1994; Creemers and Kyriakides, 2015) raises the need to differentiate between the variables on which schools and their educational communities can exert influence (process variables) and those on which it has no decision-making power (context and input variables). Thus, while context variables refer to the socio-economic and cultural environment that surrounds the school and its members, input variables are related to the personal and economic resources available and to the background of the students. From this categorization of variables, it is possible to speak about two types of school effects (Raudenbush and Willms, 1995): the first, or type-A effects, are defined as the difference between the achievement of a student and what would be obtained if he or she went to a school with certain contextual characteristics; the second, or type-B effects, can be defined as the difference between the achievement of a student in a certain school and what is expected to be obtained if that student attended a school with the same contextual conditions but with different procedural conditions (school organization, teaching methodologies, leadership process, decision-making, etc.). Thus, in the multilevel models applied in the first phase of this study, type-B effects (residuals of Level 2 of the models) are tried to be detected after controlling the type-A effects (by introducing the input and context variables as co-variables in the models).

In particular, selection of the context and input variables used in the multilevel models is based on the literature review carried out both theoretically (Creemers and Scheerens, 1994; Chapman et al., 2016) and from previous studies in similar contexts (Murillo, 2007; Joaristi et al., 2014).

Procedure and Analysis of the Data

HLM7 software was used to calculate the multilevel models, which were applied taking into account the 10 plausible values provided by PISA 2015 in each of the three areas assessed. HLM 7 computes an independent model for each of the available plausible values and returns the parameters averages. Since HLM7 does not allow the use of sample replicate weights, to minimize bias in error estimation this software employs robust estimators using the Huber-White adjustment. This adjustment compensates for the biases associated with the omission of replicate weights (Lavy, 2015; Lopez-Agudo et al., 2017).

From a significance level of 5%, we included only significant predictor variables in the multilevel models. Since the three models obtained from each achievement measurement were clearly different in terms of the predictor variables included (Gamazo et al., 2018), we consider it more appropriate to calculate three independent models.

Finally, we computed the significant models, with random intercepts and fixed slopes in school level, and calculated the residuals of the school level using empirical Bayes estimation (Raudenbush et al., 2016). In all of the cases, the values we obtained in the intraclass correlation coefficient (ICC) were greater than 10% in the null models (Eq. 1), considered as the minimum acceptable value to consider multilevel methods (Roberts et al., 2011):

$$y_{ij} = \gamma_{00} + u_{0j} + e_{ji} \tag{1}$$

where y_{ij} refers to the performance achieved by student i of school j in the corresponding area. Thus, in the null models, y_{ij} represents the sum of the overall average performance in the corresponding area (γ_{00}), the distance of average school performance j from the overall (u_{0j}), and the distance of student performance i with respect to the school j (e_{ji}).

The final models obtained in each area are specified in Eq. 2:

$$y_{ij} = \gamma_{00} + \sum_{s=1}^{S} \left(\gamma_{0s} W_{sj}\right) + \sum_{q=1}^{Q} \left(\gamma_{qj} X_{qji}\right) + \left(u_{0j} + e_{ji}\right) \tag{2}$$

where γ_{0s} and γ_{qj}, respectively, represent the main effects on the variables of school and student level, W_{sj} variables s of school, j, and X_{qij} variables q of student i of school j.

After obtaining the residuals of the schools in the three final models, we carried out the selection of high- and low-effectiveness schools. To do so, we carried out a first selection of schools (non-stratified selection), in which the schools that were placed in the first quartile in the three computed models (schools of negative residual, low effectiveness) and the schools that were placed in the last quartile in the three models (positive residual schools, high effectiveness).

Given the extensive educational competence of the autonomous communities in Spain, we made a second selection of schools (stratified selection). In this second case, we used the same criteria indicated above in each of Spain's 16 autonomous communities, implementing 16 separate selective processes for high- and low-effectiveness schools. The residuals used in this selection were the same residuals used in the original

selection. We opted for this procedure because ICC levels were below 10% in the null models of the specific samples in some autonomous communities. The decision to create a dichotomous variable from the residuals obtained in school level addresses two fundamental questions:

- The use of dichotomous criterion variables in obtaining decision trees simplifies the interpretation of the rules obtained in the models.
- The residuals used are indicators with estimation errors associated with them, so the use of their absolute values is not appropriate.

The decision trees were computed using Weka 3.8.1 free software. Given the results shown in the state of play, we decided to use the C4.5 algorithm in the construction of the models. This algorithm is an extension of ID3 (Quinlan, 1986, 1992), and its use is widespread in EDM to model student performance (Martínez-Abad and Chaparro-Caso-López, 2017; Rodrigues et al., 2018). The fit of this algorithm allowed the use of variables of all types (categorical and scale), and uses the information gain ratio for their selection. This facilitates the computation of simple models (Quinlan, 1986). With the intention of obtaining reduced trees and avoiding problems of readability and overfitting, we decided to limit the maximum number of branches to 30 (Kieskamp, 2015) to ensure that the trees obtained were easily interpretable and not overfitted. The overall fit of the decision trees was assessed based on true positive (TP), accuracy (percentage of correctly classified instances), area under the ROC curve, and kappa indicators. According to previous studies (Mitchell, 2009; Martínez-Abad and Chaparro-Caso-López, 2017), 70% of correctly classified instances were established as the cut point to determine an acceptable fit index.

As an additional control measure, we include the study of the stability of the trees obtained (Jacobucci, 2018). Taking into account that the variance of the cross-validation accuracy estimators is higher if the algorithm is unstable (Liu and Motoda, 1998), we will evaluate the internal stability of the decision trees (Aluja-Banet and Nafria, 2003) from the standard deviation and the coefficient of variation of the accuracy levels obtained in cross validations with 100-folds.

To take a known point of reference that allows the fit level of the models obtained to be assessed, logistic regression models are applied. Selection of the predictor variables included in these models is automated through the use of the LogitBoost algorithm (Landwehr et al., 2005).

It was necessary to generate a total of 120 different databases based on different determining factors:

- Informant of predictor variables (*Database*—five categories): Predictor variables from the student database; predictor variables from the teacher database; predictor variables from the school database; student and school variables with student scores aggregated at school level (aggregate data); student and school variables with the school scores included in the student level (non-aggregate data).
- Type of predictor variables included (*Items-scales*—three categories): Only items, not including scales; only scales; both items and scales.
- Stratum by Spanish region (*Stratum*—two categories): Identification and selection of high- and low-effectiveness schools in one step from the complete sample of schools in Spain; identification and selection of high- and low-effectiveness schools independently in each of the 16 regions, taking into account stratification by region.
- Significance level of the predictor variables (*Significance*—two categories): All predictor variables, both significant and non-significant, were included; only significant predictor variables were included.
- Type of sample to obtain the model (*Validation*—two categories): The models were computed from the training sample; the models were computed from a cross-validation with 10 sub-samples.

This made it possible to compute a total of 120 different decision trees based on these five determining factors (for example, one of the 120 trees calculated included as predictor variables the significance scales of the student database and as a criterion variable the identification of high- and low-effectiveness schools taking into account stratification by region, estimating the fit indices from the training sample).

After this process, we were able to compare the fit of the trees obtained based on these five determining factors. To do so, we calculated the average scores and typical overall deviations and by interest groups and used the appropriate hypothesis contrasts to compare the groups. This procedure made it possible to identify the categories or groups with best and worst fit in the predictive models.

RESULTS

Multilevel Models

The initial ICC in the three models applied was acceptable (science = 12.41%; reading = 12.04%; mathematics = 12.26%). The ICC of the final models achieved acceptable levels (science = 5.60%; reading = 5.07%; mathematics = 4.55%), since in the three models the variance explained at school level accounted for more than 50% of the total variance. The most explanatory model was the competence in mathematics predictor, in which the predictor variables accounted for 62.99% of the total variance of the second level.

The breakdown of selected schools, based on the two procedures described in the methodology, can be seen in **Table 1**. The non-stratified selection method returned a larger and more balanced sample of high- and low-effectiveness schools, while, with the stratified method, more low-effectiveness schools were selected. In no case did the selection of a school by one of the methods produce an inverse result to the other method.

Table 2 shows the breakdown of the student and teacher sample according to school selection based on stratification or without taking it into account. In this case, the use of the two

TABLE 1 | Breakdown of high- and low-effectiveness schools according to the selection procedure.

	Not stratified			
	High	**Not selected**	**Low**	**Total**
Stratified				
High	75	34	0	109
Not selected	55	518	59	632
Low	0	62	68	130
Total	130	614	127	871

TABLE 2 | Breakdown of students and teachers according to selection procedure.

		Not stratified			
		High	**Not selected**	**Low**	**Total**
Stratified					
High	Students	2569	1211	0	3780
	Teachers	294	58	0	352
Not selected	Students	2016	18,553	2230	22,799
	Teachers	241	2206	340	2787
Low	Students	0	2198	2328	4526
	Teachers	0	141	402	543
Total	Students	4585	21,962	5448	31,105
	Teachers	535	2405	742	3682

methodologies returned a sample with a similar breakdown in terms of size in high- and low-effectiveness schools.

Decision Trees: Stability and Fit

The average accuracy levels obtained according to each of the determining factors are shown in **Table 3**. Under parentheses are presented the accuracy levels obtained in the logistic regression models. Regarding *Database*, it can be observed that the highest levels of accuracy, both in the training and validated samples, are found in the samples that include school and student data, with them being higher in the aggregate data of the training sample and non-aggregate data of the validated sample. Except in the aggregate data, the models predict low effectiveness better. No significant differences were obtained regarding the use of *items and scales*. While the validated sample was more accurate if data with *significant variables* were used, in the training sample, trees in which significant and non-significant variables were used were slightly more accurate. Finally, slight differences were obtained in the training sample in terms of the use of *stratified* or *non-stratified* selection, which increased in the validated sample in favor of the stratified sample. In general, better predictive levels were maintained for low-effectiveness schools in both the training and validated samples.

Although the level of general accuracy of the logistic regression models is slightly higher, mainly in the validated data, we should bear in mind that a maximum number of predictor variables is not set in these models. It should be remembered that a maximum size of 30 branches is set in the decision trees. Thus, while the logistic regression models reach an average number of 47.12 predictor variables included (reaching a maximum of 174 variables in one of the models), the decision trees feature, on average, 15.10 rules (with a maximum of 30 rules). Therefore, this fact significantly affects the fit levels of the models.

Table 4 shows an analysis of the stability of the decision trees obtained. In general, average instabilities slightly below 40% are obtained, meaning that the level of accuracy of each repetition of the cross validation of a tree, on average, is 40% away from average accuracy. The only factor in which significant differences are found is in the databases, where it is observed that non-aggregate student and student + school data are noticeably more stable, while, in school data, high levels of instability are reached.

In general, decision trees were obtained with highly variable levels of fit depending on the determining factors proposed (**Table 5**). Certain acceptable average accuracy and TP indices can be observed in the training sample both in total prediction and the prediction of high- and low-effectiveness schools. The models were, however, significantly more accurate for low-effectiveness schools. The average fit indices in the validated data were insufficient, although, in the maximum values, acceptable scores were achieved.

TABLE 3 | Average accuracy of the decision trees according to determining factors (under parentheses accuracy of logistic regression models).

		Training			Validated		
		Total	**High**	**Low**	**Total**	**High**	**Low**
DDBB	Schools	0.803 (0.767)	0.801 (0.767)	0.812 (0.768)	0.502 (0.526)	0.483 (0.506)	0.518 (0.539)
	Teachers	0.665 (0.665)	0.647 (0.612)	0.676 (0.677)	0.581 (0.630)	0.453 (0.523)	0.610 (0.652)
	Students	0.626 (0.631)	0.617 (0.617)	0.633 (0.641)	0.591 (0.604)	0.578 (0.587)	0.603 (0.618)
	Aggr. school + student	0.934 (0.950)	0.943 (0.951)	0.927 (0.950)	0.717 (821)	0.718 (0.809)	0.717 (0.833)
	Not Aggr. school + student	0.807 (0.834)	0.786 (0.829)	0.846 (0.838)	0.786 (0.823)	0.781 (0.819)	0.797 (0.828)
Items-scales	Items + scales	0.773 (0.811)	0.759 (0.800)	0.788 (0.817)	0.638 (0.704)	0.609 (0.680)	0.650 (0.717)
	Items	0.769 (0.802)	0.758 (0.792)	0.777 (0.808)	0.632 (0.706)	0.604 (0.682)	0.642 (0.720)
	Scales	0.760 (0.695)	0.759 (0.675)	0.772 (0.700)	0.636 (0.633)	0.594 (0.585)	0.654 (0.645)
Significance	All	0.772 (0.808)	0.763 (0.795)	0.782 (0.814)	0.633 (0.704)	0.597 (0.677)	0.649 (0.716)
	Significant	0.762 (0.731)	0.754 (0.715)	0.775 (0.736)	0.637 (0.658)	0.608 (0.621)	0.648 (0.672)
Stratum	Stratified	0.776 (0.777)	0.768 (0.757)	0.784 (0.783)	0.649 (0.687)	0.597 (0.637)	0.673 (0.706)
	Not stratified	0.758 (0.762)	0.749 (0.754)	0.773 (0.767)	0.621 (0.675)	0.608 (0.661)	0.625 (0.682)

TABLE 4 | Stability of obtained decision trees (100-folds cross-validation) according to determining factors.

		Accuracy	SD	CV
DDBB	Schools	0.502	0.363	69.76%
	Teachers	0.581	0.156	26.61%
	Students	0.591	0.077	12.83%
	Aggr. school + student	0.717	0.315	44.18%
	Not Aggr. school + student	0.786	0.058	7.30%
Items-scales	Items + scales	0.638	0.257	39.99%
	Items	0.632	0.249	38.48%
	Scales	0.636	0.247	38.57%
Significance	All	0.633	0.251	38.63%
	Significant	0.637	0.250	39.38%
Stratum	Stratified	0.649	0.251	37.90%
	Not stratified	0.621	0.250	39.99%

An example of two of the very different decision trees obtained in this study is presented in **Figure 1**. It can be observed that the models are simple and intuitive. The tree on the right was the one obtained from the following determining factors: Aggr. School + Stud, Scales, All, Stratified. The accuracy obtained was 0.93 in the training sample and 0.87 in the validated sample, with very similar accuracy in the prediction of high- and low-effectiveness schools. The second tree was obtained from the following factors: Teachers, Items, Significant, Stratified. It obtained an accuracy of 0.71 in the training sample (0.63 in the validated sample). In this case, accuracy in predicting low-effectiveness schools is around 10 points higher than prediction of high effectiveness.

The rhombuses show the predictor variables included in the models and the ellipses the accuracy of each rule established by the tree. While the arrows provide information on the range of scores of the previous variable included in the rule, each ellipsis provides information on the level of accuracy of the rule (the first letter represents the prediction of schools that meet that rule as high or low effectiveness, the first number indicates the number of elements included in the rule, and the second the elements whose level of effectiveness does not match the prediction). Thus, in the right-hand tree, it can be observed that very low levels of teacher experience and job satisfaction are the fundamental variables that predict low effectiveness. Meanwhile, in schools that have teachers with higher levels of experience and job satisfaction, it is also necessary to have a staff committed to decision-making in the school and to their own teaching development.

Decision Trees: Accuracy Comparisons

Comparison of the fit of the models according to the determining factors was carried out in the validated samples through the ANOVA test, using accuracy as a dependent variable. Therefore, the determining factors *Database*, *Items-scales*, *Stratum,* and *Significance* were compared, including all of these variables as fixed factors and only the significant interactions. In **Table 6**, we can see the result obtained for the overall accuracy of the decision trees. The model obtained achieved a very high general fit (adjusted R^2 = 89.4%). Significant determining factors were observed in *Database* and *Stratum*, with a very high effect size in the first. Significant interactions resulted between *Database* and *Significance*, *Database* and *Stratum*, and *Significance* and *Stratum*. The first two interactions achieved average effect sizes.

Figure 2 shows the average accuracy in data divided by the interaction variables. Taking the school and teacher data, we can observe that the trend is different to the rest of the data, which includes student information. When student information is included, it seems preferable to use the data with all of the variables, while with data from teachers and, mainly from schools, the use of significant variables improves the accuracy of the models. If we compare *Database* with *Stratification*, we observe that it is preferable to use stratified data, unless we have data exclusively from students, in which case the improvement is significant in the non-stratified sample. Finally, a slight interaction between *Significance* and Stratum can be observed: With stratified data, the incorporation of all variables is preferable.

Table 7 shows the results obtained in the models for accuracy in high- and low-effectiveness schools, respectively. Significant determining factors varied in both models. While *Database* was significant in both cases with large effect sizes, *Stratum* was only significant in the case of low effectiveness, with a medium–low effect size. Significant interactions resulted between *Database–Significance* and *Database–Stratum*. The

TABLE 5 | Overall fit of the decision trees proposed.

		High effectiveness		Low effectiveness		Global	
		(Min, Max)	Mean (*SD*)	(Min, Max)	Mean (*SD*)	(Min, Max)	Mean (*SD*)
TP	Training	(0.18,0.95)	0.697 (0.208)	(0.59,0.97)	0.815 (0.103)	(0.59,0.95)	0.767 (0.114)
	Validated	(0.14,0.88)	0.552 (0.187)	(0.40,0.91)	0.698 (0.124)	(0.40,0.87)	0.635 (0.116)
Accu.	Training	(0.56,0.97)	0.756 (0.125)	(0.61,0.95)	0.779 (0.113)	(0.59,0.95)	0.767 (0.114)
	Validated	(0.32,0.88)	0.602 (0.146)	(0.40,0.88)	0.649 (0.116)	(0.40,0.87)	0.635 (0.116)
Kappa	Training	–	–	–	–	(0.18,0.90)	0.515 (0.246)
	Validated	–	–	–	–	(-0.20,0.76)	0.248 (0.247)
ROC	Training	–	–	–	–	(0.62,0.98)	0.802 (0.122)
	Validated	–	–	–	–	(0.38,0.92)	0.652 (0.135)

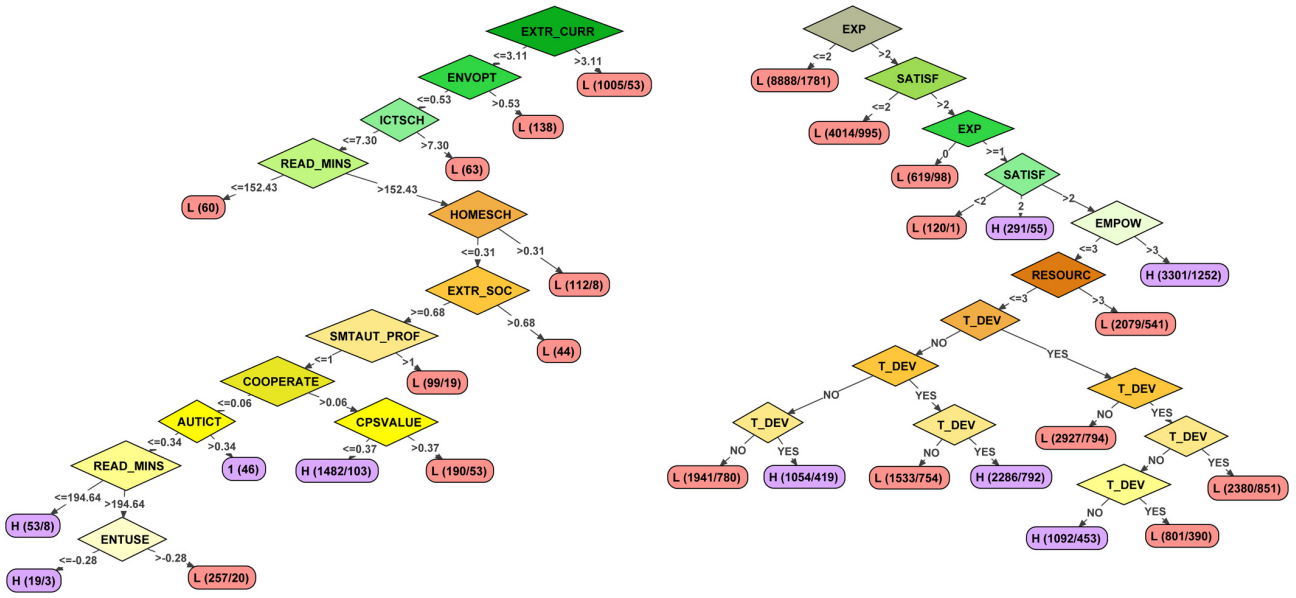


FIGURE 1 | Example of two trees obtained in the study.

models achieved a high overall fit both in the high-effectiveness model (adjusted R^2 = 87.8%) and in the low-effectiveness model (adjusted R^2 = 83.6%).

The significant interaction between Database and Significance is further analyzed in **Figure 3**. We observe an interaction with a trend similar to that obtained in the overall data. However, in the low-effectiveness data, teachers achieved a better fit, with a similar fit in the data with significant variables and all variables.

Finally, **Figure 4** presents the graphs of interaction between *Database* and *Stratum*. While the trend in low effectiveness was similar to the overall one, in high effectiveness, differences were observed: In the teacher data, the use of stratified data improved the accuracy, and in the school data, no great differences were observed between the two *Stratum* cases.

DISCUSSION AND CONCLUSION

The main aim of this work was to study the relevance of the use of decision trees for the study of educational factors associated with school effectiveness using data from large-scale assessments.

TABLE 6 | Decision tree fit comparison ANOVA table—overall accuracy.

	F	*p*	η^2
Intercept	16, 860.041	<0.001	0.998
DDBB	108.829	<0.001	0.912
Items-scales	0.154	0.857	0.007
Significance	0.140	0.710	0.003
Stratum	7.997	0.007	0.160
DDBB*significance	8.902	<0.001	0.459
DDBB*stratum	6.979	<0.001	0.399
Significance*stratum	6.186	0.017	0.128

We decided to use the C4.5 algorithm since it allows the use of variables of all kinds (Quinlan, 1986) and facilitates the obtaining of simple and easily interpretable models (Martínez-Abad and Chaparro-Caso-López, 2017). In addition, evidence from previous studies shows that it is an algorithm with highly satisfactory levels of fit both in the study of secondary school performance (Oskouei and Askari, 2014) and in the prediction of performance in large-scale assessments (Kılıç et al., 2017), higher than other classification algorithms when the criterion variable is dichotomous (Jamain and Hand, 2008). The results obtained allowed us to affirm that the fundamental aim proposed in the work was satisfactorily fulfilled thanks to the analysis performed.

On the one hand, a descriptive study of the level of fit of the decision trees was carried out based on several important determining factors. The results seem to indicate that the factor that creates the most differences in the accuracy achieved and in the stability is the *Database* used. As expected by the increase in the number of available variables, it seems that the use of school and student data combined produces better levels of fit. Observing the validated models, it seems that the use of non-aggregate data is preferable since it returns results with no overfitting in the training sample. For its part, as the previous evidence pointed out (Martínez-Abad and Chaparro-Caso-López, 2017), school data show high overfitting in the training sample. According to this evidences, stability levels were higher when students were established as the unit of analysis. Both the inclusion of significant variables and the selection of high- and low-effectiveness schools based on the stratification of the sample also result in less overfitted data, with better fit indices in the validated sample. These results coincide with the state of play, which shows significant improvements in the fit of the models when adequate pre-processing of the data is performed (Romero et al., 2013; Costa et al., 2017).

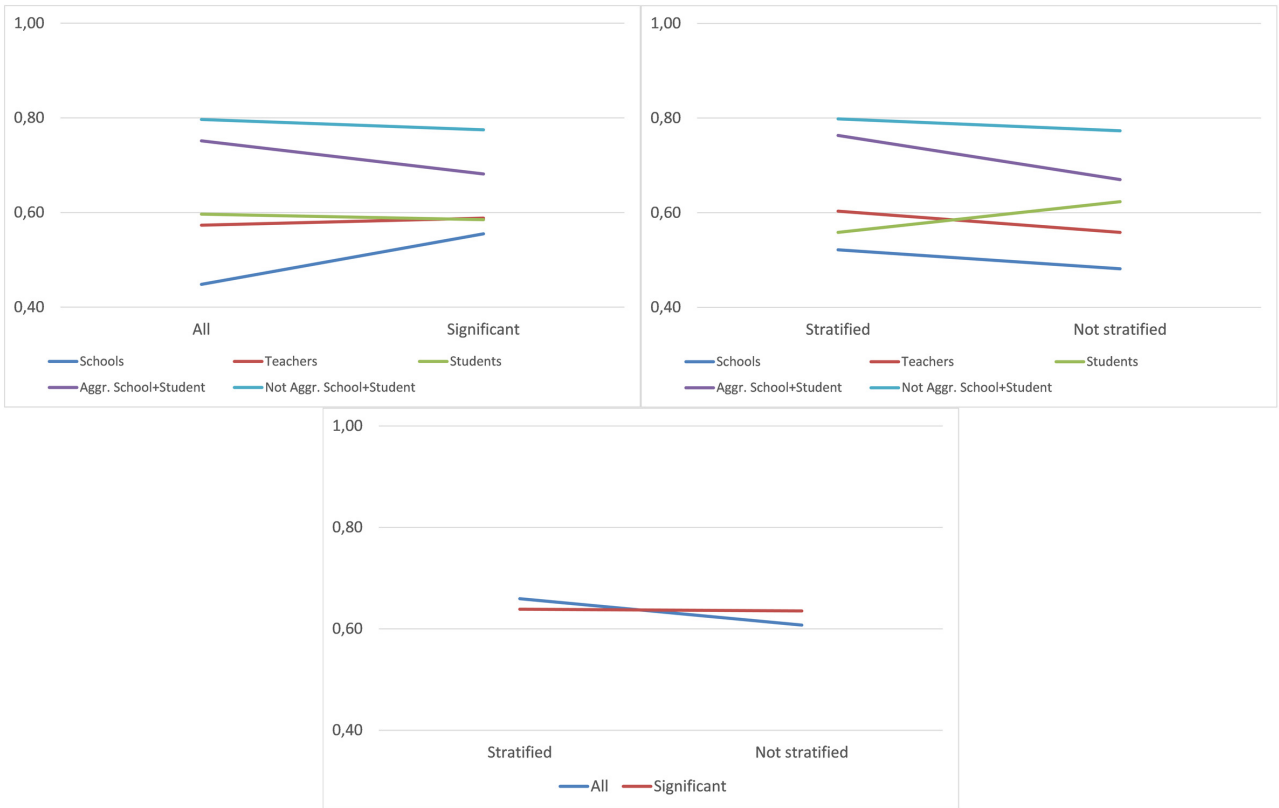


FIGURE 2 | Interaction between determining factors—overall accuracy.

Taking into account that in this study decision trees with a very small size (maximum 30 branches) were selected, the levels of average accuracy achieved, which were above 0.76 in the training sample, were satisfactory. Other previous works that did not limit the size of the decision trees achieved overall accuracy levels of between 0.7 and 0.8 (Oskouei and Askari, 2014; Akçapinar et al., 2015; Costa et al., 2017; Kılıç et al., 2017). The study by Liu and Whitford (2011) is particularly interesting since they developed a predictive model of performance from the data of students in PISA 2006 using the C4.5 algorithm. They obtained a tree with more than 100 branches, achieving an accuracy in the validated sample (10-fold cross-validation) of 0.70. Our results achieved similar levels of accuracy using trees of a significantly smaller size and from a criterion variable to which the variability of contextual factors was previously eliminated.

TABLE 7 | Decision tree fit comparison ANOVA table—high and low effectiveness.

	High effectiveness			Low effectiveness		
	F	*p.*	η^2	*F*	*p.*	η^2
Intercept	8372.433	<0.001	0.995	11, 456.530	<0.001	0.996
DDBB	95.475	<0.001	0.899	64.474	<0.001	0.857
Items-scales	0.449	0.641	0.020	0.339	0.714	0.016
Significance	0.656	0.422	0.015	0.012	0.913	<0.001
Stratum	0.741	0.394	0.017	15.683	<0.001	0.267
DDBB* significance	6.573	<0.001	0.379	5.388	0.001	0.334
DDBB* stratum	7.347	<0.001	0.406	5.021	0.002	0.318

Also notable in the results obtained was the superior fit of the models for the prediction of low-effectiveness schools. The calculated models achieved a TP rate almost 15% higher in these schools than in those of high effectiveness in the validated results (and 5% higher accuracy). Since other works seem to point to this trend based on more applied analysis (Martínez-Abad et al., 2018), it is worth delving deepen into its explanation and practical implications.

In addition, an inferential study was carried out both on the significance of the main effects of the determining factors analyzed and on the interactions between these factors. The results reflected those indicated above, highlighting that the strongest determining factor in the accuracy of the models was *Database*. *Stratum*, the other significant factor identified, reflected the importance of performing good pre-processing of the data. In contrast to the works of Romero et al. (2013) and Costa et al. (2017), in which a prior reduction in the number of variables produced an improvement in the fit of the predictive models, in our study, this factor did not achieve significant main effects. This difference may have to do with the criterion

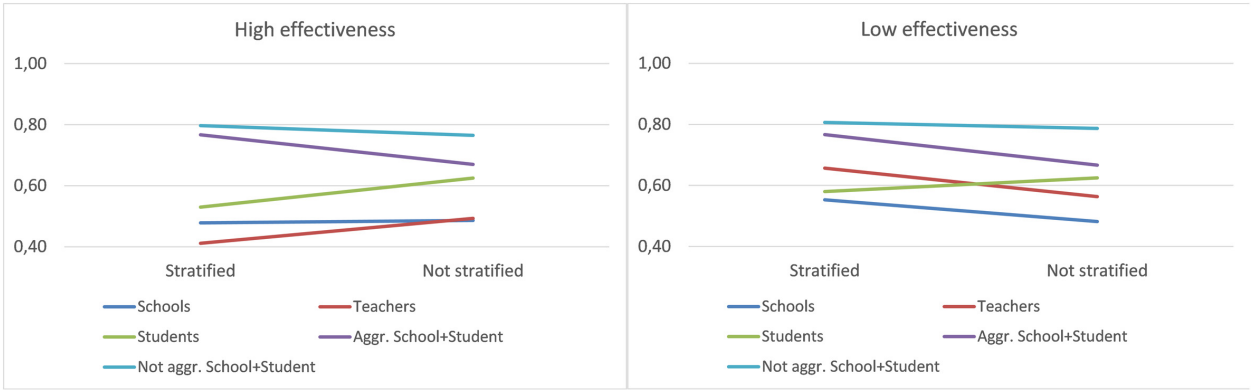


FIGURE 3 | Interaction between Database and Significance—high and low effectiveness.

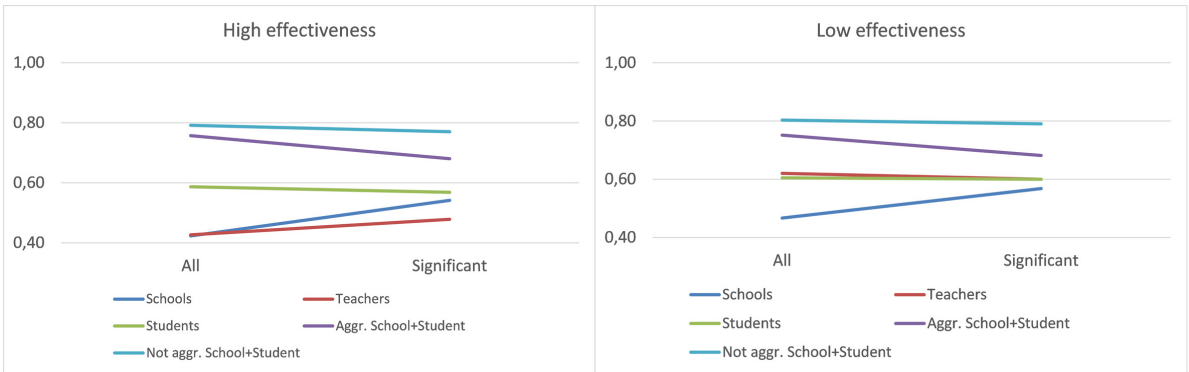


FIGURE 4 | Interaction between Database and Stratum—high and low effectiveness.

variable used in our study, which was qualitatively different to gross performance.

The interactions studied show interesting trends not assessed in previous works: on the one hand, *Databases* with better fits work worse when only significant variables are used as predictors. These effects are maintained in the independent prediction of high- and low-effectiveness schools. Regarding the interaction with *Stratum*, the student database is the only one with better levels of fit with non-stratified data. In high-effectiveness schools, however, teachers show this same trend. There is no clear pattern in these last results, and a more detailed study of them is necessary.

Several consistent strengths and contributions in our study are, therefore, confirmed, mainly in relation to the good fit shown by a good number of the models applied in general, and especially by one of them in particular. It seems that the use of decision trees from correctly pre-processed data, which includes abundant and combined school, teacher, and student information, returns predictive models of school effectiveness with good fits both in the training and validated samples. Some weaknesses inherent to this work should, however, also be highlighted. On the one hand, we find the restrictive nature of the categorization performed in the criterion variable to obtain a dichotomous variable. This decision eliminates much of the variability of this variable, limiting the possibilities of pattern identification in the data. In this regard, we decided to prioritize the obtaining of easily interpretable decision trees over trees that are very tight fitting but difficult to apply to educational reality and decision-making. We believe that this is the most appropriate procedure to facilitate the transfer of the results obtained given the level of development and current possibilities of the techniques used. In this sense, we must also point out that the computation of small trees, easily interpretable, could make it difficult to obtain trusted trees (Jacobucci, 2018). On the other hand, a more in-depth study of the implications of using aggregate data from the averages for the use of classification algorithms, or incorporating data from higher levels (school) to databases of lower levels (students) should be carried out. Comparison of the fit obtained in the database achieved by each of these two procedures shows that both methodologies have their advantages and disadvantages, with it being necessary to clear up any uncertainties about the biases that go with each one of them. Finally, it should be noted that in the first phase of the study, three independent multilevel models were calculated, one for each subject assessed by PISA, so the calculations do not take into account the covariance between the three achievement variables. Although this decision

is justified by the clear differences between the variables included in the three final models, future studies could focus on the obtaining of models that integrate all of the information from the three dependent variables (e.g., from MIMIC models).

Many future research lines of great interest for the educational scientific community, mainly in two areas, are therefore opened up for the near future. Regarding the carrying out of more basic studies similar to this one, we need to increase the volume of evidence and contributions, since there are no similar studies that use school effectiveness as a criterion variable: Works that compare the operation of various classification algorithms; systematic analysis of the implications of using combined aggregate and non-aggregate data; and studies similar to this one in which the gross residual of school effectiveness (scale variable) or politomic categorization is used. With respect to the use of studies closer to that used, there is an undeniable potential regarding the use and interpretation of specific decision trees in various databases to try to identify factors associated with effectiveness, thereby contributing to the educational characterization of high- and low-effectiveness schools.

DATA AVAILABILITY STATEMENT

The datasets for this study can be found in the OECD webpage: https://www.oecd.org/pisa/data/2015database/.

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. This study is based on the public databases of the PISA 2015 assessment (OECD). Data collection for OECD-PISA studies is under the responsibility of the governments from the participating countries.

AUTHOR CONTRIBUTIONS

This work was completely conducted and developed only by FM-A.

REFERENCES

Akçapinar, G., Altun, A., and Aşkar, P. (2015). Modeling students' academic performance based on their interactions in an online learning environment. *Elem. Educ. Online* 14, 815–824. doi: 10.17051/io.2015.03160

Aluja-Banet, T., and Nafria, E. (2003). Stability and scalability in decision trees. *Comput. Stat.* 18, 505–520. doi: 10.1007/BF03354613

Asensio, I., Carpintero, E., Expósito, E., and López, E. (2018). How much gold is in the sand? Data mining with spain's PISA 2015 results. *Revista Española de Pedagogía* 76, 225–245.

Asif, R., Merceron, A., Ali, S. A., and Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Comput. Educ.* 113, 177–194. doi: 10.1016/j.compedu.2017.05.007

Bronfenbrenner, U. (1979). *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge: Harvard University Press.

Bryk, A. S., and Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park: Sage Publications.

Casey, K., and Azcona, D. (2017). Utilizing student activity patterns to predict performance. *Int. J. Educ. Technol. High. Educ.* 14:4. doi: 10.1186/s41239-017-0044-3

Chapman, C., Muijs, D., Raynolds, D., Sammons, P., and Teddlie, C. (eds). (2016). *The Routledge International Handbook of Educational Effectiveness and Improvement Research, Policy, and Practice*. New York, NY: Routledge.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of Educational Opportunity (N.o OE-38001)*. Washington, DC: National Center for Educational Statistics.

Cordero, J. M., Pedraja, F., and Simancas, R. (2015). Success factors for educational attainment in unfavorable socioeconomic conditions. *Revista de Educacion* 2015, 163–187. doi: 10.4438/1988-592X-RE-2015-370-302

Costa, E. B., Fonseca, B., Almeida, M., Ferreira, F., and Rego, J. (2017). Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses. *Comput. Hum. Behav.* 73, 247–256. doi: 10.1016/j.chb.2017.01.047

Costa, P., and Araújo, L. (2017). Skilled students and effective schools: reading achievement in denmark, sweden, and france. *Scand. J. Educ. Res.* 62, 850–864. doi: 10.1080/00313831.2017.1307274

Creemers, B., and Kyriakides, L. (2015). Developing, testing, and using theoretical models for promoting quality in education. *Sch. Eff. Sch. Improv.* 26, 102–119. doi: 10.1080/09243453.2013.869233

Creemers, B., and Scheerens, J. (1994). Developments in the educational effectiveness research programme. *Int. J. Educ. Res.* 21, 125–140. doi: 10.1016/0883-0355(94)90028-0

Fernandez-Cano, A. (2016). A methodological critique of the PISA evaluations. *RELIEVE* 22:Art M15. doi: 10.7203/relieve.22.1.8806

Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., and Rodríguez-Conde, M. J. (2018). Assessment of factors related to school effectiveness in PISA 2015. a multilevel analysis. *Revista de Educacion* 2017, 56–84. doi: 10.4438/1988-592X-RE-2017-379-369

Goldstein, H. (1995). *Multilevel Statistical Models*. Oxford: Oxford University Press.

González, C., Caso, J., Díaz, K., and López, M. (2012). Rendimiento académico y factores asociados: aportaciones de algunas evaluaciones a gran escala. Bordón. *Revista de pedagogía* 64, 51–68.

González-Such, J., Sancho-Álvarez, C., and Sánchez-Delgado, P. (2016). Background questionnaires of PISA: a study of the assessment indicators. *RELIEVE* 22:M7. doi: 10.7203/relieve.22.1.8274

Guruler, H., Istanbullu, A., and Karahasan, M. (2010). A new student performance analysing system using knowledge discovery in higher educational databases. *Comput. Educ.* 55, 247–254. doi: 10.1016/j.compedu.2010.01.010

Hopfenbeck, T. N., Lenkeit, J., El, M., Cantrell, K., Ryan, J., and Baird, J.-A. (2018). Lessons learned from PISA: a systematic review of peer-reviewed articles on the programme for international student assessment. *Scand. J. Educ. Res.* 62, 333–353. doi: 10.1080/00313831.2016.1258726

Hox, J. (2010). *Multilevel Analysis: Techniques and Applications*, 2nd Edn. New York, NY: Routledge.

Huang, X., Wilson, M., and Wang, L. (2016). Exploring plausible causes of differential item functioning in the PISA science assessment: language, curriculum or culture. *Educ. Psychol.* 36, 378–390. doi: 10.1080/01443410.2014.946890

Hung, J. L., Hsu, Y. C., and Rice, K. (2012). Integrating data mining in program evaluation of K-12 online education. *Educ. Technol. Soc.* 15, 27–41.

Jacobucci, R. (2018). Decision tree stability and its effect on interpretation. *PsyArXiv* [Preprint].

Jamain, A., and Hand, D. J. (2008). Mining supervised classification performance studies: a meta-analytic investigation. *J. Classif.* 25, 87–112. doi: 10.1007/s00357-008-9003-y

Joaristi, L., Lizasoain, L., and Azpillaga, V. (2014). Detection and characterization of highly effective schools in the autonomous community of the basque country

using contextualized cross-sectional attainment models and hierarchical linear models. *Estudios Sobre Educacion* 27, 37–61. doi: 10.15581/004.27.37-61

Jornet, M. (2016). Methodological analysis of the PISA project as international assessment. *RELIEVE* 22:Art.M1. doi: 10.7203/relieve22.1.8293

Kasih, J., Ayub, M., and Susanto, S. (2013). Predicting students' final passing results using the classification and regression trees (cart) algorithm. *World Trans. Eng. Technol. Educ.* 11, 46–49.

Kieskamp, M. H. (2015). Please Fasten Your Seat Belts. Data Analysis Towards the Origin of Departure Delay of KLM Cityhopper at Schiphol Airport. Available at: https://essay.utwente.nl/68918/1/Kieskamp_MA_BMS.pdf (accessed August 19, 2019).

Kılıç, D., Aşkın, ÖE., and Öz, E. (2017). Identifying the classification performances of educational data mining methods: a case study for TIMSS. *Kuram ve Uygulamada Egitim Bilimleri* 17, 1605–1623. doi: 10.12738/estp.2017.5.0634

Kirby, N. F., and Dempster, E. R. (2014). Using decision tree analysis to understand foundation science student performance. Insight gained at one south african university. *Int. J. Sci. Edu.* 36, 2825–2847. doi: 10.1080/09500693.2014.936921

Landwehr, N., Hall, M., and Frank, E. (2005). Logistic model trees. *Mach. Learn.* 59, 161–205. doi: 10.1007/s10994-005-0466-3

Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *Econ. J.* 125, F397–F424. doi: 10.1111/ecoj.12233

Lenkeit, J., and Caro, D. H. (2014). Performance status and change—measuring education system effectiveness with data from PISA 2000-2009. *Educ. Res. Eval.* 20, 146–174. doi: 10.1080/13803611.2014.891462

Lezotte, L. W. (1989). School improvement based on the effective schools research. *Int. J. Educ. Res.* 13, 815–825. doi: 10.1016/0883-0355(89)90031-1

Liu, H., and Motoda, H. (1998). *Feature Selection for Knowledge Discovery and Data Mining*. Norwell, MA: Kluwer Academic Publishers.

Liu, X., and Ruiz, M. E. (2008). Using data mining to predict K-12 students' performance on large-scale assessment items related to energy. *J. Res. Sci. Teach.* 45, 554–573. doi: 10.1002/tea.20232

Liu, X., and Whitford, M. (2011). Opportunities-to-learn at home: profiles of students with and without reaching science proficiency. *J. Sci. Educ. Technol.* 20, 375–387. doi: 10.1007/s10956-010-9259-y

Lopez-Agudo, L. A., Jerrim, J., Marcenaro-Gutierrez, O. D., and Shure, N. (2017). "To weight or not to weight? The case of PISA data," in *Investigaciones de Economía de la Educación*, eds J. C. G. ómez Gallego, M. C. P. érez Cárceles, and L. Nieto Torrejón. (Madrid: Asociación de Economía de la Educación).

Martínez, R., Gaviria, J. L., and Castro, M. (2009). Concept and evolution of educational value-added models. *Revista de Educacion* 348, 15–45.

Martínez-Abad, F., and Chaparro-Caso-López, A. A. (2017). Data-mining techniques in detecting factors linked to academic achievement. *Sch. Eff. Sch. Improv.* 28, 39–55. doi: 10.1080/09243453.2016.1235591

Martínez-Abad, F., Gamazo, A., and Rodríguez-Conde, M. J. (2018). Big data in education: detection of ICT factors associated with school effectiveness with data mining techniques. *Proc. TEEM* 18, 145–150. doi: 10.1145/3284179.3284206

Martínez-Garrido, C., and Murillo, F. J. (2016). Investigación iberoamericana sobre enseñanza eficaz. *Revista mexicana de investigación educativa* 21, 471–499.

McNeish, D. M., and Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: a review and illustration. *Educ. Psychol. Rev.* 28, 295–314. doi: 10.1007/s10648-014-9287-x

Mitchell, A. J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *J. Psychiatr. Res.* 43, 411–431. doi: 10.1016/j.jpsychires.2008.04.014

Muijs, D., Kyriakides, L., Werf, G., van der Creemers, B., Timperley, H., and Earl, L. (2014). State of the art – teacher effectiveness and professional learning. *Sch. Eff. Sch. Imp.* 25, 231–256. doi: 10.1080/09243453.2014.885451

Murillo, J. (2007). "School effectiveness research in latin america," in *International Handbook of School Effectiveness and Improvement*, ed. T. Townsend, (Netherlands: Springer), 75–92. doi: 10.1007/978-1-4020-5747-2_5

OECD, (2017). *PISA 2015. Technical Report*. París: OECD Publishing.

OECD, (2019). *"PISA: Programme for International Student Assessment", OECD Education Statistics (database)*. París: OECD Publishing.,

Oskouei, R. J., and Askari, M. (2014). Predicting academic performance with applying data mining techniques (generalizing the results of two different case studies). *Comput. Eng. Appl. J.* 3, 79–88. doi: 10.18495/comengapp.v3i2.81

Pitsia, V., Biggart, A., and Karakolidis, A. (2017). The role of students' self-beliefs, motivation and attitudes in predicting mathematics achievement a multilevel analysis of the programme for international student assessment data. *Learn. Individ. Differ.* 55, 163–173. doi: 10.1016/j.lindif.2017.03.014

Quinlan, J. R. (1986). Induction of decision trees. *Mach. Learn.* 1, 81–106. doi: 10.1023/A:1022643204877

Quinlan, R. (1992). *C4.5: Programs for Machine Learning. San Mateo*. California: Morgan Kaufmann Publishers Inc.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., and Du Toit, M. (2016). *HLM7 Hierarchical Linear and Nonlinear Modeling User Manual: User Guide for Scientific Software International's (S.S.I.) Program*, Edn 1. Skokie, IL: Scientific Software International, Incorporated.

Raudenbush, S. W., and Willms, J. D. (1995). The estimation of school effects. *J. Educ. Behav. Stat.* 20, 307–335. doi: 10.3102/10769986020004307

Roberts, J. K., Monaco, J. P., Stovall, H., and Foster, V. (2011). "Explained variance in multilevel models," in, *Handbook of Advanced Multilevel Analysis* eds J. Hox, and J. K. Roberts, (New York, NY: Routledge), 219–230.

Rodrigues, M. W., Isotani, S., and Zárate, L. E. (2018). Educational data mining: a review of evaluation process in the e-learning. *Telem. Inform.* 35, 1701–1717. doi: 10.1016/j.tele.2018.04.015

Romero, C., López, M. I., Luna, J.-M., and Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. *Comput. Educ.* 68, 458–472. doi: 10.1016/j.compedu.2013.06.009

Rutkowski, L., and Rutkowski, D. (2010). Getting it 'better': the importance of improving background questionnaires in international large-scale assessment. *J. Curriculum Stud.* 42, 411–430. doi: 10.1080/00220272.2010.487546

Rutter, M., and Maughan, B. (2002). School effectiveness findings 1979–2002. *J. Sch. Psychol. J. Sch. Psychol.* 40, 451–475. doi: 10.1016/S0022-4405(02)00124-3

Şara, N.-B., Halland, R., Igel, C., and Alstrup, S. (2015). "High-school dropout prediction using machine learning: a danish large-scale study," in *Proceedings of the European Symposium on Artificial Neural Networks, (ESANN 2015)*, (Bruges: Computational Intelligence and Machine Learning), 319–324.

Tan, C. Y., and Hew, K. F. (2017). Information technology, mathematics achievement and educational equity in developed economies. *Educ. Stud.* 43, 371–390. doi: 10.1080/03055698.2016.1277137

Tan, M., and Shao, P. (2015). Prediction of student dropout in E-learning program through the use of machine learning method. *Int. J. Emerg. Technol. Learn.* 10, 11–17. doi: 10.3991/ijet.v10i1.4189

Turney, P. D. (1995). Technical note: bias and the quantification of stability. *Mach. Learn.* 20, 23–33. doi: 10.1007/BF00993473

Wrigley, T. (2013). Rethinking school effectiveness and improvement: a question of paradigms. *Discourse* 34, 31–47. doi: 10.1080/01596306.2012.698862

Xu, Y. J. (2005). An exploration of using data mining in educational research. *J. Mod. Appl. Stat. Methods* 4, 251–274. doi: 10.22237/jmasm/1114906980

Yu, X., Shi, Y., Zhang, L., Nie, G., and Huang, A. (2014). Intelligent knowledge beyond data mining: influences of habitual domains. *Commun.Assoc. Inform. Sys.* 34, 985–1000. doi: 10.17705/1CAIS.03453

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

Copyright © 2019 Martínez-Abad. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 13 December 2019
doi: 10.3389/fpsyg.2019.02843

Evaluating Impact on Motivation and Academic Performance of a Game-Based Learning Experience Using Kahoot

Andrés Fuster-Guilló[1], María Luisa Pertegal-Felices[2], Antonio Jimeno-Morenilla[1], Jorge Azorín-López[1], María Luisa Rico-Soliveres[1] and Felipe Restrepo-Calle[3]*

[1] Department of Computer Science and Technology, University of Alicante, Alicante, Spain, [2] Department of Development Psychology and Teaching, University of Alicante, Alicante, Spain, [3] Department of Systems and Industrial Engineering, Universidad Nacional de Colombia, Bogotá, Colombia

OPEN ACCESS

Edited by:
José Carlos Núñez,
Universidad de Oviedo, Spain

Reviewed by:
Anna Visvizi,
The American College of Greece, Greece
Mirna Muñoz,
Centro de Investigación en Matemáticas, Mexico
Alejandro Calderón Sánchez,
University of Cádiz, Spain

***Correspondence:**
María Luisa Pertegal-Felices
ml.pertegal@ua.es

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

Received: 16 September 2019
Accepted: 02 December 2019
Published: 13 December 2019

Citation:
Fuster-Guilló A, Pertegal-Felices ML, Jimeno-Morenilla A, Azorín-López J, Rico-Soliveres ML and Restrepo-Calle F (2019) Evaluating Impact on Motivation and Academic Performance of a Game-Based Learning Experience Using Kahoot. Front. Psychol. 10:2843. doi: 10.3389/fpsyg.2019.02843

Gamification methods adapt the mechanics of games to educational environments for the improvement of the teaching-learning process. Serious games play an important role as tools for gamification, in particular in the context of software engineering courses because of the idiosyncratic nature of the topic. However, the studies on the improvement of student performance resulting from the use of gamification and serious games in courses with different contexts are not conclusive. More empirical research is thus needed to obtain reliable results on the effectiveness, benefits and drawbacks. The overall objective of this work is to study the benefits generated by serious games in the teaching-learning process of Computer Engineering degrees, analyzing the impact on the motivation and student satisfaction, as well as on the learning outcomes and results finally achieved. To this end, an intervention is proposed in the subject of Computer Architecture based on two components covering theoretical and practical sessions. In the theoretical sessions, a serious game experience using Kahoot has been introduced, complementing the master classes and class exercises. For the practical sessions, the development of projects with groups of students has been proposed, whose results in terms of computer performance can be compared through a competition (hackathon). Evaluation of the serious game-based intervention has been approached in terms of student satisfaction and motivation, as well as improved academic performance. In order to assess student satisfaction, surveys have been used to assess the effect on student motivation and satisfaction. For the evaluation of academic performance, a comparative analysis between an experimental and a control group has been carried out, noting a slight increase in the experimental group students' marks.

Keywords: gamification, serious games, motivation, teaching-learning, computer engineering

INTRODUCTION

The present work aims to contribute to the study of the potential benefits of serious games on the teaching-learning process in a higher education context, and specifically in relation to computer engineering studies. A dual experience is proposed which considers the particularities of the theoretical and practical components. The impact on student motivation and satisfaction

of a serious game-based learning experience using Kahoot is analyzed, as well as on the learning outcomes finally achieved.

Gamification has become more relevant at the end of the first decade of the 21st century. Some authors define it as the use of typical elements of games in contexts outside the game environment (Deterding et al., 2011). According to Villagrasa et al. (2014) the main objective of gamification is to increase commitment and motivation. Gamification has been widely and successfully used in marketing to influence consumer behavior (Zichermann and Cunningham, 2011). In education, gamification techniques transfer the mechanics of games to educational environments with the aim of improving motivation, and consequently the teaching-learning process (Lee and Hammer, 2011; De-Marcos et al., 2017). The aim is to encourage interaction between teacher and student in order to increase motivation, leading to an improvement in the capacity to assimilate knowledge and acquire skills.

Experiences based on gamification have had an impact on basic and intermediate levels of education and are gradually being incorporated into university environments (Wiggins, 2016) as a response to a demand to understand the learning processes of younger generations. The work of Subhash and Cudney (2018) conducts an extensive review of gamification in the higher education environment. Moreover, the use technology in education may be a way of improving overall performance of Higher Education Institutions (HEI), due to possibly also improving professors' satisfaction and motivation (Lytras et al., 2019); and may have an impact on the overall broad question of HEI sustainability (Visvizi and Daniela, 2019).

The work by Caton and Greenhill (2014) shows how awards motivate students to produce higher quality results and attempt challenging tasks. The methods most commonly used in gamification in the context of higher education are those based on points, badges, leader boards, levels, missions or challenges. The work of Mayer et al. (2014) explores the contributions and weaknesses of Game Based Learning (GBL) and Serious Games (SGs). Serious games play an important role as components for gamification in learning process. Tools to perform interactive quizzes like kahoot have proven their effectiveness on student motivation (Orhan Göksün and Gürsoy, 2019). However, studies on the improvement of student performance resulting from the use of gamification and serious games in different subject contexts are not conclusive (Subhash and Cudney, 2018).

The work of Alhammad and Moreno (2018) provides a systematic mapping of the state of the art in gamification in software engineering studies. This study introduces a number of interesting conclusions highlighting the greater importance of gamification in the educational process of software engineering due to the idiosyncrasy of the subject area. This study concludes that more empirical research is needed to arrive at reliable results on the effectiveness, benefits and drawbacks of gamification. It is worth remembering Gartner's warning that around 80% of gamification applications will not meet business objectives, mainly because the processes have been inappropriately adapted to gamification. Moreover, the works performed by Petri et al. (2018) and Marín et al. (2019) highlight the relevance of serious games especially in this area.

From the review of the state of the art we can infer an interest to contribute with empirical studies to the clarification of the benefits of gamification and serious games, especially in the context of the higher level studies of computer engineering where the idiosyncrasies of the subject area create a case of special application interest.

For these reasons, an intervention in the subject of Computer Architecture within the Computer Engineering Degree at the University of Alicante is proposed as an objective. It is further proposed to use a double method to consider the specificities of the theoretical and practical sessions. In the case of the theoretical sessions, a serious game with interactive questionnaires is proposed by using "Kahoot" (Licorish et al., 2018), an online free game-based learning platform which allows the creation of different questions types like multiple choice quizzes, discussion questions, or surveys. The gamification method of the practical sessions is based on competitions in "hackathon" format, specifically the "CUDATHON" competition. Once the intervention has been developed, the benefits of the serious game-based experience in terms of student motivation and satisfaction, as well as in terms of learning outcomes, can be analyzed.

The rest of the paper is structured as follows: in the second section "Materials and Methods" the context is explained and the participants of the intervention, the instruments of the proposal and the procedure are also detailed. In the third section "Results," the data obtained in the intervention are reported. Finally, in the fourth section "Discussion and Conclusion" the data obtained are analyzed and the conclusions reached are summarized.

MATERIALS AND METHODS

Once the state of the art has been reviewed and the general and specific objectives of the research defined, this section explains the methodology of case of study (Wohlin et al., 2003) that uses the satisfaction survey as a tool to obtain student feedback. It describes, in detail, the teaching context and objects of study, the instrument to be used and the procedures planned for its development, including the proposed serious game tool.

Description of the Context and Participants

The subject in which the serious game experience is developed is "Computer Architecture" from the "Degree in Computer Engineering" at the University of Alicante. This course is compulsory in the second semester of the second year. It has six ECTS credits (1.2 theoretical, 1.2 practical, and 3.6 non-presential load). This means 30 classroom hours of theory and 30 classroom hours of practice, organized into 15 sessions of 2 h for both theoretical and practical elements.

The 15 2-h theoretical sessions aim to provide the student with knowledge on computer performance assessment, computer concepts and models, instruction set design, instruction level parallelism, segmentation, memory, and I/O performance. To this end, content is organized into six topics. The teaching methodology combines traditional resources such as master

classes with highly experimental mechanisms and student participation. In this context, in the academic year 2018–2019 games have been introduced into the theoretical sessions using "gaming" tools (Kahoot) to encourage student participation and motivation. The theoretical part is evaluated by means of tests during the course and a final examination of problems. In the academic year 2018–2019, gaming leader boards have been incorporated into the theoretical evaluation with the possibility of students obtaining up to one extra point for participating in the games.

The 15 2-h practical sessions aim to provide the student with the skills to be able to implement test programs to evaluate specific aspects of the computer; to use standard benchmarks to make performance reports; to design optimal software solutions taking advantage of the parallelism provided by the architectures (80 × 86, SIMD, MMX, SSE, CUDA). For the development of the practices, a project organized in three phases is proposed. Phase I deals with performance evaluation. Phase II is about taking advantage of the parallelism of the architectures through the use of 80 × 86, SIMD, MMX, SSE technologies. Phase III deals with the use of massively parallel architectures with technologies such as CUDA. Individual work is carried out to guarantee the individual acquisition of skills, along with group exercises to develop organization and integration capabilities within a work group. In this context, since the academic year 2015–2016 a "hackathon" competition has been introduced to encourage student participation and motivation. Different groups of students compete with the solution then implemented using CUDA technology which achieves a better performance in solving a given problem. The competition is organized in a day called "CUDATHON." The individual part of the student's work is evaluated by means of multiple-choice tests. The group part is evaluated by means of reports and classroom exhibitions. Since 2015–2016 the "CUDATHON" has been incorporated into the practical evaluation with the possibility of obtaining an extra point.

The number of students enrolled in Computer Architecture is usually around 140 organized into four theory groups and seven practice groups. The theory groups are taught in different languages: two in Spanish (morning and afternoon), one in Valencian and one in English (HAP: High Academic Performance). Of the seven practice groups there are five in Spanish (three in the morning and two in the afternoon), one in Valencian and one in English (HAP: High Academic Performance).

Instruments

At this point, it is appropriate to recall the specific measurement objectives of the intervention. On the one hand, it is intended to measure the effect of Kahoot-based intervention on motivation and academic performance; on the other hand, to measure the effect of CUDATHON-based experience. These last results are not detailed in the present work, given that the CUDATHON experience began the 2015–2016 academic year, requiring comparison with data from previous courses in which practices with assimilable CUDA were not developed. For this reason, in this paper we focus on detailing the effects of Kahoot-based experience with the following objectives:

(1) To measure the effect of Kahoot-based intervention on student motivation through satisfaction surveys – To cover this objective, a satisfaction survey has been prepared using the Kahoot platform itself. This survey has been carried out in the tenth theoretical session, with nine Kahoots carried out, which allows the student to formulate their own opinion on the interest and details of the use of Kahoot in the classroom. Specifically, the satisfaction survey asks seven questions with the following answers that appear in **Table 1**.

(2) To measure the effect of Kahoot-based intervention on learning outcomes by analysing the results of theoretical evaluation: To measure the effect on academic outcomes it is proposed to compare the results of the theoretical evaluation of the present academic year 2018–2019 with the results of the previous academic year 2017–2018 in

TABLE 1 | Satisfaction survey questions and answers about Kahoot use.

Q/Resp	Question	Response
Q1Resp1	When I play Kahoot	"I have fun but I don't learn"
Q1Resp2		"I have fun and I learn"
Q1Resp3		"I don't have fun but I learn"
Q1Resp4		"I don't have fun or learn"
Q2Resp1	Making Kahoots helps me reinforce what I learned in class	"It doesn't help me at all"
Q2Resp2		"It helps me a little"
Q2Resp3		"It helps me"
Q2Resp4		"It helps me a lot"
Q3Resp1	Making Kahoots motivates me to learn the subject	"Nothing"
Q3Resp2		"Little"
Q3Resp3		"Quite a lot"
Q3Resp4		"A lot"
Q4Resp1	I prefer to do the Kahoot	"As soon as class starts"
Q4Resp2		"In the middle of class"
Q4Resp3		"At the end of class"
Q5Resp1	I would like the Kahoot's length to be	"Short (<5 min)"
Q5Resp2		"Medium (between 5 and 15 min)"
Q5Resp3		"Long (>15 min)"
Q6Resp1	I prefer the teacher to use to explain the theory	"Exclusively his explanation"
Q6Resp2		"His explanation combined with Kahoot"
Q6Resp3		"His explanation combined with practical exercises"
Q6Resp4		"His explanation combined with Kahoot and practical exercises"
Q7Resp1	In general, I consider Kahoot to be	"Unnecessary"
Q7Resp2		"Unimportant"
Q7Resp3		"Necessary"
Q7Resp4		"Essential"

which the experience with Kahoot was not developed. Therefore, the proposal contemplates the comparison of the marks of this test with respect to the previous course.

The statistic Alfa de Cronbach has been chosen to carry out the reliability analysis of the survey. According to the recommendations of Gliem and Gliem (2003), Cronbach alpha coefficients can be evaluated as follows: alpha > 0.9 excellent; alpha > 0.8 good; alpha > 0.7 acceptable; alpha > 0.6 questionable; alpha > 0.5 poor; and alpha < 0.5 unacceptable. On the other hand, Nunnally (1978) mentions that in a standard exploratory analysis, an estimated alpha coefficient of 0.7 is considered adequate. The reliability result, after applying Cronbach's Alfa statistic with the instrument items was 0.75 which is acceptable considering the low number of items and the previous considerations.

Regarding the validation of the instrument, it was carried out by means of the Content Validity Index (CVI) proposed by Lawshe (1975). Under this validity index, the variables or items of an instrument are subject to expert review and are quantified as follows:

$$\mathrm{CVI} = \frac{\mathrm{ne} - N/2}{N/2}$$

ne: number of experts who rate the item favourably; N: total number of experts valuing the item.

In this research, nine professors from outside the research validated the instrument. In this sense, eight experts rated all the items favourably and one expert rated all the items positively except for Q5 item. In order for the instrument to be validated with nine experts, a CVI of 0.75 or higher is required for all of its items and in this case the CVI was 0.78 for all the items. Therefore, the instrument was validated under this index.

Procedure

As mentioned, in the academic year 2018–2019 the incorporation of games into the theoretical sessions was formalized using serious games tools such as Kahoot. The aim is to complement the masterclass method by incorporating quizzes to encourage student participation and motivation. In previous courses this type of activity was already carried out but in an unplanned way, nor generalized in relation to the subject area. In the present 2019 academic year the systematized realization of "quizzes" has been planned in all the theoretical sessions of all the groups of the subject. The Kahoot tool is an online platform that allows the development by the teacher of questionnaire "quizzes" that can be raised interactively during class sessions, to get feedback from students on the assimilation of some concept previously exposed. The Kahoot platform enables students to answer a series of questions with answers in the form of options, so that during the game everyone can observe for each question the number of student answers to each of the options, as well as the ranking achieved by the participants according to the points obtained by correct answers.

Through the use of the Kahoot tool, an interactive questionnaire linked to each of the 15 theoretical sessions has been proposed. These questionnaires are designed to be developed in the last 15 min of each theoretical session. The questions are short statements (about 20 words) and four answer options. Depending on the topic explained in class, the questions may or may not require calculations, so response times can range from 10 to 90 s. The number of questions ranges from 5 to 12 depending on the length of the response times.

A total of 15 Kahoots (with their corresponding translations into Valencian and English) have been prepared for use in each of the 15 theoretical sessions. **Table 2** shows the titles of the 15 Kahoots proposed and their links to thematic units:

The theoretical part represents 50% of the overall mark for the subject. This part is assessed by means of two multiple-choice tests (30% of the theoretical assessment) with theoretical questions test1 (topics 1 and 2) and test2 (topics 3, 4, and 5) and a final exam of problems (70% of the theoretical assessment). In the academic year 2018–2019, gaming leader boards have been incorporated into the theoretical evaluation through the possibility of obtaining one extra point for participating in the games. In each Kahoot of the theoretical sessions the students accumulate kahoot-points for the leader board, depending on the number of correct answers. The one who accumulates the most kahoot-points, once all the kahoots have been completed, gets the extra point. The rest of the students obtain a fraction of the extra point calculated according to their kahoot-points in relation to the kahoot-points for the first classified. The leader board of accumulated kahoot-points is shown in each session.

RESULTS

The following section presents the results of the research. First, the effect of Kahoot-based experience on student motivation is analyzed through their responses to the satisfaction survey. This survey was carried out during the 10th theoretical session and was answered by a total of 65 students.

TABLE 2 | List of Kahoot quizzes for each theoretical session.

Kahoot/Session	Title	Lesson
Kahoot1: Session1	Initial concepts	T1. Introduction
Kahoot2: Session2	Performance	T2. Performance
Kahoot3: Session3	Amdahl	T2. Performance
Kahoot4: Session4	CPU Performance	T2. Performance
Kahoot5: Session5	Instruction Set Architecture ISA	T3. Instruction Set Architecture
Kahoot6: Session6	Instruction Set Architecture ISA 2	T3. Instruction Set Architecture
Kahoot7: Session7	Introducing segmentation	T4. Segmentation
Kahoot8: Session8	Segmented performance	T4. Segmentation
Kahoot9: Session9	Pipeline segmentation	T4. Segmentation
Kahoot10: Session10	Satisfaction survey	
Kahoot11: Session11	Pipeline segmentation 2	T4. Segmentation
Kahoot12: Session12	Pipeline segmentation 3	T4. Segmentation
Kahoot13: Session13	Memory 1	T5. Memory
Kahoot14: Session14	Memory 2	T5. Memory
Kahoot15: Session15	Input Output	T6. Input Output

It is observed that most students (88%) "have fun and learn" when they play Kahoot (**Figure 1**), which is a clear positive indicator of motivation. In addition, does making Kahoots reinforce what you have learned in class? Most of them (54%) (**Figure 2**) answer that it helps them.

To the direct question about motivation, does Kahoots motivate me to learn the subject? The majority answer is "Quite a lot" (48%) (**Figure 3**).

Questions 4 and 5 refer to the dynamics of the questionnaires, asking for the best time to perform them Q4 and the preferred duration Q5. It is observed that the students prefer to ask the questionnaires at the end of each session (58%) (**Figure 4**). It can also be seen that the length preferred by the students for the questionnaires is between 5 and 15 min (**Figure 5**).

The students also show a clear preference for a teaching methodology that makes use of theoretical explanations combined with Kahoots and practical exercises (88%) (**Figure 6**).

Regarding the generic assessment question on the use of Kahoot, it should be noted that 61% of students consider the use of Kahoot to be essential or necessary (**Figure 7**).

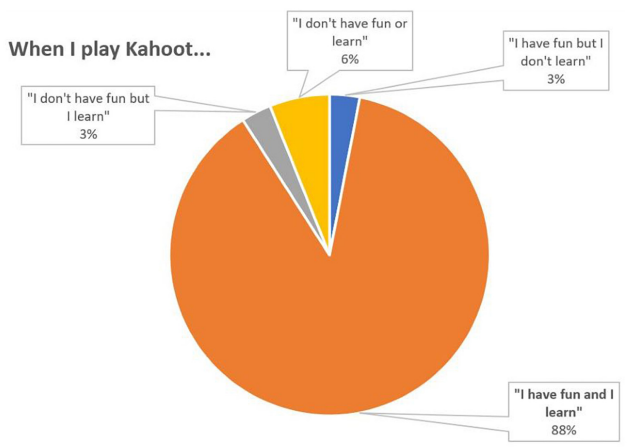


FIGURE 1 | Percentage distribution of responses to the question 1.

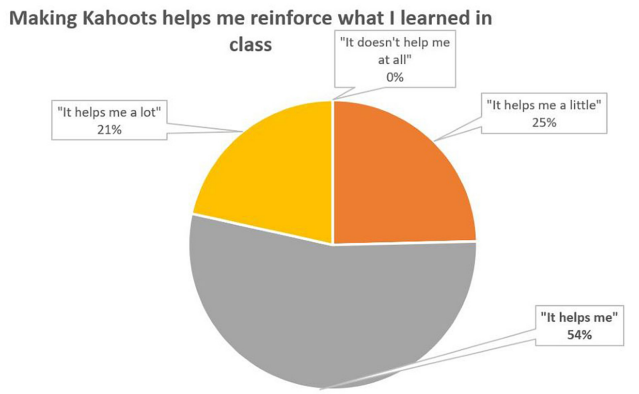


FIGURE 2 | Percentage distribution of responses to the question 2.

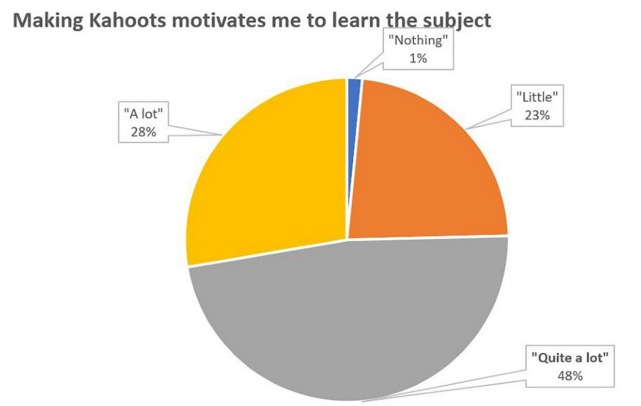


FIGURE 3 | Percentage distribution of responses to the question 3.

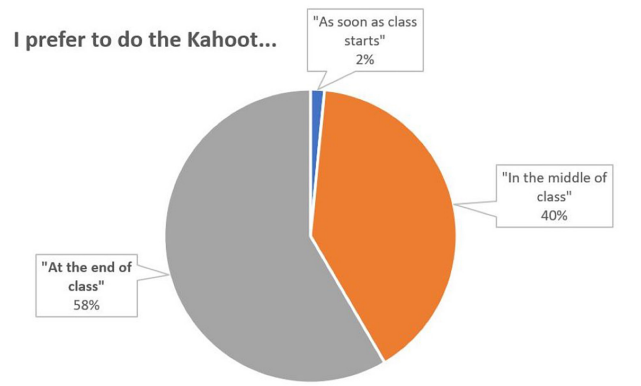


FIGURE 4 | Percentage distribution of responses to the question 4.

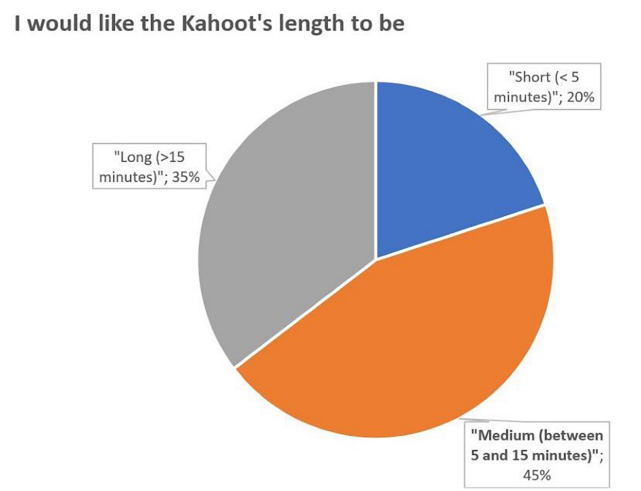


FIGURE 5 | Percentage distribution of responses to the question 5.

Finally, in order to measure the effect of Kahoot-based experience on learning outcomes, average results are provided by theory group for the current academic year 2018–2019 (called

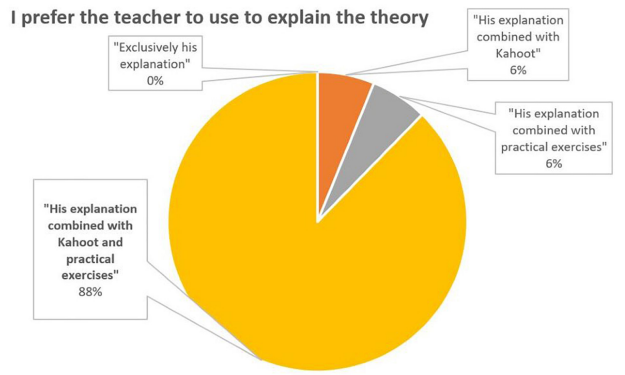


FIGURE 6 | Percentage distribution of responses to the question 6.

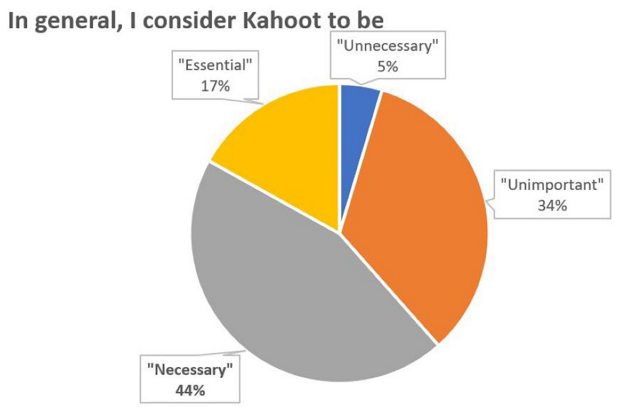


FIGURE 7 | Percentage distribution of responses to the question 7.

experimental group) in relation to the results for the previous academic year 2017–2018, in which the serious game experience was not developed (called control group). In both experimental and control groups, the assignment of tasks and the assessment methods were the same.

To assess the program's impact on student performance, two grades obtained by both groups were compared: prior to (pretest) and after the experiment (posttest). The pretest consisted of the assessment of a preliminary examination made in the very beginning of the course (first week) to test the initial knowledge of the students for facing the subject. The posttest consisted of the final examinations of the course (after week 15). Belonging to one group or the other was the independent factor or variable, and the scores obtained by the students in these examinations were the criteria or dependent variables.

The statistical procedure used the general linear model with repeated measures, with the score obtained for the examinations being taken as the dependent variable. The time of assessment (pretest and posttest) was used as the intra-subject factor; and participation in the experiment (belonging to the experimental or control group) was the inter-subject factor. All statistical analyses were conducted using SPSS (version 24.0).

The values of the inter-subject test (see **Table 3**) indicate that the means of all observations differ from 0 because the tests have been shown to be significant ($p < 0.000$) for intersection but not for group belonging ($p = 0.244$). This finding confirms that there are no initial significant differences between the two groups of students.

For the implementation of the program, **Table 4** shows the test for intra-subject effects. The values resulting from the test show that the effect of the interaction between the time of assessment (pretest and posttest) and the intervention is significant ($p = 0.000$). The observed power is 0.989, rejecting the null hypothesis of equality of means. The effect size (η^2), proportion of total variability attributable to a factor (Gardner, 2003), or the magnitude of the difference between one time and another (Ledesma et al., 2008), resulting from the interaction between the time of the assessment and the implementation of the program is 0.105.

Finally, to test whether there is any difference between the experimental group and control group, at the time of pretest and posttest, a Student's *t*-test on the difference in means was conducted, **Table 5**, which shows that there were no significant differences at the time of pretest ($p = 0.343$). This finding could mean that both groups began in comparable situations, which was already suggested by the inter-subject test. For the posttest, the test shows a significant difference between the two groups ($p = 0.000$); this difference is 1.08 out of 10 points higher in the experimental group.

Figure 8 shows the scores obtained by both groups before and after the intervention. In the posttest, the experimental group, who had used the gaming strategy, had higher scores, whereas the control group who had had no gaming interaction had worse performance.

TABLE 3 | Test of inter-subject effects.

Source	Type III error	Gl	*F*	Sig.
Intersection	12888.686	1	1849.842	0.000
Group	9.516	1	1.366	0.244
Error	1079.955	155		

TABLE 4 | Test of intra-subject effects.

Source	Type III error	Gl	*F*	Sig.	η^2 partial	Ob. Power
Apl	26.390	1	11.572	0.001	0.069	0.922
Gr × Apl	41.450	1	18.176	0.000	0.105	0.989
Error	353.469	155				

TABLE 5 | Student's *t*-test on the difference of means between the experimental and control groups.

Moment	*t*	Gl	Sig.	Diff.*	Std. dev.
PRE	0.951	155	0.343	−0.380	0.399
POST	−4.323	155	0.000	1.079	0.249

**Difference expressed as experimental group scores minus control group scores.*

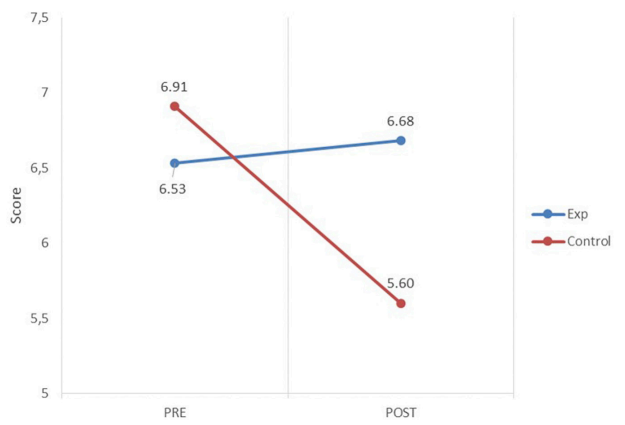


FIGURE 8 | Academic performance score (out of a maximum of 10) of the groups at pretest (PRE) and posttest (POST).

DISCUSSION AND CONCLUSION

The purpose of this case study was to evaluate a game-based intervention designed to improve student motivation and satisfaction, as well as the results and outcomes of learning and the evaluation of student satisfaction after the intervention (Wieringa, 2010).

According to all measures the introduction of Kahoot in theory classes appears to have been succeeded. Most students reported a positive perception and a positive attitude toward learning (56% believe that it reinforces what have been learned and 48% believe that it motivates a lot to learn, **Figures 2**, **3**). In general, students felt that the use of Kahoot is essential or necessary (61% according to **Figure 7**). Students overwhelmingly showed preference in sharing theoretical lesson with Kahoot and practical exercises (88% according to **Figure 6**). In addition, the improved academic performance in the experimental group into which Kahoot was introduced further supports the success of the intervention.

This work proposes a game-based experience applied in the higher educational environment of Computer Engineering based on a double method that contemplates the particularities of the theoretical and practical components of the subject. The aim of the study is to measure the improvement in student motivation and academic performance as a result of introducing the game-based experience.

In the theoretical sessions a serious game-based experience has been introduced. Specifically, interactive quizzes using the Kahoot tool have been used, incorporating questionnaires for each of the theoretical sessions. In addition, leader boards have been used to encourage participation. The result allows us to state that the serious game experience is clearly positive from the point of view of the motivation and degree of satisfaction of the student, which is clearly observed in the satisfaction survey carried out on them. This conclusion is consistent with most of the gamification and game-based experiences reviewed in the state of the art, in which a clear impact on motivation is observed (Villagrasa et al., 2014).

The gamification experience of the practical sessions is based on the proposal of competitive practices, in which groups of students compete to find the best solution to some of the problems proposed, interacting with other groups that solve the same problem. This competition called CUDATHON culminates with a hackathon format. Nevertheless, the analysis of the data provided in the results section have not been influenced by CUDATHON, since this experience has been developed in the same way in the two academic courses that are compared in the study. Although our perception in the classroom allows us to qualitatively affirm the positive acceptance of CUDATHON by students, we cannot infer or quantify its effect on motivation or academic performance.

Regarding the impact of the serious game experience on learning outcomes, the experimental results show different findings for the experimental group and the control group. First, when focusing exclusively on traditional teaching (control group, and experimental group before the pretest), a worse group performance can be observed. This decrease is significant in comparison with the experimental group, as shown in the intra-subject test in **Table 4**. This decrease may be due to the fact that the posttest is more difficult than the pretest, which is conducted at a time when student knowledge is still limited. For the experimental group, the usual tendency of obtaining lower scores than the initial test is not seen. Despite the difficulty of the posttest, there was a mild improvement, and the relatively improved results achieved by the experimental group are considered to be relevant.

Studies on the improvement of academic performance as a consequence of the use of gamification and serious games differ depending on the application context (Subhash and Cudney, 2018). The main contribution of this paper is to reinforce the idea of the improvement of learning outcomes as a consequence of introducing a serious game experience in the context of computer engineering courses at the higher education level.

Because Kahoot is widely used, the experience can be easily extrapolated to other educational fields. The benefits in other areas such as social sciences or humanities can be comparable to those obtained by this research. In this way it is proposed as future work to replicate this experience in other degrees and make a comparison with the results obtained in this research.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Alicante. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AF-G, MP-F, and AJ-M: conceptualization, methodology, and reviewing and approving the final manuscript. JA-L, FR-C, and MR-S: experimentation and validation. AF-G and MR-S: writing and reviewing.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. David Bradley for his valuable comments and reviews which have made an important contribution during the preparation of the manuscript.

REFERENCES

Alhammad, M. M., and Moreno, A. M. (2018). Gamification in software engineering education: a systematic mapping. *J. Syst. Softw.* 141, 131–150. doi: 10.1016/j.jss.2018.03.065

Caton, H., and Greenhill, D. (2014). Rewards and penalties: a gamification approach for increasing attendance and engagement in an undergraduate computing module. *Int. J. Game Based Learn.* 4, 1–12. doi: 10.4018/ijgbl.2014070101

De-Marcos, L., Garcia-Cabot, A., and Garcia-Lopez, E. (2017). Towards the social gamification of e-learning: a practical experiment. *Int. J. Eng. Educ.* 33, 66–73.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., and Dixon, D. (2011). "Gamification. using game-design elements in non-gaming contexts," in *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems*. New York, NY: ACM Press, 2425–2428. doi: 10.1145/1979742.1979575

Gardner, H. (2003). Multiple Intelligences after Twenty Years, Vol. 21. Chicago, IL: American Educational Research Association, 1–15.

Gliem, J. A., and Gliem, R. R. (2003). "Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales," in *Proceedings of the Midwest Research to Practice Conference in Adult, Continuing, and Community Education*, Columbus, OH, 82–88.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Pers. Psychol.* 28, 563–575. doi: 10.1111/j.1744-6570.1975.tb01393.x

Ledesma, R., Macbeth, G., and De Kohan, N. C. (2008). Effect size: a conceptual review and applications with the ViSta statistical system. *Rev. Latinoam. Psicol.* 40, 425–439.

Lee, J. J., and Hammer, J. (2011). Gamification in education: what, how, why bother? *Acad. Exch. Q.* 15, 1–5.

Licorish, S. A., Owen, H. E., Daniel, B., and George, J. L. (2018). Students' perception of Kahoot!'s influence on teaching and learning. *Res. Pract. Technol. Enhanc. Learn.* 13:9. doi: 10.1186/s41039-018-0078-8

Lytras, M. D., Visvizi, A., Damiani, D., and Mthkour, H. (2019). The cognitive computing turn in education: prospects and application. *Comput. Human Behav.* 92, 446–449. doi: 10.1016/j.chb.2018.11.011

Marín, B., Frez, J., Cruz-Lemus, J., and Genero, M. (2019). An empirical investigation on the benefits of gamification in programming courses. *ACM Trans. Comput. Educ.* 19. doi: 10.1145/3231709

Mayer, I., Bekebrede, G., Harteveld, C., Warmelink, H., Zhou, Q., Van Ruijven, T., et al. (2014). The research and evaluation of serious games: toward a comprehensive methodology. *Br. J. Educ. Technol.* 45, 502–527. doi: 10.1111/bjet.12067

Nunnally, J. (1978). *Psychometric Methods*. New York, NY: McGraw-Hill.

Orhan Göksün, D., and Gürsoy, G. (2019). Comparing success and engagement in gamified learning experiences via Kahoot and Quizizz. *Comput. Educ.* 135, 15–29. doi: 10.1016/j.compedu.2019.02.015

Petri, G., Calderón, A., Von Wangenheim, C. G., Borgatto, A. F., and Ruiz, M. (2018). Games for teaching software project management: an analysis of the benefits of digital and non-digital games. *J. Univers. Comput. Sci.* 24, 1424–1451.

Subhash, S., and Cudney, E. A. (2018). Gamified learning in higher education: a systematic review of the literature. *Comput. Human Behav.* 87, 192–206. doi: 10.1016/j.chb.2018.05.028

Villagrasa, S., Fonseca, D., Redondo, E., and Duran, J. (2014). Teaching case of Gamification and visual technologies for education. *J. Cases Inform. Technol.* 16, 38–57. doi: 10.4018/jcit.2014100104

Visvizi, A., and Daniela, L. (2019). Technology-enhanced learning and the pursuit of sustainability. *Sustainability* 11:4022. doi: 10.3390/su11154022

Wieringa, R. (2010). "Design science methodology: principles and practice," in *Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering*, Cape Town, 493–494. doi: 10.1145/1810295.1810446

Wiggins, B. E. (2016). An overview and study on the use of games, simulations, and gamification in higher education. *Int. J. Game Based Learn.* 6, 18–29. doi: 10.4018/ijgbl.2016010102

Wohlin, C., Höst, M., and Henningsson, K. (2003). "Empirical research methods in software engineering," in *Empirical Methods and Studies in Software Engineering. Lecture Notes in Computer Science*, Vol. 2765, eds R. Conradi, and A. I. Wang, (Heidelberg: Springer), 7–23. doi: 10.1007/978-3-540-45143-3_2

Zichermann, G., and Cunningham, C. (2011). *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. Newton, MA: O'Reilly Media, Inc.

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.

Copyright © 2019 Fuster-Guilló, Pertegal-Felices, Jimeno-Morenilla, Azorín-López, Rico-Soliveres and Restrepo-Calle. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 07 January 2020
doi: 10.3389/fpsyg.2019.02513

Relationship Between SES and Academic Achievement of Junior High School Students in China: The Mediating Effect of Self-Concept

Shifeng Li[1,2], Qiongying Xu[1,2] and Ruixue Xia[1,2]**

[1] School of Psychology, Northwest Normal University, Lanzhou, China, [2] Key Laboratory of Behavioral and Mental Health, Lanzhou, China

Over the past decades, the relationship between family socioeconomic status (SES) and academic achievement in school-age children has been well documented. However, the underlying mechanism of how family SES works on academic achievement remains unclear. In this study, we examine the possible role of self-concept in the relationship between SES and school academic achievement among 345 junior high school students in China. The results showed that both family SES and self-concept were significantly associated with the children's Chinese and mathematics performance, and family SES was also significantly correlated with self-concept. The mediation analysis showed that self-concept partially mediated the relationship between SES and school academic achievement. These findings suggest that interventions targeting self-concept may be an effective way in which to improve children's school academic achievement.

Keywords: socioeconomic status, self-concept, school academic achievement, adolescents, China

OPEN ACCESS

Edited by:
Raquel Gilar,
University of Alicante, Spain

Reviewed by:
Fernando Doménech-Betoret,
Jaume I University, Spain
Evangelia Karagiannopoulou,
University of Ioannina, Greece

****Correspondence:***
Qiongying Xu
xuqiongying@nwnu.edu.cn
Ruixue Xia
xrx9391@163.com

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

Received: *02 July 2019*
Accepted: *23 October 2019*
Published: *07 January 2020*

Citation:
Li S, Xu Q and Xia R (2020) Relationship Between SES and Academic Achievement of Junior High School Students in China: The Mediating Effect of Self-Concept. Front. Psychol. 10:2513. doi: 10.3389/fpsyg.2019.02513

INTRODUCTION

Socioeconomic status (SES), an index of one's overall social status or prestige in society, is one of the most widely studied constructs in the social sciences. It is usually measured alongside education, occupational status, and income (Conger and Donnellan, 2007). Over the past decades, the relationship between SES and child development has been well documented (Bradley and Corwyn, 2002; Hackman et al., 2010; Aizer and Currie, 2014). Compared to children and adolescents growing up in families with high SES, those growing up in families with low SES demonstrated an increased health risk (Chen et al., 2002), higher rates of anxiety, depression, and conduct disorders (Wadsworth and Achenbach, 2005). Numerous studies also associated SES with the IQ level and academic achievement of children and adolescents (National Institute of Child Health and Human Development [Nicdh], 2005; White et al., 1993).

Over recent decades, the relationship between family SES and academic achievement in school-age children has been well documented across different sociocultural contexts. As early as 1966, the well-known Coleman Report revealed that family SES explained most variances in academic achievement (Coleman et al., 1966). Sirin (2005) conducted a comprehensive meta-analysis of 58 studies. For a sample from the United States, Sirin confirmed a medium to strong relation between SES and achievement with an average effect size of 0.27 (95% CI: 0.28–0.29). Similarly, Liu et al. (2019) conducted a meta-analysis based on 215,649 students from 78 independent samples. The sample from Mainland China demonstrated a moderate relation between SES and academic

achievement (r = 0.243). However, the mechanism underlying the relationship between SES and child development remains unclear.

Several explanations have been proposed to interpret how family SES impacts child development. The most influential are the social causation model, social selection model, and sociocultural self model. The social causation model argues that social and economic conditions may influence children's functioning and development (Conger et al., 2002). Some empirical studies supported this view, demonstrating that family economic hardship negatively affected parent emotion, relationship, and parenting behavior, which influence child development (Conger and Conger, 2002). Likewise, the investment of resources (including financial, social, and human capital) by families promotes the development of their children (Bradley et al., 2001). The social selection model takes a different approach to the relationship between SES and child development, arguing that individuals' characteristics or attributes may influence their social and economic status (Mayer, 1997; Rowe and Rodgers, 1997). Some empirical research supports the social selection argument, showing that the positive characteristics of parents will reduce exposure to economic pressure in the family; decrease the likelihood of parents' emotional, interparental, and parenting problems; and improve child well-being (Linver et al., 2002).

The sociocultural self model integrated and extended the key tenets of the social causation and social selection models (Stephens et al., 2012). It proposed that (1) social economic conditions and individual characteristics or attributes are interdependent forces that influence each other, and (2) both social economic conditions and individual characteristics or attributes indirectly influence individuals' behavior through the self. In this model, self is defined as "a product of the ongoing mutual constitution of individuals and structures and serve to guide people's behavior by systematically shaping how people construe situations" (Stephens et al., 2012, p. 733). Compatible with this view, recent studies indicated that academic self-efficacy mediated the relationship between SES and anticipated and actualized school performance (Wiederkehr et al., 2015). Interventions targeting the self and identity have been effective in reducing the racial/ethnic achievement gap for college (Cohen et al., 2006, 2009) and high school students in the United States (Sherman et al., 2013; Goyer et al., 2017).

Although there is evidence that improving family economic conditions reduces children's risk of psychiatric disorders (Costello et al., 2003), and that interventions targeting individual attributes (e.g., attention) can significantly facilitate child development in low SES families (Neville et al., 2013), SES and individual characteristics remain relatively stable over a certain period. The sociocultural self model proposed a new and promising perspective in facilitating child development in families with low SES by changing the selves of students that emerged in a certain situation. However, previous studies on the mechanism of self mediating the relationship between SES and child development focused on western samples.

To our knowledge, no study has examined the role of self-concept in the relationship between SES and child development using a Chinese sample. It is well known that cultural experiences influence and determine one's self. In China, under the influence of the Confucian culture, parents tend to pay much attention to their children's learning activities and academic achievement. In most families, parents do their best to provide good learning conditions regardless of SES (Wong et al., 2012; Wang et al., 2014). Relevant research has shown that this type of parent support may influence the self-concept of children, which influences their school achievement (Xiao and Liu, 2017). Accordingly, some studies indicated the relation between good academic achievement and the praise and respect children received (Zhou et al., 2010). Whether this difference in cultural value influences the relationships between SES, self-concept, and school achievement remains unclear.

Therefore, in this study, we investigate the relationships between family SES, self-concept, and school academic achievement of Chinese junior school students in Mainland China. We hypothesized that family SES measured through parents' education, occupation, and income would be significantly associated with children's self-concept, which will influence their school achievement.

MATERIALS AND METHODS

Participants

In total, 345 first-year students (age range = 9–17 years, M_{age} = 13.40 years, SD = 0.73; 52.4% female) at a junior middle school in Lanzhou, China were recruited as participants in this study. This school is a medium to large-sized community-based public elementary school with approximately 40–50 students in each class and approximately 3,000 students in total in grades 1–3. All students are native Mandarin speakers and native to Mainland China. We also followed ethics guidelines and obtained permission from the school principals, teachers, parents, and children. Consent was first obtained from the school principals and teachers. Then, parents indicated their consent by signing a form distributed at a parent meeting or brought home by their children.

Measures and Procedure

All children completed a demographic question (age, gender, parents' level of education, parents' occupational status or what jobs the parents held, and annual household income) and a self-concept scale in the middle of the second semester. Children completed the demographic questions by taking the questionnaire home and consulting with their parents. The children were asked to complete the self-concept scale by themselves. At the end of the semester, we obtained their final exam scores for two subjects (Chinese and mathematics) as indicators of their school academic achievement.

SES

Although there is no consensus on how to measure SES, it is agreed that a stable measure thereof should incorporate education, occupation, and income (Bradley and Corwyn, 2002).

Therefore, in this study, we used parents' level of education, occupation, and annual household income to estimate the children's family SES. Parents' level of education was measured on a seven-point Likert scale: 1 = primary grade 3 or below, 2 = primary grade 4 to 6, 3 = middle school, 4 = high school, 5 = 3 year college, 6 = 4-year university, 7 = postgraduate. Parents' occupation was measured using the Occupational Prestige Scale (Li, 2005), in which 81 occupations are rated and their scores standardized as 0–100. A higher score represents the higher prestige of that occupation. Annual household income was measured on a ten-point Likert scale: 1 = less than 10,000; 2 = between 10,000 and 30,000; 3 = between 30,000 and 50,000; 4 = between 50,000 and 100,000; 5 = between 100,000 and 150,000; 6 = between 150,000 and 200,000; 7 = between 200,000 and 300,000; 8 = between 300,000 and 500,000; 9 = between 500,000 and 1,000,000; 10 = more than 1,000,000 Chinese Yuan per year.

Self-Concept

Children's self-concept was assessed using the Children and Adolescents Self-Recognition Scale (CASRS) developed by Dong and Lin (2011). This scale includes 18 items, which assess the children's perceived self through their past experience and understanding of this past experience (Shavelson et al., 1976). Sample items are: "Most of my courses are very good," "I did well in most of my courses." Children were asked to rate all items on a four-point Likert scale (1 = strongly disagree, 4 = strongly agree). The sum of the scores of each item was the final score of this scale, with higher scores indicating a more positive perceived self-concept. Previous studies confirmed the high validity and reliability of the CASRS (Dong and Lin, 2011). In this study, the internal consistency reliability coefficient (Cronbach's α) of the scale was 0.82.

School Academic Achievement

School academic achievement in this study was defined as the children's performance in school subject areas such as language literacy and mathematics. Because there are no standardized language and mathematics tests in China, following previous studies (Donnelly et al., 2016; Xiang et al., 2017), we collected the children's final exam scores for two subjects (Chinese and mathematics) as indicators of their school academic achievement. In China, school achievement is usually assessed through a Teacher-Edited Test, which examines students' learning and understanding in school subject areas such as mathematics and language literacy in the middle and at the end of each semester. In this study, the test raw score for each subject area ranged from 0–150, with higher scores indicating higher performance in that subject area.

Data Analysis

First, descriptive statistics (mean scores, standard deviations, and range) and Pearson's correlations were calculated using SPSS 22.0 for each variable. Then, the mediation model was tested in Mplus 7.0 (Muthen and Muthen, 1998/2012). In mediation model, we use latent construct to estimate the SES with five observed variables (father' level of education, mother' level of education, father's occupational prestige, mother's occupational prestige, and annual household income).

RESULTS

Descriptive Statistics, Internal Reliability, and Inter-Correlations

Table 1 shows the means, standard deviations, and ranges of all measures in this study. **Table 2** shows the inter-correlations for all measures controlling for sex and age. As **Table 2** shows, most SES measures (parents' level of education, occupation, and annual household income) were significantly correlated with the self-concept as well as performance in Chinese and mathematics, respectively. In addition, self-concept was also significantly correlated with performance in Chinese and mathematics.

Mediation Model

To examine the possibility that the relationship between SES and school academic achievement was mediated by self-concept, we conducted two mediation analyses. The first analysis tested the hypothesis that self-concept mediated the relationship between SES and performance in Chinese. The second tested the hypothesis that self-concept mediated the relationship between SES and performance in mathematics.

Performance in Chinese

Figure 1 shows the results for performance in Chinese. As shown in **Figure 1A**, first, a direct model was used to test the relationship between SES and school Chinese achievement. As expected, the direct effect ($\beta = 0.23$, $p < 0.01$) from SES to school Chinese achievement was significant. Then, the mediation model was used to examine the potential mediating effects of self-concept on the relationship between SES and school Chinese achievement. As shown in **Figure 1B**, when self-concept entered as a mediator, the direct effect ($\beta = 0.20$, $p = 0.01$) from SES to school Chinese achievement was still statistically significant, and the indirect effect of self-concept was also significant ($\beta = 0.027$, $p < 0.05$). These results confirmed that self-concept partially mediated the relationship between SES and Chinese achievement.

TABLE 1 | Means, standard deviations, range, and reliability for all measures (n = 345).

Variable	*M*	*SD*	Range	Cronbach's
Mother's education level	3.40	1.22	1–7	–
Father's education level	3.30	1.47	1–7	–
Mother's occupational prestige	38.83	13.43	9.73–87.92	–
Father's occupational prestige	46.92	12.95	26.35–87.92	–
Annual household income level	3.80	1.81	1–10	–
Self-concept	39.83	5.05	20–53	0.82
Chinese achievement	79.26	10.84	31–121	–
Mathematics achievement	76.89	19.55	12–141	–

TABLE 2 | Correlations among SES measures, self-concept, Chinese, and mathematics performance controlling for sex and age.

Variable	1	2	3	4	5	6	7	8
1. Mother's education level	1.00							
2. Father's education level	0.47***	1.00						
3. Mother's occupational prestige	0.33***	0.14*	1.00					
4. Father's occupational prestige	0.29***	0.25***	0.51***	1.00				
5. Annual household income level	0.16*	0.21***	0.21***	0.21***	1.00			
6. Self-concept	0.12*	0.16**	0.12*	0.04	0.20***	1.00		
7. Chinese performance	0.22***	0.19***	0.08	0.01	0.14*	0.18**	1.00	
8. Mathematics performance	0.16**	0.08	0.16**	0.05	0.16**	0.23***	0.63***	1.00

*n = 345; *P < 0.05; **P < 0.01; and ***P < 0.001.*

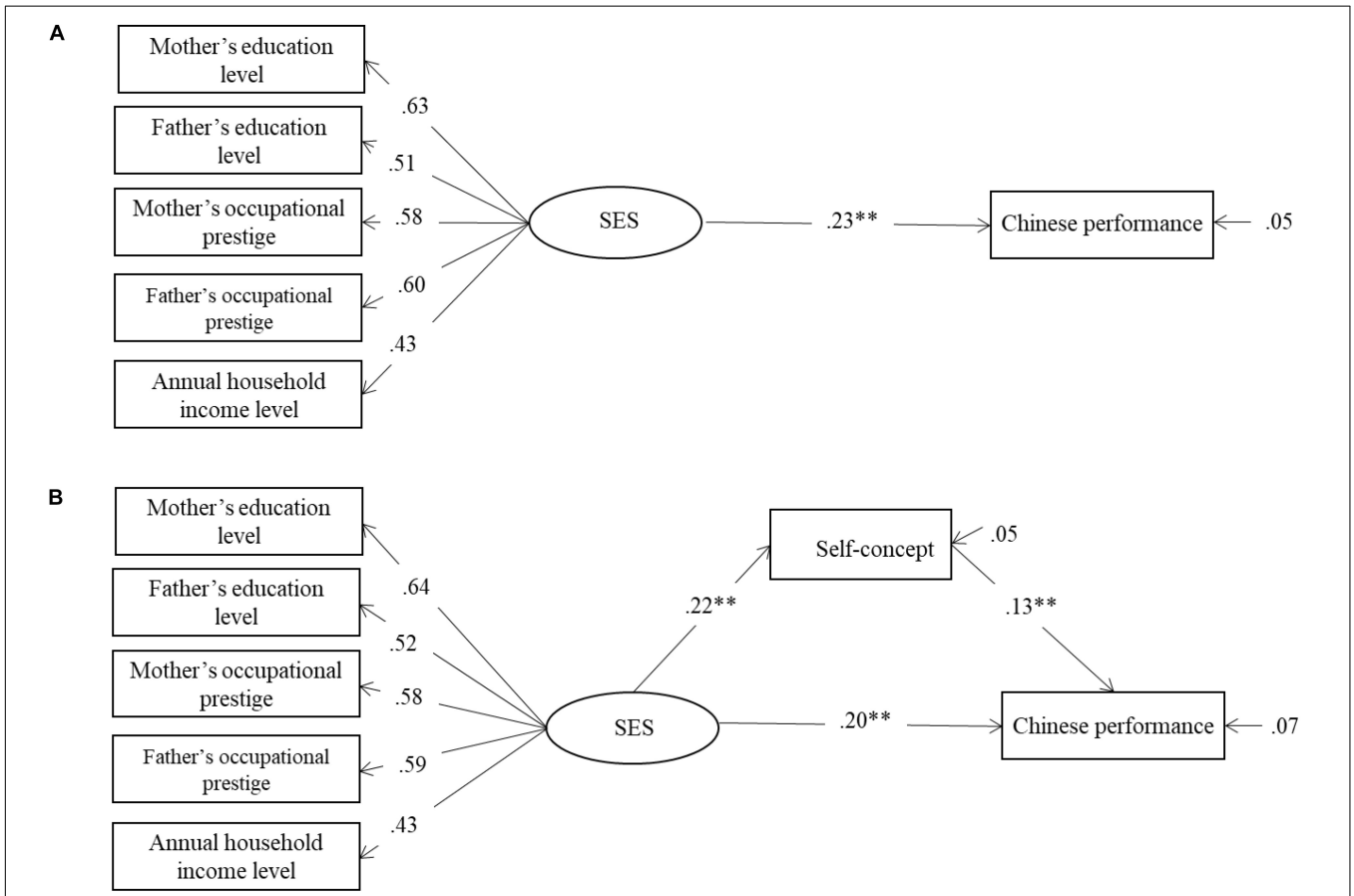


FIGURE 1 | (A) The relationship between SES and Chinese performance. **(B)** Self-concept mediate the relationship between SES and Chinese performance. SES = socioeconomic status. *$p < 0.01$; ***$p < 0.001$; $n = 345$.

Performance in Mathematics

The same analyses were repeated for performance in mathematics. As shown in **Figure 2A**, the direct model showed that the direct effect ($\beta = 0.20$, $p < 0.01$) from SES to school mathematics achievement was significant. The mediation model showed that when self-concept entered as a mediator (**Figure 2B**), the direct effect ($\beta = 0.17$, $p < 0.01$) from SES to school mathematics achievement was still statistically significant, and likewise, the indirect effect of self-concept was also significant ($\beta = 0.042$, $p < 0.05$).

Moderation Model

The possible moderating effects of self-concept in the relationship between SES and academic achievement (in both Chinese and mathematics) were examined using hierarchical regression analyses. The interactions of SES with self-concept were represented by multiplying the standard SES score with the standard score of the self-concept measures. However, these interactions demonstrated no significant effects after sex, age, SES, and self-concept were put into previous blocks in the corresponding models (all $p > 0.05$).

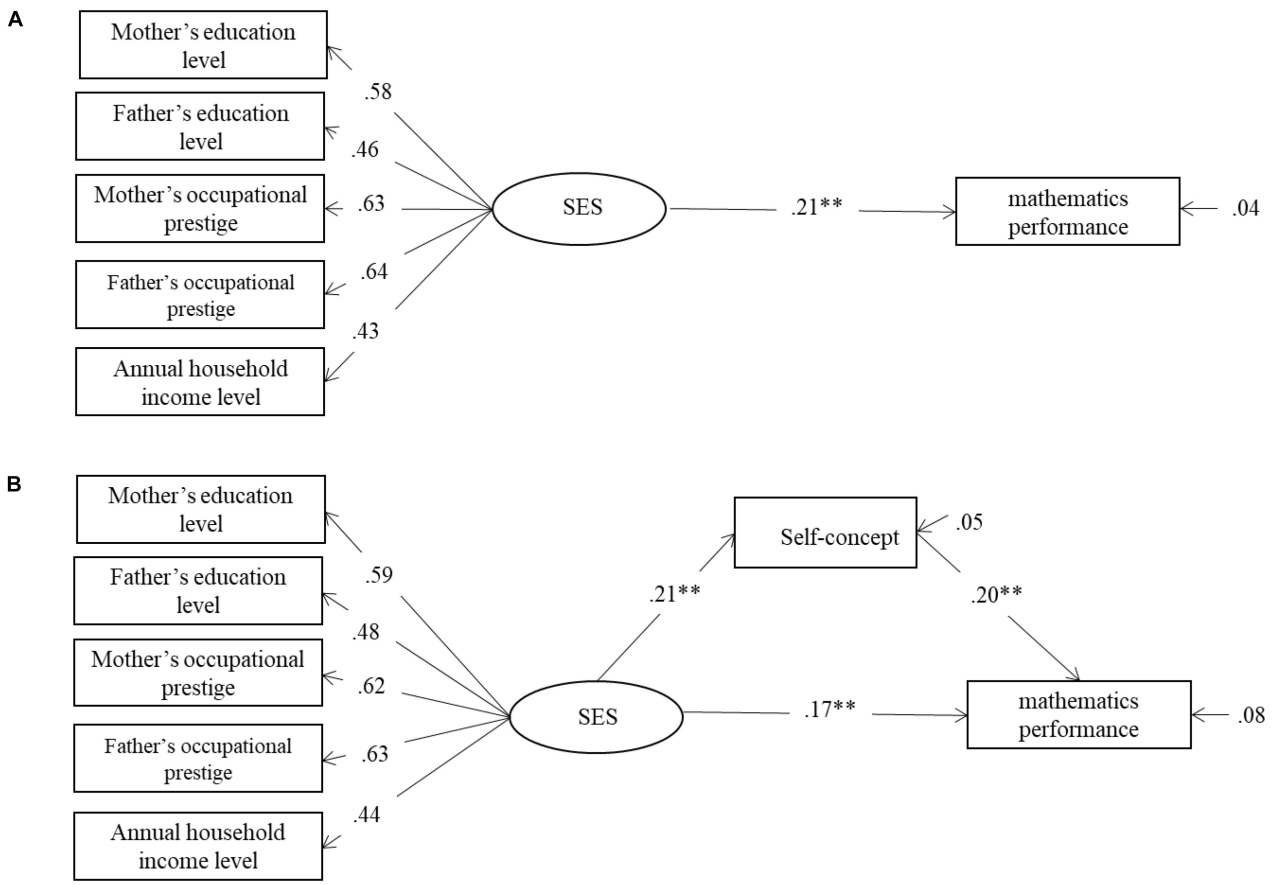


FIGURE 2 | (A) The relationship between SES and mathematics performance. **(B)** Self-concept mediate the relationship between SES and mathematics performance. SES = socioeconomic status. $^{**}p < 0.01$; $^{***}p < 0.001$; $n = 345$.

DISCUSSION

Social inequalities have profound effects on the development of children and adolescents. Reducing social class disparities in their development is on the global research agenda (Adler et al., 1993; Cohen et al., 2006; Stephens et al., 2012). However, the mechanism underlying the relationship between SES and child development remains unclear. This study was designed to explore the possible mediating role of self-concept in the relationship between SES and academic achievement among junior high school students in China. We found a moderate relation between SES and academic achievement (Chinese: $r = 0.18$; mathematics: $r = 0.23$). This finding is aligned with that of Liu et al. (2019), which reported that the overall relationship between Chinese students' SES and academic achievement was moderate ($r = 0.24$). However, our finding differs from that of the PISA report ($r^2 = 0.18$ is about $r = 0.42$) in the B-S-J-G area (Beijing-Shanghai-Jiangsu-Guangdong) of China (OECD, 2016). One reason for the differences in the results of studies may be related to the level of socio-economic and cultural development of the selected sample. B-S-J-G are among the top four provinces in China in terms of economics, residents' incomes, and school education, whereas Gansu (in the current study) is among the bottom provinces in terms of these aspects (NBSC, 2016). Compatible with this explanation, relevant studies showed that the relationship between SES and academic achievement in developing countries, especially in low-income countries, was weaker than that in developed countries (Heyneman and Loxley, 1983; OECD, 2016). Further studies are needed to examine how the levels of socio-economic and cultural development modulate the relationship between SES and academic achievement in different provinces or regions of China.

One important implication of this study is that SES predicted academic achievement partially through the mediating effect of self-concept. This is generally consistent with previous studies using a western sample, which revealed that self-efficacy mediated the relationship between SES and anticipated and actualized school performance (Wiederkehr et al., 2015). The overall findings support the theoretical framework that children's academic achievement is indirectly associated with SES through the mediating effect of self-concept across cultures. These results suggest that families with a high SES should help children form and sustain a positive self-concept, which is associated with better academic school performance. Indeed, research has documented that children from families with low SES usually experience more economic hardship, a lack of various resources, and higher threats

to social identity such as negative stereotypes regarding their intellectual ability (Croizet and Claire, 1998) and social belonging (Veland et al., 2009). These disadvantaged economic and psychological conditions may pose a chronic threat to children's self-integrity, undermining their academic performance (Cohen et al., 2006; Walton and Cohen, 2011). These results were also consistent with the view of the sociocultural self model, namely that socioeconomic conditions influence individuals' behavior through how they define themselves in a certain situation (Stephens et al., 2012).

In addition to the significant indirect effect of family SES on school achievement through self-concept, we also found that family SES directly affects school achievement. This implies that self-concept may not completely explain the relationship between family SES and school achievement. This finding differed somewhat from that of a study by Wiederkehr et al. (2015) on French children that identified the fully mediating role of self-concept in the relationship between family SES and school performance. Differences in other important factors closely related to family SES such as learning materials available in the home and parents' stimulation of their children to learn may also play important roles in the link between family SES and the school achievement of Chinese children. According to the family investment model, parents' material and interpersonal investment in their children may at least partially explain the association between family SES and children's development (Kalil and Deleire, 2004; Conger and Donnellan, 2007). Nevertheless, the results of this study demonstrated the important role of self-concept in the relationship between family SES and the academic achievement of Chinese children.

The findings of this study have important implications for possible interventions to improve academic school achievement. Although SES is a relatively stable condition and difficult to change, our results suggest that helping children form and sustain a positive self-concept may improve their academic school achievement and reduce social inequalities in child development. Accordingly, recent studies based on the sociocultural self model showed that interventions targeting the self and identity were effective in reducing the racial/ethnic achievement gap for college and high school students in the United States (Cohen et al., 2006, 2009; Sherman et al., 2013; Goyer et al., 2017). Further studies are needed to examine whether these interventions are also an effective way to reduce the social class (SES) disparities in academic achievement in China.

Finally, some limitations of this study should be acknowledged. First, convenience sampling may hinder the generalizability of the results. Second, the cross-sectional design of the study may have caused difficulties in establishing the causal relationships between variables. Previous studies showed a reciprocal relationship between self-concept and academic achievement (Marsh et al., 1999; Marsh and O'Mara, 2008). Longitudinal studies are needed to examine the dynamic relationship between SES, self-concept, and school academic achievement.

In summary, the findings of this study confirmed the role of family SES in academic achievement among junior high school students in northwest China. Furthermore, the current study also extended the extant literature by demonstrating that family SES influences children's academic school achievement partly through their self-concept in China. These findings suggested that disadvantageous family backgrounds may have a negative impact on how children defined themselves in school situation, and ultimately influence on their academic school achievement. Intervention target on help low SES students to maintain self-integrity may an effective way in reducing the social stand achievement gap for middle school students in China.

DATA AVAILABILITY STATEMENT

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

SL and RX conceived the research. QX participated in performing the research. All authors participated in writing the manuscript.

FUNDING

This research was supported by the China Postdoctoral Science Foundation (2017M613298XB) and the Science and Technology Project of Gansu Province (18JR3RA082).

REFERENCES

Adler, N. E., Boyce, W. T., Chesney, M. A., Folkman, S., and Syme, S. L. (1993). Socioeconomic inequalities in health: no easy solution. *J. Am. Med. Assoc.* 269, 3140–3145. doi: 10.1001/jama.1993.03500240084031

Aizer, A., and Currie, J. (2014). The intergenerational transmission of inequality: maternal disadvantage and health at birth. *Science* 344, 856–861. doi: 10.1126/science.1251872

Bradley, R. H., and Corwyn, R. F. (2002). Socioeconomic status and child development. *Annu. Rev. Psychol.* 21, 371–399. doi: 10.1177/1063426611421007

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., and Coll, C. G. (2001). The home environments of children in the United States part I: variations by age, ethnicity, and poverty status. *Child Dev.* 72, 1844–1867. doi: 10.1111/1467-8624.t01-1-00382

Chen, E., Matthews, K. A., and Boyce, W. T. (2002). Socioeconomic differences in children's health: how and why do these relationships change with age? *Psychol. Bull.* 128, 295–329. doi: 10.1037/0033-2909.128.2.295

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

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of Educational Opportunity*. Washington, DC: Government Printing Office, 325.

Conger, R. D., and Conger, K. J. (2002). Resilience in Midwestern families: selected findings from the first decade of a prospective, longitudinal study. *J Marriage Fam*. 64, 361–73. doi: 10.2307/3600110

Conger, R. D., and Donnellan, M. B. (2007). An interactionist perspective on the socioeconomic context of human development. *Annu. Rev. Psychol.* 58, 175–199. doi: 10.1146/annurev.psych.58.110405.085551

Conger, R. D., Wallace, L. E., Sun, Y., Simons, R. L., Mcloyd, V. C., and Brody, G. H. (2002). Economic pressure in African American families: a replication and extension of the family stress model. *Dev. Psychol.* 38, 179–193. doi: 10.1037/0012-1649.38.2.179

Costello, E. J., Compton, S. N., Keeler, G., and Angold, A. (2003). Relationships between poverty and psychopathology: a natural experiment. *J. Am. Med. Assoc.* 290, 2023–2029. doi: 10.1001/jama.290.15.2023

Croizet, J.-C., and Claire, T. (1998). Extending the concept of stereotype threat to social class: the intellectual underperformance of students from low socioeconomic backgrounds. *Pers. Soc. Psychol. Bull.* 24, 588–594. doi: 10.1177/0146167298246003

Dong, Q., and Lin, C. D. (2011). *Standardized Instruments for the Psychological Development of Chinese Children and Adolescents*. Beijing: Science Press.

Donnelly, J. E., Hillman, C. H., Castelli, D., Etnier, J. L., Lee, S., Tomporowski, P., et al. (2016). Physical activity, fitness, cognitive function, and academic achievement in children: a systematic review. *Med. Sci. Sports Exerc.* 48, 1197–1222. doi: 10.1249/MSS.0000000000000901

Goyer, J. P., Garcia, J., Purdie-Vaughns, V., Binning, K. R., Cook, J. E., Reeves, S. L., et al. (2017). Self-affirmation facilitates minority middle schoolers' progress along college trajectories. *Proc. Natl. Acad. Sci. U.S.A.* 114, 7594–7599. doi: 10.1073/pnas.1617923114

Hackman, D. A., Farah, M. J., and Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nat. Rev. Neurosci.* 11, 651–659. doi: 10.1038/nrn2897

Heyneman, S. P., and Loxley, W. A. (1983). The effect of primary-school quality on academic achievement across twenty-nine high-and low-income countries. *Am. J. Sociol.* 88, 1162–1194. doi: 10.1086/227799

Kalil, A., and Deleire, T. (2004). *Family Investments in Children's Potential: Resources and Parenting Behaviors That Promote Success*. New York: Psychology Press.

Li, C. L. (2005). Prestige stratification in the contemporary China: occupational prestige measures and socio-economic index. *Sociol. Stud.* 20, 74–102.

Linver, M. R., Brooksgunn, J., and Kohen, D. E. (2002). Family processes as pathways from income to young children's development. *Dev. Psychol.* 38, 719–734. doi: 10.1037//0012-1649.38.5.719

Liu, J., Peng, P., and Luo, L. (2019). The relation between family socioeconomic status and academic achievement in China: a meta-analysis. *Edu. Psycho. Review*. 1–28. doi: 10.1007/s10648-019-09494-0

Marsh, H. W., Byrne, B. M., and Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: reanalysis of a pioneering study and revised recommendations. *Educ. Psychol.* 34, 155–167. doi: 10.1207/s15326985ep3403_2

Marsh, H. W., and O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: unidimensional and multidimensional perspectives of self-concept. *Pers. Soc. Psychol. Bull.* 34, 542–552. doi: 10.1177/0146167207312313

Mayer, S. (1997). *What Money Can't Buy: Family Income and Children's Life Chances*. Cambridge, MA: Harvard University Press.

Muthen, L. K., and Muthen, B. O. (1998/2012). *Mplus 7.0 [Computer software]*. Los Angeles: CA: Muthen and Muthen.

National Institute of Child Health and Human Development [Nicdh]. (2005). Early child care research network: duration and developmental timing of poverty and children's cognitive and social development from birth through third grade. *Child Dev.* 76, 795–810. doi: 10.1111/j.1467-8624.2005.00878.x

NBSC (2016). *China Statistical Yearbook*. Beijing: National Bureau of Statistics of China. Available at: http://www.stats.gov.cn/tjsj/ndsj/

Neville, H. J., Stevens, C., Pakulak, E., Bell, T. A., Fanning, J., Klein, S., et al. (2013). Family-based training program improves brain function, cognition, and behavior in lower socioeconomic status preschoolers. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12138–12143. doi: 10.1073/pnas.1304437110

OECD (2016). *PISA 2015 Results: Excellence and Equity in Education, PISA*. Paris: OECD Publishing, doi: 10.1787/9789264266490-en

Rowe, D. C., and Rodgers, J. L. (1997). Poverty and behavior: are environmental measures nature and nurture? *Dev. Rev.* 17, 358–375. doi: 10.1006/drev.1997.0434

Shavelson, R. J., Hubner, J. J., and Stanton, G. C. (1976). Self-concept: validation of construct interpretation. *Rev. Educ. Res.* 46, 407–441. doi: 10.3102/00346543046003407

Sherman, D. K., Hartson, K. A., Binning, K. R., Purdievaughns, V., Garcia, J., Taborskybarba, S., et al. (2013). Deflecting the trajectory and changing the narrative: how self-affirmation affects academic performance and motivation under identity threat. *J. Pers. Soc. Psychol.* 104, 591–618. doi: 10.1037/a0031495

Sirin, S. R. (2005). Socioeconomic status and academic achievement: a meta-analytic review of research. *Rev. Educ. Res.* 75, 417–453. doi: 10.2307/3515987

Stephens, N. M., Markus, H. R., and Fryberg, S. A. (2012). Social class disparities in health and education: reducing inequality by applying a sociocultural self model of behavior. *Psychol. Rev.* 119, 723–744. doi: 10.1037/a0029028

Veland, J., Midthassel, U. V., and Idsoe, T. (2009). Perceived socio-economic status and social inclusion in school: interactions of disadvantages. *Scand. J. Educ. Res.* 53, 515–531. doi: 10.1080/00313830903301994

Wadsworth, M. E., and Achenbach, T. M. (2005). Explaining the link between low socioeconomic status and psychopathology: testing two mechanisms of the social causation hypothesis. *J. Consult. Clin. Psycho.* 73, 1146–1153. doi: 10.1037/0022-006x.73.6.1146

Walton, G. M., and Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* 331, 1447–1451. doi: 10.1126/science.1198364

Wang, L., Li, X., and Li, N. (2014). Socio-economic status and mathematics achievement in china: a review. *ZDM* 46, 1051–1060. doi: 10.1007/s11858-014-0617-8

White, S. B., Reynolds, P. D., Thomas, M. M., and Gitzlaff, N. J. (1993). Socioeconomic status and achievement revisited. *Urban Educ.* 28, 328–343. doi: 10.1177/0042085993028003007

Wiederkehr, V., Darnon, C., Chazal, S., Guimond, S., and Martinot, D. (2015). From social class to self-efficacy: internalization of low social status pupils' school performance. *Soc. Psychol. Educ.* 18, 769–784. doi: 10.1007/s11218-015-9308-8

Wong, N. Y., Wong, W. Y., and Wong, E. W. (2012). What do the Chinese value in (mathematics) education? *ZDM* 44, 9–19. doi: 10.1007/s11858-012-0383-4

Xiang, M., Gu, X., Jackson, A., Zhang, T., and Guo, Q. (2017). Understanding adolescents' mental health and academic achievement: does physical fitness matter? *Sch. Psychol. Int.* 38, 647–663. doi: 10.1177/0143034317717582

Xiao, L., and Liu, J. (2017). 家庭社会经济地位对学生学业成就的影响[The Impact of Family Socio-economic Status on Students' Academic Achievement]. *Educ. Sci. Res.* 17, 61–66.

Zhou, Q., Main, A., and Wang, Y. (2010). The relations of temperamental effortful control and anger/frustration to Chinese children's academic achievement and social adjustment: a longitudinal study. *J. Educ. Psychol.* 102, 180–196. doi: 10.1037/a0015908

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.

Copyright © 2020 Li, Xu and Xia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 10 January 2020
doi: 10.3389/fpsyg.2019.02976

Can Academic Achievement in Primary School Students Be Improved Through Teacher Training on Emotional Intelligence as a Key Academic Competency?

Teresa Pozo-Rico[1] and Ivan Sandoval[2]*

[1] Department of Developmental Psychology and Didactics, University of Alicante, Alicante, Spain, [2] Departamento de Formación Básica, Escuela Politécnica Nacional, Quito, Ecuador

Background: The primary aim of the current study was to develop a multi-methodological teacher training program based on emotional intelligence (EI) as a key competency in order to improve student academic achievement based on two methods: face-to-face instruction and game-based e-learning instruction.

Methods: Seventy-four primary education teachers and their 2069 students were randomly assigned to three groups. The first group of teachers (n = 23) were trained to use a face-to-face method. The objective of the training was that the teachers would be able to implement EI into their teaching to improve academic achievement in their students (n = 645) using face-to-face instruction. For the second group (n = 28), the teachers were trained to use an e-learning gamification method. Similar to the first group, the objective of the training was that the teachers would be able to implement EI into their teaching to improve academic achievement in their students (n = 758) using e-learning gamification instruction. The third group of teachers (n = 23) served as the controls and did not receive any special training, nor did they implement EI into their teaching (n = 666).

Results: Implementation of EI into classroom teaching effectively improved academic achievement in primary school students using both methods. However, there was a greater increase in academic achievement and higher teacher satisfaction in the game-based e-learning group. No significant differences in student achievement were observed in the control group.

Conclusion: Emotional intelligence as a key academic competency.

Keywords: primary school, academic achievement, emotional intelligence, teacher continued training, methodology and didactics, satisfaction

OPEN ACCESS

Edited by:
José Carlos Núñez, Universidad de Oviedo, Spain

Reviewed by:
Jakub Swacha, University of Szczecin, Poland
Richard Eleftherios Boyatzis, Case Western Reserve University, United States

****Correspondence:***
Teresa Pozo-Rico
teresa.pozo@ua.es

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

Received: *18 October 2019*
Accepted: *16 December 2019*
Published: *10 January 2020*

Citation:
Pozo-Rico T and Sandoval I (2020) Can Academic Achievement in Primary School Students Be Improved Through Teacher Training on Emotional Intelligence as a Key Academic Competency? Front. Psychol. 10:2976. doi: 10.3389/fpsyg.2019.02976

INTRODUCTION

Theoretical Framework

Emotional intelligence (EI) was first defined by Mayer and Salovey (1993) as a "type of social intelligence that involves the ability to monitor one's own and other's emotions, to discriminate among them, and to use the information to guide one's thinking and actions" (p. 433). Since then, numerous studies have been carried out to conceptualize the topic, develop an EI measure,

and explore EI as a key competency in education and training to improve performance in different areas (Schutte et al., 1998; Petrides and Furnham, 2000; Mayer et al., 2001; Caruso et al., 2002; Hall et al., 2019).

The use of EI as a key competency to improve academic achievement in primary education has been the subject of many studies (Pangiras et al., 2012; Temple-Harvey and Vannest, 2012; Pozo et al., 2014; Morente et al., 2017; Gilar-Corbi et al., 2018; Howard and Cogswell, 2018; Petrides et al., 2018). However, it is necessary to investigate applied programs with effective methodologies in order to train teachers to serve as enterprising leaders of EI educational projects in the classroom.

In addition, emerging evidence indicates there may be an association between EI as a key competency and school performance (Viguer et al., 2017; Petrides et al., 2018; Wolf et al., 2018) and as a competency with a mediating effect on well-being (Pangiras et al., 2012), happiness (Karavasilis et al., 2010; O'Connor, 2012), positive relationships, and a democratic and tolerant environment in the classroom (Morente et al., 2017; Ferres et al., 2018; Macias Garcia et al., 2018).

In the same way, solid evidence in the scientific literature link competencies with a behavioral approach to emotional, social, and cognitive intelligence (Boyatzis, 2008). Additionally, several studies link competencies with a behavioral approach to emotional EI and as behavioral manifestations of talent (Boyatzis, 2006, 2009). In fact, this robust perspective of the EI behavioral approach provides a more holistic theory of EI because it explains how recognizing, understanding, and using emotional information leads to effective or superior performance (Boyatzis, 2018). So, the findings of previous studies raises three key research questions:

Research Question 1: Does academic performance measured by student school records improve after the 7-week teacher training program on EI as a key competency in primary school education?

Research Question 2: Is there a difference in the change in academic performance between face-to-face instruction and game-based e-learning instruction?

Research Question 3: How does the 7-week teacher training program affect teacher satisfaction?

This study's hypotheses were as follows: (1) academic performance improves (according to student school records) after the 7-week teacher training program, (2) both of the proposed methods (face-to-face instruction and game-based e-learning instruction) improve academic performance in primary school students, and (3) teachers are satisfied with their participation in the 7-week teacher training program.

The Content, Methods, and Characteristics of the 7-Week Teacher Training Program

The 7-week teacher training program was designed to provide multi-methodological training to primary school teachers on implementing EI into their teaching as a key competency to improve academic achievement in their students. As explained above, there were two training modalities:

(1) Implementation of EI into teaching as a key competency to improve academic achievement through face-to-face instruction.
(2) Implementation of EI into teaching as a key competency to improve academic achievement through game-based e-learning instruction.

Although the methods were different, the objectives and contents of the programs were the same. The two programs were both delivered in seven sessions (one 5-h session per week, for a total of 7 weeks). The design of the program was based on Mayer and Salovey's four-branch model (1997): (1) understanding emotions; (2) identifying emotions; (3) expressing and using emotions; and (4) managing emotions. The session topics were as follows: (1) EI as a key competency in primary school, (2) understanding emotions, (3) identifying emotions, (4) expressing emotions, (5) using emotions, (6) managing emotions, and (7) conclusions.

After each session, an education project was completed in the classroom with the primary school students in order to transfer the knowledge on EI from teacher to student and capacity building was focused on to improve academic achievement. In addition, application of the two different methods were as follows:

- First session: *EI key concepts* were taught in order to improve academic achievement in primary school students through face-to-face lessons in the first method and a virtual learning environment in the second method.
- Second session: *Understanding emotions* (and also academic self-esteem, self-realization, and emotional self-consciousness) was taught in order to improve academic achievement in primary school students through roleplay illustrating the importance of understanding emotions. A final theater performance on understanding emotions was conducted for the families of students in the face-to-face group, and an understanding emotions video game created by the students for playing with their families was used for the game-based e-learning group.
- Third session: *Identifying one's own emotions and the emotions of others* (along with intrapersonal and interpersonal relationships, empathy, and social responsibility) was taught to the primary school students. In the face-to-face method, the students participated in an assembly in which they drew pictures of their classmates with different facial expressions that reflected different emotions and played a game that involved identifying emotions. In the game-based e-learning method, the activity involved virtual painting of classmates.
- Fourth session: Throughout the academic year, *expressing emotions* was taught to the students by having them use self-drawn pictures of their own facial expressions to communicate with the teacher or with the other students about their feelings, especially in conflict situations, in dealing with anxiety before exams, or for frustration over

TABLE 1 | Content, methods, and characteristics of the 7-week teacher training program.

Lesson	Goal	Face-to-face	Game-based e-learning
1	Emotional intelligence (EI) key concepts	Traditional lesson	Virtual learning environment
2	Understanding emotions	Illustration of role play	Theater performance
3	Identifying emotions	Assembly for drawing pictures of classmates	Virtual paintings of classmates
4	Expressing emotions	Assembly regarding self-drawn emotion pictures	E-learning discussion forum about self-drawn emotion pictures
5	Positive emotions	Roadmap mural painting	Mural painting upload in a learning environment
6	EI strategies	Positive reappraisal through role play and drills	Positive reappraisal through a video recording
7	Summary	Conclusions reached through an assembly	Conclusions reached through an e-learning discussion forum

grades. A weekly assembly was used for the face-to-face group, and a weekly e-learning discussion forum was used in the game-based e-learning group.

- Fifth session: *Using the power of positive emotions* in order to improve one's own positive feelings and the positive feelings of others and *using emotions to solve problems* was taught in order to improve academic achievement in the students through the emotional roadmap mural painting in the face-to-face method and a virtual mural painting upload in a learning environment in the game-based e-learning method.
- Sixth session: *EI strategies and their effectiveness* were taught to the teachers in order to improve academic achievement in the students through theory and group discussion on the analysis and solution of real EI difficulties in class. The students participated in activities that involved positive reappraisal through roleplay and drills in which families participated. A video recording was created and delivered to families by students in the face-to-face group or uploaded and made available online in the game-based e-learning group.
- Seventh session: A *summary of the program, questions, and conclusions* were discussed with the students through a weekly assembly in the face-to-face method and a weekly e-learning discussion forum in the game-based e-learning method.

Finally, for further clarification, the content, methods, and characteristics of the 7-week teacher training program have been summarized (see **Table 1**).

MATERIALS AND METHODS

Participants

A total of 74 primary school teachers (with their 2069 students) participated in this study. They participated voluntarily in this research study and were randomly assigned to three experimental conditions. The first group (n = 23) participated in the teacher training program with the objective of implementing EI into their teaching to improve academic achievement in their students (n = 645) through face-to-face instruction. The second group (n = 28) participated in the teacher training program with the objective of implementing EI into their teaching to improve academic achievement in their students (n = 758) through game-based e-learning instruction. The third group (n = 23) served as the control group and did not receive an special training, nor did the members of the group implement EI into their teaching (666 students). Of the entire sample of 74 teachers, 48.6% were women and 51.4% were men, and the average age of the teachers was 42.5 years, with a standard deviation of 1.48. There were 2069 students, 51.1% were girls and 48.9% were boys. The average age of the students was 8.48 years, with a standard deviation of 1.49.

Instruments

Two forms of data were collected in this study: (1) scores obtained by the students to appraise their academic achievement, and (2) data from a survey on teacher satisfaction with the program. Therefore, the following instruments were used:

(1) Academic performance according to student school records. Scores were assessed in seven areas: Natural Science, Social Science, Spanish Language and Literature, Mathematics, Foreign Language, Regional Language, Physical Education, and an average grade (an average of the scores obtained in each area of knowledge).
(2) Satisfaction survey. The survey was administered with the objective of determining teachers' degree of satisfaction with the 7-week training program upon its completion. The survey consisted of 10 statements for which subjects rated the extent to which they agreed or disagreed with each on a 5-point Likert scale [1 ("strongly disagree") to 5 ("strongly agree")]. The statements were as follows:

    (1) The program has facilitated my *understanding of the use of EI* as a key competency to improve academic primary school achievement.
    (2) The objectives were appropriate for the planned duration and the established *work schedule*.
    (3) I consider the *method* used to be effective.
    (4) I consider the *transfer of knowledge* to my teaching practice understandable.
    (5) I believe the knowledge that I have gained will positively affect my *methods* in the classroom.
    (6) I believe that the competencies that I have gained from the program will help me to better understand my emotions (and the emotions of my students) and to *manage* them successfully.
    (7) The 7-week teacher training program met my *expectations*.
    (8) My *motivation and interest* during the 7-week teacher training program was high.

TABLE 2 | Summary of intra-inter subject univariate analysis of variance.

	Source	Type III	df	*F*	Sig.	η^2 partial	Ob. Power	*Post hoc*
Natural Science	Intra	634.13	1	668.77	0.00**	0.24	1.00	1 < 2; 2 > 3;
	Intra*Entre	120.61	2	63.60	0.00**	0.06	1.00	3 < 1
	Error intra	1958.96	2066					
	Inter	213929.36	1	91157.38	0.00**	0.99	1.00	
	Condition	58.56	2	12.47	0.00**	0.01	0.99	
	Error inter	4848.51	2066					
Social Science	Intra	751.87	1	805.11	0.00**	0.28	1.00	1 < 2; 2 > 3;
	Intra*Entre	123.89	2	66.33	0.00**	0.06	1.00	3 < 1
	Error Intra	1929.38	2066					
	Inter	215283.29	1	89673.94	0.00**	0.97	1.00	
	Condition	91.04	2	18.96	0.00**	0.01	1.00	
	Error inter	4959.91	2066					
Spanish Language	Intra	655.79	1	688.20	0.00**	0.25	1.00	1 < 2; 2 > 3;
	Intra*Entre	144.41	2	75.77	0.00**	0.06	1.00	3 < 1
	Error Intra	1968.69	2066					
	Inter	216462.77	1	87227.84	0.00**	0.97	1.00	
	Condition	64.72	2	13.15	0.00**	0.01	0.99	
	Error inter	5084.30	2066					
Mathematics	Intra	713.73	1	790.67	0.00**	0.27	1.00	1 < 2; 2 > 3;
	Intra*Entre	123.44	2	68.37	0.00**	0.06	1.00	3 < 1
	Error intra	1864.98	2066					
	Inter	215187.86	1	90113.56	000.00**	0.98	1.00	
	Condition	90.20	2	18.88	0.00**	0.01	1.00	
	Error inter	4933.53	2066					
Foreign Language	Intra	670.43	1	743.88	0.00**	0.26	1.00	1 < 2; 2 > 3;
	Intra*Entre	160.33	2	88.95	0.00**	0.07	1.00	3 < 1
	Error intra	1861.99	2066					
	Inter	215536.29	1	87022.59	000.00**	0.97	1.00	
	Condition	64.20	2	32.10	0.00**	0.01	0.99	
	Error inter	5117.03	2066					
Regional Language	Intra	691.91	1	750.84	0.00**	0.26	1.00	1 < 2; 2 > 3;
	Intra*Entre	134.53	2	67.26	0.00**	0.06	1.00	3 < 1
	Error intra	1903.86	2066					
	Inter	214919.85	1	91597.95	0.00**	0.98	1.00	
	Condition	120.42	2	25.64	0.00**	0.02	1.00	
	Error inter	4860.71	2066					
Physical Education	Intra	759.40	1	872.68	0.00**	0.29	1.00	1 < 2; 2 > 3;
	Intra*Entre	1118.59	2	68.14	0.00**	0.06	1.00	3 < 1
	Error intra	1797.81	2066					
	Inter	215001.99	1	86121.07	0.00**	0.98	1.00	
	Condition	74.46	2	14.91	0.00**	0.01	0.99	
	Error inter	5157.78	2066					
Average grade	Intra	695.79	1	3389.17	0.00**	0.62	1.00	1 < 2; 2 > 3;
	Intra*Entre	131.09	2	319.28	0.00**	0.23	1.00	3 < 1
	Error intra	195.656	2066					
	Inter	214936.00	1	126391.42	0.00**	0.98	1.00	
	Condition	79.09	2	23.25	0.00**	0.02	1.00	
	Error inter	3513.35	2066					

*$p < 0.05$, **$p < 0.01$.

(9) I consider this teacher training program a good stimulus for *teacher performance improvement.*
(10) I consider the training engaging and beneficial for the improvement of *academic achievement in primary school.*

Procedure

All the teachers were fully informed of the details of the study (including the objectives, the responsible team, the confidentiality of their answers on the survey, and the confidentiality regarding the grades obtained by their students) prior to participation, and their written informed consent was obtained. This study was conducted in accordance with the Declaration of Helsinki.

The participants were randomly assigned to one of two experimental groups or the control group. The first experimental group participated in a 7-week teacher training program designed to improve academic achievement in their primary school students through a face-to-face method. The second experimental group participated in a 7-week teacher training program designed to improve academic achievement in their primary school students through a game-based e-learning method. The control group consisted of primary school teachers who did not participate in the program or receive any other intervention during the 7-week period.

The grades obtained by the students were recorded before and after the 7-week teacher training program in all experimental conditions. In addition, after the 7-week teacher training program, the two experimental groups completed a survey to assess their satisfaction with the program.

Experimental Design and Data Analysis

To analyze the effect of the 7-week training program on primary school academic achievement, within the general linear model procedure, the univariate split-plot variance was analyzed. Data from the students' school records were analyzed using repeated-measures analysis of variance (factors: group and time). Tests of within-subjects interaction effects (time × group) were used to identify the effects of the transfer of EI knowledge by the teacher on academic achievement in the students after the 7-week training program. Subsequently, a comparison of means was performed to analyze whether there were significant differences in academic achievement among the experimental (in the two modalities) and control groups. In addition, teacher satisfaction with the 7-week training program was analyzed. Finally, graphs of interactions, illustrating the differences found and their meanings, were created for the experimental groups (in their two modalities) and for the control group in the pre-test and post-test settings. For all statistical analyses, SPSS, version 21.0 (IBM Corporation, Armonk, NY, United States) was used.

RESULTS

To analyze whether there were differences in academic primary school achievement between the three groups before the intervention, mean contrast for the independent samples was

TABLE 3 | Marginal means.

	Group	*X*	SD	*N*
Natural Science (T1)*	1	6,79	1,16	645
	2	6,77	1,27	758
	3	6,89	1,29	666
Social Science (T1)*	1	6,80	1,18	645
	2	6,77	1,30	758
	3	6,83	1,24	666
Spanish Language (T1)*	1	6,84	1,16	645
	2	6,73	1,32	758
	3	6,89	1,33	666
Mathematics (T1)*	1	6,84	1,12	645
	2	6,76	1,29	758
	3	6,83	1,31	666
Foreign Language (T1)*	1	6,81	1,17	645
	2	6,75	1,35	758
	3	6,93	1,31	666
Regional Language (T1)*	1	6,83	1,14	645
	2	6,79	1,35	758
	3	6,82	1,26	666
Physical Education (T1)*	1	6,83	1,16	645
	2	6,73	1,27	758
	3	6,82	1,29	666
Average Grade (T1)*	1	6,82	0,94	645
	2	6,76	0,98	758
	3	6,86	0,96	666
Natural Science (T2)**	1	7,67	1,16	645
	2	7,91	1,23	758
	3	7,22	1,54	666
Social Science (T2)**	1	7,72	1,18	645
	2	8,01	1,22	758
	3	7,24	1,58	666
Spanish Language (T2)**	1	7,70	1,17	645
	2	7,95	1,24	758
	3	7,21	1,56	666
Mathematics (T2)**	1	7,69	1,15	645
	2	8,01	1,25	758
	3	7,24	1,53	666
Foreign Language (T2)**	1	7,71	1,17	645
	2	7,99	1,23	758
	3	7,22	1,53	666
Regional Language (T2)**	1	7,76	1,14	645
	2	7,99	1,24	758
	3	7,16	1,49	666
Physical Education (T2)**	1	7,69	1,16	645
	2	8,00	1,30	758
	3	7,27	1,55	666
Average Grade (T2)**	1	7,71	0,89	7,71
	2	7,98	0,74	7,98
	3	7,22	1,28	7,22

**Time 1 (T1)* = pre-training, ** Time 2 (T2)* = post-training.*

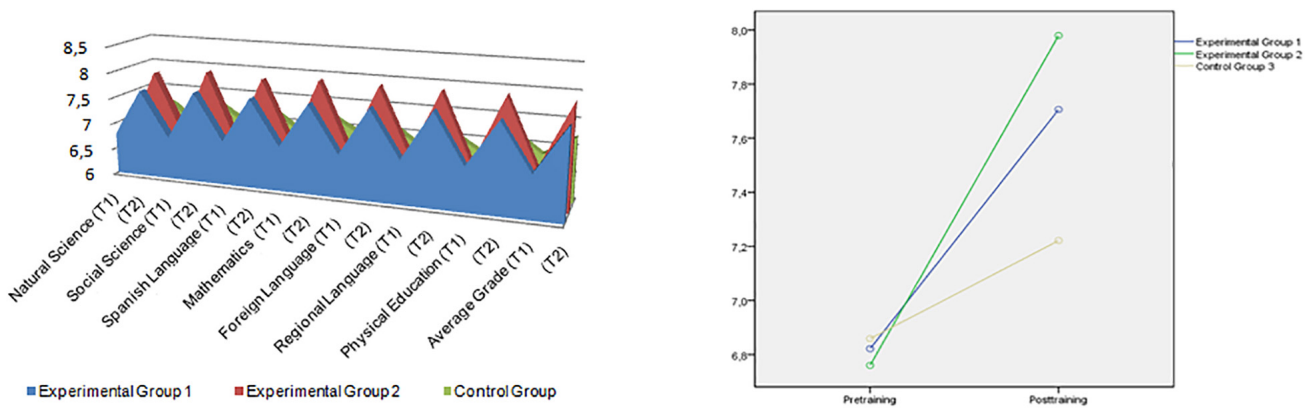


FIGURE 1 | Changes by subject and average grade for group 1 (face-to-face methodology), group 2 (e-learning gamification methodology), and control group with the implementation of a 7-week training program and administration of two tests (before and after the training).

completed. The results showed that there were no significant differences in any of the measured variables between the two groups in the pre-training. In addition, results of a Box's *M* test did not show homogeneity of the variance-covariance matrix for the Natural Science student records ($M = 96.25$; $F = 16.02$; $p \geq 0.00$), Social Science student records ($M = 104.17$; $F = 17.33$; $p \geq 0.00$), Spanish Language and Literature student records ($M = 119.25$; $F = 19.84$; $p \geq 0.00$), Mathematics student records ($M = 98.21$; $F = 16.34$; $p \geq 0.00$), Foreign Language student records ($M = 104.13$; $F = 17.33$; $p \geq 0.00$), Regional Language student records ($M = 91.56$; $F = 15.23$; $p \geq 0.00$), Physical Education student records ($M = 75.14$; $F = 12.50$; $p \geq 0.00$), or on the average grade ($M = 397.67$; $F = 66.18$; $p \geq 0.00$). It should be noted that a violation of this assumption has a minimum effect if the groups are approximately equal in size (Hair et al., 1999).

Next, contrast tests using analysis of variance showed a significant time × group interaction, as seen in **Table 2**, of intra-subject and inter-subject effects. The resulting grades from the student school records indicate that the effect of the interaction between the time of evaluation (pre-training and post-training) and the implementation of the program was significant ($p = <0.01$) between the face-to-face instruction group and the game-based e-learning instruction group compared with the control group. Differences between the experimental (in the two modalities) and control groups as reflected by changes in the average grade (and in the score in all knowledge areas) after the 7-week training program.

In addition, this improvement is clear in the two experimental groups, but the best results are achieved in the game-based e-learning group. **Table 3** presents the pre-training and post-training means of each of the groups.

Figure 1 presents interaction graphs that illustrate directions of the differences by subject and, as a conclusion, for average grades. It is important to state that there were just two tests (before and after the training) for each of the subjects and for the average mark. In this way, the results show that after the 7-week training period, significant changes were observed in students' school records overall for the face-to-face instruction and game-based e-learning instruction groups, compared with the control group, demonstrating the enhancement of coping strategies after completion of the 7-week program.

With regard to teachers' degree of satisfaction with the 7-week training program, the two experimental groups demonstrated high scores for all the statements. Ninety-seven percent of the teachers in the two groups strongly agreed with all the survey items. However, the game-based e-learning instruction group demonstrated significantly high scores for statement 3 ("I consider the method used to be effective") and statement 5 ("I believe the knowledge that I have gained will positively affect my methods in the classroom."). In that group, 100% of teachers strongly agreed with both items.

DISCUSSION

In recent decades, there has been growing evidence of the importance of EI competency in primary school (Karavasilis et al., 2010; O'Rourke et al., 2010; Kopecki and Katavic, 2011; O'Connor, 2012; Pangiras et al., 2012; Baumeister et al., 2014; Barmpouti et al., 2015; Ferres et al., 2018). The new competencies needed for academic achievement in primary school are no longer exclusively academic (Aslam, 2018; Howard and Cogswell, 2018; Vahabzadeh et al., 2018; Vasilieva, 2018). Other types of skills such as EI and everything that EI implies (e.g., self-esteem, self-realization, emotional self-consciousness, positive relationships, empathy, and social responsibility) are also factors in academic achievement.

In addition, the EI behavioral approach demonstrates that competencies and talent can be developed across the right training programs in different contexts (education, organizations, etc.) and improvement can be sustained for years (Boyatzis, 2006, 2008). Therefore, it is important in EI training to provide evidence for the relationships between EI competencies and a person's actions, life outcomes, performance, engagement, citizenship, and innovation beyond general mental ability and personality traits (Boyatzis, 2018).

This study showed that academic performance, measured by student school records, improved after teacher participation in a 7-week teacher training program on the implementation of EI as a key competency in teaching, across two different methods: (1) face-to-face instruction and (2) game-based e-learning instruction. The findings of this study indicate that it is possible to improve academic achievement across the two methods; however, similar to other studies (Mysirlaki and Paraskeva, 2015; Ros-Morente et al., 2018; Gray et al., 2019; Nikolaou et al., 2019), the improvement was found to be more significant when game-based e-learning was used. In addition, teacher satisfaction with the 7-week training program was very high for both methods.

CONCLUSION

In conclusion, EI is a key factor for educational achievement. Therefore, teachers must be qualified to teach EI in the classroom. These findings suggest that EI can be taught with different training methods using new pedagogical designs. However, one limitation of the present paper is that no follow-up survey was administered until some time had passed after the training had ended; therefore, the time period for realizing positive results of the training is not known. For this reason, further research on this topic is warranted to improve the quality of teaching in primary school.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Alicante Ethics Committee (UA-2015-07-06). The ethics committee waived the requirement of written informed consent for participation.

AUTHOR CONTRIBUTIONS

TP-R was responsible of the research, carried out the quantitative methods, and did fieldwork and program implementation. IS conducted the theoretical review of the topic.

FUNDING

This research was supported by the Spanish Ministry of Economy and Competitiveness (EDU2015-64562-R).

REFERENCES

Aslam, M. (2018). Current trends and issues affecting academic libraries and leadership skills. *Libr. Manag.* 39, 78–92. doi: 10.1108/Lm-10-2016-0076

Barmpouti, I., Vrettou, M., and Oungrinis, K. A. (2015). "The role of phenomenology in education procedure," in *Proceedings of ADVED15 International Conference on Advances in Education and Social Sciences*, Istanbul, 656–665.

Baumeister, A. E. E., Rindermann, H., and Barnett, W. S. (2014). Creche attendance and children's intelligence and behavior development. *Learn. Individ. Differ.* 30, 1–10. doi: 10.1016/j.lindif.2013.11.002

Boyatzis, R. E. (2006). An overview of intentional change from a complexity perspective. *J. Manag. Dev.* 25, 607–623. doi: 10.1108/02621710610678445

Boyatzis, R. E. (2008). Competencies in the 21st century. *J. Manag. Dev.* 27, 5–12.

Boyatzis, R. E. (2009). Competencies as a behavioral approach to emotional intelligence. *J. Manag. Dev.* 28, 749–770. doi: 10.1108/02621710910987647

Boyatzis, R. E. (2018). The behavioral level of emotional intelligence and its measurement. *Front. Psychol.* 9:1438. doi: 10.3389/fpsyg.2018.01438

Caruso, D. R., Mayer, J. D., and Salovey, P. (2002). Relation of an ability measure of emotional intelligence to personality. *J. Pers. Assess.* 79, 306–320. doi: 10.1207/s15327752jpa7902_12

Ferres, M. A., de Luna, E. B., and Sanchez, M. J. (2018). Study on emotional intelligence and contextual factors in students of fourth grade of elementary school in the province of Granada. *Rie-Rev. Investig. Educ.* 36, 141–158. doi: 10.6018/rie.36.1.281441

Gilar-Corbi, R., Pozo-Rico, T., Pertegal-Felices, M. L., and Sanchez, B. (2018). Emotional intelligence training intervention among trainee teachers: a quasiexperimental study. *Psicol. Reflexao E Crit.* 31, 1–13. doi: 10.1186/s41155-018-0112-1

Gray, S. I., Robertson, J., Manches, A., and Rajendran, G. (2019). BrainQuest: the use of motivational design theories to create a cognitive training game supporting hot executive function. *Int. J. Hum. Comp. Stud.* 127, 124–149. doi: 10.1016/j.ijhcs.2018.08.004

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1999). *Análisis Multivariante*, Vol. 491. Madrid: Prentice Hall.

Hall, S. E., Wrench, J. M., Connellan, M., Ott, N., and Wilson, S. J. (2019). The role of emotional intelligence in community integration and return to work after acquired brain injury. *Arch. Phys. Med. Rehabil.* 100, 464–473. doi: 10.1016/j.apmr.2018.06.029

Howard, M. C., and Cogswell, J. E. (2018). The "other" relationships of self-assessed intelligence: a meta-analysis. *J. Res. Pers.* 77, 31–46. doi: 10.1016/j.jrp.2018.09.006

Karavasilis, I., Samoladas, I., and Nedos, A. (2010). From empiricism to total quality management in Greek education. *Technol. Enhanced Learn. Qual. Teach. Educ. Reform* 73, 572–582. doi: 10.1007/978-3-642-13166-0_80

Kopecki, D., and Katavic, I. (2011). "Emotional intelligence: the challenge for managing resources in the globalized and changing world - a case of the insurance industry," in *Proceedings of the 5th International Scientific Conference Entrepreneurship and Macroeconomic Management: Reflections on the World in Turmoil*, Pula, CO.

Macias Garcia, D., Gonzalez Lopez, I., and Eslava-Suanes, M. D. (2018). Football as a strategy for the development of the emotional intelligence of the pupils in primary education - a didactic experience. *Educ. Form.* 3, 17–36.

Mayer, J. D., and Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence* 17, 433–442. doi: 10.1016/0160-2896(93)90010-3

Mayer, J. D., Salovey, P., Caruso, D. R., and Sitarenios, G. (2001). Emotional intelligence as a standard intelligence. *Emotion* 1, 232–242. doi: 10.1037//1528-3542.1.3.232

Morente, A. R., Guiu, G. F., Castells, R. R., and Escoda, N. P. (2017). Analysis of the relationship between emotional competences, self-esteem, classroom climate, academic achievement, and level of well-being in primary education. *Rev. Espanola Orient. Y Psicopedag.* 28, 8–18.

Mysirlaki, S., and Paraskeva, F. (2015). "Training emotionally intelligent leaders: the case of massively multiplayer online games," in *Proceedings of the 9th European Conference on Games Based Learning (Ecgbl 2015)*, Steinkjer, 687–693.

Nikolaou, I., Georgiou, K., and Kotsasarlidou, V. (2019). Exploring the relationship of a gamified assessment with performance. *Spanish J. Psychol.* 22:E6. doi: 10.1017/sjp.2019.5

O'Connor, D. (2012). "Pedagogy of freedom: why primary school teachers should embrace educational emancipation," in *Proceedings of the 5th International*

Conference of Education, Research and Innovation (Iceri 2012) (Sydney), 1100–1106.

O'Rourke, M., Hammond, S., O'Flynn, S., and Boylan, G. (2010). The medical student stress profile: a tool for stress audit in medical training. *Med. Educ.* 44, 1027–1037. doi: 10.1111/j.1365-2923.2010.03734.x

Pangiras, G., Kumarasuriar, V., and Koran, S. A. L. (2012). "A literature review of the familial and social influence on academic achievement," in *Proceedings of the Inted2012: International Technology, Education and Development Conference* (Valencia), 4527–4535.

Petrides, K. V., and Furnham, A. (2000). On the dimensional structure of emotional intelligence. *Pers. Indiv. Differ.* 29, 313–320. doi: 10.1016/s0191-8869(99)00195-6

Petrides, K. V., Sanchez-Ruiz, M. J., Siegling, A. B., Saklofske, D. H., and Mavroveli, S. (2018). "Emotional intelligence as personality: measurement and role of trait emotional intelligence in educational contexts," in *Emotional Intelligence in Education: Integrating Research with Practice*, eds K. Keefer, J. Parker, and D. Saklofske (Cham: Springer). 49-81.

Pozo, T., Gilar, R., and Gonzalez, C. (2014). "What we know about emotional intelligence, social skills and big five personality traits: improving university students' learning," in *Proceedings of the Edulearn14: 6th International Conference on Education and New Learning Technologies*, Barcelona, 6236–6242.

Ros-Morente, A., Cuenca, E. C., and Filella-Guiu, G. (2018). Analysis of the effects of two gamified emotional education software's in emotional and well-being variables in Spanish children and adolescents. *Int. J. Emerg. Technol. Learn.* 13, 148–159. doi: 10.3991/ijet.v13i09.7841

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., et al. (1998). Development and validation of a measure of emotional intelligence. *Pers. Indiv. Differ.* 25, 167–177. doi: 10.1016/s0191-8869(98)00001-4

Temple-Harvey, K. K., and Vannest, K. J. (2012). Participation and performance of students with emotional disturbance on a statewide accountability assessment in math. *Remedial Spec. Educ.* 33, 226–236. doi: 10.1177/0741932510391815

Vahabzadeh, A., Keshav, N. U., Abdus-Sabur, R., Huey, K., Liu, R. P., and Sahin, N. T. (2018). Improved socio-emotional and behavioral functioning in students with autism following school-based smartglasses intervention: multi-stage feasibility and controlled efficacy study. *Behav. Sci.* 8:E85. doi: 10.3390/Bs8100085

Vasilieva, E. V. (2018). Developing the creative abilities and competencies of future digital professionals. *Automat. Document. Math. Linguist.* 52, 248–256. doi: 10.3103/S0005105518050060

Viguer, P., Cantero, M. J., and Banuls, R. (2017). Enhancing emotional intelligence at school: evaluation of the effectiveness of a two-year intervention program in Spanish pre-adolescents. *Pers. Indiv. Differ.* 113, 193–200. doi: 10.1016/j.paid.2017.03.036

Wolf, S. J., Lockspeiser, T. M., Gong, J., and Guiton, G. (2018). Identification of foundational non-clinical attributes necessary for successful transition to residency: a modified Delphi study with experienced medical educators. *BMC Med. Educ.* 18:150. doi: 10.1186/S12909-018-1247-6

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.

Copyright © 2020 Pozo-Rico and Sandoval. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 10 January 2020
doi: 10.3389/fpsyg.2019.02841

Longitudinal Stability of Reading Difficulties: Examining the Effects of Measurement Error, Cut-Offs, and Buffer Zones in Identification

Maria Psyridou[1], Asko Tolvanen[2], Marja-Kristiina Lerkkanen[1], Anna-Maija Poikkeus[1] and Minna Torppa[1]*

[1] Department of Teacher Education, Faculty of Education and Psychology, University of Jyväskylä, Jyväskylä, Finland,
[2] Department of Psychology, Faculty of Education and Psychology, University of Jyväskylä, Jyväskylä, Finland

This study examined the stability of reading difficulties (RD) from grades 2 to 6 and focused on the effects of measurement error and cut-off selection in the identification of RD and its stability with the use of simulations. It addressed methodological limitations of prior studies by (a) applying a model-based simulation analysis to examine the effects of measurement error and cut-offs in the identification of RD, (b) analyzing a non-English and larger sample, and (c) examining RD in both reading fluency and reading comprehension. Reading fluency and reading comprehension of 1,432 Finnish-speaking children were assessed in grades 2 and 6. In addition to the use of single cut-off points on observed data, we used a simulation approach based on an estimated structural equation model (SEM) in order to examine the effect of measurement error on RD identification stability. We also examined the effect of single cut-offs by using a simulation-based buffer zone. Our results showed that measurement error affects the identification of RD over time. The use of a simulation-based buffer zone could control both the effects of measurement error and the arbitrariness of single cut-offs and lead to more accurate classification into RD groups, especially for those with scores close to the cut-offs. However, even after controlling for measurement error and using buffer zones, RD was not stable over time for all children, but both resolving and late-emerging groups existed. The findings suggest that reading development needs to be followed closely beyond the early grades and that reading instruction should be planned according to individual needs at specific time points. There is a clear need for further consideration of the mechanisms underlying the stability and instability of RD.

Keywords: reading difficulties, stability, measurement error, cut-offs, simulation, reading fluency, reading comprehension

OPEN ACCESS

Edited by:
José Carlos Núñez,
Universidad de Oviedo Mieres, Spain

Reviewed by:
Rui Alexandre Alves,
University of Porto, Portugal
Celestino Rodríguez,
Universidad de Oviedo Mieres, Spain
Raquel Fidalgo,
Universidad de León, Spain

****Correspondence:***
Maria Psyridou
maria.m.psyridou@jyu.fi

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

Received: *20 September 2019*
Accepted: *02 December 2019*
Published: *10 January 2020*

Citation:
Psyridou M, Tolvanen A, Lerkkanen M-K, Poikkeus A-M and Torppa M (2020) Longitudinal Stability of Reading Difficulties: Examining the Effects of Measurement Error, Cut-Offs, and Buffer Zones in Identification. Front. Psychol. 10:2841. doi: 10.3389/fpsyg.2019.02841

INTRODUCTION

Studies on reading difficulties (RD) have mainly focused on reading development during the early grades, and long-term longitudinal follow-up studies are scarce. The customary assumption about the persistence of RD is contested by findings in recent studies on the stability of RD. These studies indicate that in addition to persistent RD cases, there are those with resolving RD (i.e., normal

reading skills in later grades despite RD being identified during early grades) and those with late-emerging RD (i.e., RD identified during the later grades despite normal early reading skills) (e.g., Catts et al., 2012; Torppa et al., 2015; Etmanskie et al., 2016). These unstable groups are of particular interest because they may provide more understanding on the developmental risk and supportive mechanisms in RD. Such knowledge is useful for more accurate identification of children with RD and for planning support for children and youth with RD. But do these groups truly exist, or are they simply a result of measurement error? When we categorize continuous reading distribution at two time points using arbitrary cut-offs, is it possible that the changes in RD classification are actually due to measurement error?

We cannot answer these questions based on the previous studies on RD stability because of methodological limitations. Such limitations include the use of (arbitrary) cut-offs in reading skill distributions to identify RD cases (Francis et al., 2005; Branum-Martin et al., 2013), not examining the role of measurement error played in the stability of diagnosis across time, small sample sizes (Leach et al., 2003; Lipka et al., 2006), and the use of very lenient 20–25% cut-offs in RD identification (e.g., Lipka et al., 2006; Etmanskie et al., 2016). Furthermore, nearly all studies except for Torppa et al. (2015) have been conducted in English, which constrains the generalization of the findings to other orthographies. Torppa et al. (2015), however, involved only a small number of Finnish children with RD from a specific sample with family risk for dyslexia and examined only reading fluency. Our study examines the effects of measurement error and cut-offs in the longitudinal stability of RD identification across two time points, grades 2 and 6, by using a structural equation model (SEM)-based simulation approach. Examining the effects of measurement error with the use of observed data is not possible because measurement error is constant in observed variables and cannot be manipulated. It can be manipulated though with the use of simulations (Schatschneider et al., 2016). In addition to the use of single cut-off, we examined the stability of RD identification by using a simulation-based buffer zone in order to handle both the effects of the use of single cut-off on continuous distributions (Shankweiler et al., 1999) and the effects of the measurement error. In our study, we used a large-population-based sample of Finnish-speaking children and included both reading fluency and reading comprehension.

Reading Fluency and Reading Comprehension Difficulties

One of the main aims of education is to teach young children to read and use reading for learning. To be considered a skilled reader, an individual must be able to read accurately and fluently and comprehend the meaning of what was read. Reading accuracy, fluency, and comprehension are closely linked skills, particularly in the early phases of reading development, and reading comprehension can only develop after some basic word identification skills have been achieved (e.g., Florit and Cain, 2011). According to the verbal efficiency theory (Perfetti, 1985), good reading fluency skills facilitate reading comprehension because automaticity in decoding reduces the resource demands of cognitive processes (e.g., memory and attention), which can then be allocated to comprehension. The strength of the relationship between reading fluency and reading comprehension diminishes over time (Florit and Cain, 2011; Torppa et al., 2016; Santos et al., 2019) as children become "fluent enough". This finding is in accordance with the Simple View of Reading (SVR) model (Gough and Tunmer, 1986; Hoover and Gough, 1990), which states that reading comprehension is a product of two separable abilities – decoding and linguistic comprehension. Several studies over the past 30 years have supported the SVR model (e.g., Catts et al., 2006; Tunmer and Chapman, 2012; Torppa et al., 2016; see García and Cain, 2014, for a meta-analysis). Based on SVR, it is thus expected that, although decoding and reading comprehension are highly related skills, four discrete groups of RD can be identified: only decoding difficulties, only reading comprehension difficulties, both decoding and reading comprehension difficulties, and no RD. Evidence for the dissociated reading skill development in decoding (accuracy and/or fluency) and reading comprehension has been shown repeatedly (e.g., Catts et al., 2003; Torppa et al., 2007; Florit and Cain, 2011).

As reading fluency and reading comprehension are associated, but difficulties with either can emerge; the examination of RD stability should include both. We base our RD stability analysis on a model that includes reading fluency and comprehension, their stability in time, their correlations at both time points, and cross-lagged effects. We do not include reading accuracy in our model because in transparent orthographies (Aro and Wimmer, 2003; Seymour et al., 2003; Aro, 2017) acquisition of high accuracy is a fast process and because almost all children read accurately after 1 year of formal instruction of reading (Lerkkanen et al., 2004; Landerl and Wimmer, 2008; Soodla et al., 2015).

RD Stability Over Time and the Challenge With Measurement Error in Cut-Off-Based RD Identification

Five studies have addressed the stability of RD classification over time (Leach et al., 2003; Lipka et al., 2006; Catts et al., 2012; Torppa et al., 2015; Etmanskie et al., 2016). Overall, these studies suggest that while some children have stable RD, many also show difficulties only in later grades despite good early reading skills (late-emerging RD). Some studies have also reported that there are individuals whose RD resolve over time (Catts et al., 2012; Torppa et al., 2015). One of the most cited studies of late-emerging poor readers was conducted by Leach et al. (2003), in which 161 grade 4 and grade 5 students were assessed based on school and parental reports. As a cut-off, they used a combination of the research-based estimation that 10–20% of children have RD and the range of reading abilities (including both word reading and reading comprehension measures) of the group with early school-identified and persistent RD. Thirty-one children with late-emerging RD were identified: 11 with word RD, 10 with reading comprehension difficulties, and 10 with both word reading and reading comprehension difficulties. In another longitudinal study, Lipka et al. (2006) examined word reading among 44 children with early RD, of which 22 had RD in grade 4

and 22 were typical readers. Using the 25th percentile as a cut-off in a standardized oral reading test, three subgroups of RD were identified: persistent RD (N = 7), borderline RD (N = 7), and late-emerging RD (N = 8). However, in their studies Leach et al. (2003); Lipka et al. (2006), and Torppa et al. (2015) used single cut-off points on raw variables with measurement error included. Consequently, it is possible to include false-positive and false-negative cases, which can affect the reliability of the results and the prevalence of each group.

Setting a cut-off value in reading ability distribution is a practical tool for identifying RD. However, it causes uncertainty in research findings because of measurement error (Francis et al., 2005; Branum-Martin et al., 2013; Schatschneider et al., 2016). When we use a reading test, there is going to be some measurement error, and as a result, setting a cut-off based on the raw scores will lead to misclassifications. If an individual's score is slightly above the cut-off, the measurement error could cause his/her observed score to fall below the cut-off, leading him/her to be falsely identified as having RD. Measurement error can thus affect the accurate identification of children with RD. It can also affect the stability of RD, because changes in RD status can reflect either a true change or be due to the effect of measurement error at either or both time points.

The effects of the use of cut-off points on raw variables with measurement error included were tested in the only non-English RD stability study (Torppa et al., 2015). In this Finnish study, the stability of RD status was examined from grades 2 to 8. In the study, 182 children participated, of which 101 had family risk for RD and 81 had no risk. Three reading speed tasks were used for the identification of RD in grades 2 and 8, and the 10th percentile was used as a cut-off point in the distribution of the children without family risk. Four groups were identified: no RD (N = 127), late-emerging RD (N = 18), resolving RD (N = 15), and persistent RD (N = 22). In addition to the cut-off-based identification of RD groups, a simulation approach was used to examine how many children would have changed their group due to the unreliability of the measurement. The simulation results showed that 10 out of 33 RD children were misclassified. While the simulation did confirm that RD seem not to be stable for all participants, it also showed that measurement error had a clear effect.

A different methodology was used by Catts et al. (2012), who examined the prevalence of late-emerging RD with a form of latent transition analysis (LTA) in a sample of 493 children followed through grades 2, 4, 8, and 10. Importantly, they examined RD in both word reading and reading comprehension. The LTA provides a good solution for modeling the transitions between latent classes over time. At each time point, four latent classes were allowed in the model: typical reader, word RD, reading comprehension difficulties, and both word reading and reading comprehension difficulties. The cut-off of 1 standard deviation below the weighted sample mean was used for the identification of RD at each time point. In their analysis, six groups were identified across time, two stable over time, and four with transitions. This study controlled measurement error in the LTA mover–stayer model by using multiple binary indicators. Although measurement error was controlled, the actual effects on transitions were not addressed. In addition, although all possible combinations for late-emerging RD were examined (late-emerging word RD, late-emerging reading comprehension difficulties, late-emerging word RD and reading comprehension difficulties), the persistent and resolving classes were not examined.

It is possible though that the single cut-offs contribute to false impressions of the distinctness of the RD groups if many resolving and late-emerging individuals are scoring just above the cut-off value. Shankweiler et al. (1999) argued that using a buffer zone instead of simple cut-off points divides better those with and those without RD. The latest edition of the American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) includes severity ratings, which reflect the idea of continuous reading distribution. Developmental disorders are the results of many risk factors and are better seen as dimensional disorders rather than diagnostic categories (Pennington, 2006; Snowling and Hulme, 2012; van Bergen et al., 2014). This is also in line with the notion of the arbitrariness of cut-off points for the classification into learning disorder groups (Moll et al., 2014). Etmanskie et al. (2016) used a buffer zone in their examination of the prevalence and persistence of late-emerging and early identified RD in a sample of 964 children. They used the 25th percentile as a cut-off for the identification of RD (word reading and/or reading comprehension), but to be considered a typical reader, a child's score needed to be at or above the 35th percentile (a buffer zone between the 25th and 35th percentiles). In grade 4, five groups were identified: typical reading skills (N = 694), word RD (N = 7), reading comprehension difficulties (N = 121), word reading and reading comprehension difficulties (N = 24), and borderline reading skills (N = 118). The children with poor word reading and/or reading comprehension difficulties were further regrouped into early identified, late-emerging, and inconsistent readers based on their performance in grades 1, 2, and 3.

The Present Study

The aim of the current study is to examine the effects of measurement error and the effects of single cut-offs in the stability of RD identification from grades 2 to 6. The study aims to address the methodological limitations of the previous studies on RD stability by (a) applying a model-based simulation analysis to examine the effects of measurement error and single cut-offs in the identification of RD, (b) analyzing a non-English and larger sample relative to previous ones, and (c) examining RD in both reading fluency and reading comprehension. The advantages of the use of simulation are threefold: we can examine the effects of measurement error and the effects of single cut-offs on transitions; we can examine all possible combinations for reading fluency and reading comprehension between grades 2 and 6 for all the groups (persistent RD, late-emerging RD, resolving RD); and our larger sample allows us to identify more groups that probably would not have been identified in observed data because some of these groups might not have been big enough. In addition, in our study, we focus on the beginning and the end of primary school in Finland. Because in grade 1 (during which formal instruction of reading begins), it would

be difficult to assess accurately reading comprehension, our first assessment point was grade 2. Grade 6 is the last grade of primary school. After that, children enter junior high school and high school where they are taught by subject teachers instead of the classroom teacher.

The research questions of the present study are as follows:

1. How stable are reading fluency and reading comprehension RD from grades 2 to 6?
2. What is the effect of measurement error on the estimation of RD stability over time from grades 2 to 6?
3. What is the effect of using single cut-offs compared with a buffer zone when examining RD stability over time from grades 2 to 6?

We expect that cases of late-emerging and resolving RD would also emerge in the present data (e.g., Leach et al., 2003; Catts et al., 2012; Torppa et al., 2015; Etmanskie et al., 2016). We also anticipated that the simulation approach would reveal an effect of measurement error (Schatschneider et al., 2016) and that the use of a buffer zone would lower the percentages of changing RD groups (resolving and late-emerging).

MATERIALS AND METHODS

Participants and Procedure

The present study is part of the longitudinal First Steps Study (Lerkkanen et al., 2006), a follow-up of approximately 2,000 children from kindergarten to grade 6. The aim of the First Steps Study is to examine the development of children's learning and motivation in the family and school contexts. The sample was drawn from four municipalities: two in central, one in western, and one in eastern Finland. In three of the municipalities, an invitation was sent through schools to the whole age cohort of children, and in the fourth (urban) municipality, the invitation for participation was sent to approximately half of the age cohort. At the beginning of the study, the children's parents and teachers were asked for written consent. Of the parents who were contacted, 78–89%, depending on town or municipality, agreed to take part in the study. Of the children's mothers, 7.6% had no education beyond secondary school, 30.2% had a vocational school degree, 23.8% a vocational college degree, 9.9% a bachelor's degree, 24% a master's degree, and 4.6% a doctoral degree. Of the children's fathers, 7.9% had no education beyond secondary school, 33.2% had a vocational school degree, 23.7% a vocational college degree, 9.9% a bachelor's degree, 19% a master's degree, and 6.3% a doctoral degree. Parental education distribution was very close to the national distribution of Finland (Statistics Finland, 2007). The sample was highly homogeneous in ethnic and cultural background (e.g., Finnish-speaking schools and students). The study was approved by the Ethical Committee of the University of Jyväskylä and at the beginning of the study the children's parents, and teachers provided informed written consents for participation. During the study, also the children gave their written consent to participate.

The present study involved assessments at two time points – the end of grade 2 (Spring 2009) (8 years) and the end of grade 6 (Spring 2013) (13 years). Only children (N = 1,432; 662 girls and 770 boys) for whom data were available for both the grade 2 and grade 6 assessments were included in the analyses. All participants who were assessed in grade 6 were included in the current sample. In grade 6 spring, 1,824 12- to 13-year-old children participated: 863 girls (47.31%) and 961 (52.69%) boys. Of them, 1,458 participated also in grade 2 spring (8–9 years old): 680 girls (46.64%) and 778 (53.36%) boys. The sample size of the First Steps Study changed somewhat each year due to factors as shifts in teaching groups or absences during the testing days. In the present study, data from 72 schools and 147 classrooms were used. The SEM described below for the development of reading fluency and reading comprehension from grades 2 to 6 was also constructed using full grade 6 data. There were only minor differences to some of the path estimates ranging from 0.01 to 0.05.

Conducting such a long-term follow-up study is challenging, and some changes in the sample from one assessment time point to another are inevitable. We conducted a missing value analysis in order to examine if missingness was random for the data we used at the current study (z-scores for reading fluency composite and reading comprehension in grades 2 and 6). We used the Little's (1988) tests of missing completely at random (MCAR), which showed that the data were not MCAR, $\chi^2(14) = 62.29$, $p < 0.001$. Reading fluency composite score and reading comprehension in grade 2 had 20.34 and 21.27% of the cases missing. Reading fluency composite score and reading comprehension in grade 6 had 0.22 and 0.16% of the cases missing. The one-way ANOVA analysis comparing the reading fluency and reading comprehension performance of the sample included in this study (only those who were assessed in both grades 2 and 6; N = 1,432) and the whole sample (N = 1,824) showed that those who participated in both grades were somewhat better readers than those who participated only in grade 6 [for reading fluency: $F(1,1,812) = 47.08$, $p < 0.001$; for reading comprehension: $F(1,1,819) = 7.46$, $p < 0.01$]. However, the effect sizes were small for reading fluency ($d = -0.39$) and negligible for reading comprehension ($d = -0.15$).

The comprehensive education, grades 1–9, starts at the fall from the year in which the child turns 7 years of age, which is rather late compared with other countries. Before entrance to elementary school, all 6-year-olds attend 1-year kindergarten education. One goal of kindergarten education is to arouse children's interest in texts and reading and to support emerging pre-reading skills, instead of a systematic instruction of decoding (Lerkkanen, 2018). However, children are read to, and they are also encouraged to play with letters, phonemes, and words (Lerkkanen, 2019). Reading instruction begins at grade 1, and it is based on grapheme–phoneme correspondence (phonics) and a highly transparent Finnish orthography, which makes reading acquisition relatively easy and quick for children (Lerkkanen, 2007). At the end of kindergarten, around 30% of the children can read fluently, around 30% can decode easy words while around 30% of the children show no sign of reading (Lerkkanen et al., 2010; Soodla et al., 2015). Largely due to the consistent nature of the highly transparent orthography of the Finnish language (Aro, 2017), reading accuracy hits a ceiling

after a few months of formal reading instruction in grade 1 (Lerkkanen et al., 2004), and basically all children can read accurately by the end of the first school year (Soodla et al., 2015). However, even a highly consistent orthography does not guarantee efficient reading acquisition for all children. RD are typically identified for approximately 5–20% of children in either reading fluency or comprehension, depending on the criteria (Lerkkanen et al., 2010).

In basic education, children do not need to have an official diagnosis in order to have access to special educational services. Teachers and parents along with the students assess the need for extra support (Björn et al., 2016). The most common form of special educational services is the part-time special education provided by a special education teacher (Statistics Finland, 2005). In this form of special education, students study in general education classes and receive support 1–2 h/week from a special education teacher. This kind of support focuses on reading, spelling, and math difficulties. It is implemented in small groups (typically three to four students) or individually if the student faces long-lasting or more severe difficulties or if the student faces difficulties in more than one learning areas (Holopainen et al., 2018).

Measures

Reading Fluency

There were three group-administered tests for the assessment of reading fluency: a word reading fluency task, a word-chain task, and a sentence reading task. Cronbach's alpha reliability coefficient for the fluency composite was 0.79 in grade 2 and 0.77 in grade 6.

Word reading fluency task

The word reading fluency task is a subtest of the nationally normed reading test battery [ALLU–Ala-asteen lukutesti (ALLU–Reading Test for Primary School); Lindeman, 2000]. Each of the 80 items consisted of a picture with four phonologically similar words attached to it. The child silently read the four words and then drew a line to connect the picture with the word, semantically matching it. The words and pictures were frequently used words familiar to young children. For example, there was a picture of a bunny (pupu in Finnish) and the correct word along with three distractors (English word is in parentheses): pipo (cap), papu (bean), and apu (help). Completing the test requires fluent decoding. The score was the number of correct answers within a 2-min time limit. Because of the nature of this timed test, the score reflects both the child's fluency in reading the stimulus words and accuracy in making the correct choice from among the alternatives. According to the test manual (Lindeman, 1998), the Kuder–Richardson reliability was 0.82 in grade 2 and 0.97 in grade 6.

Word-chain task

The word-chain task (Nevala and Lyytinen, 2000) is a timed test with 10 rows of word chains comprising four to six words written together without spaces. The child silently read the words in the chains and, while reading them, indicated the word boundaries by drawing a division line between words. The score was the number of correct responses (maximum 40) within the time limit (1 min 25 s in grade 2). In our sample, the Pearson correlation coefficient between grades 2 and 6 was 0.52.

Sentence reading task

The Test of Silent Reading Efficiency and Comprehension (TOSREC; Wagner et al., 2009; Finnish version by Lerkkanen and Poikkeus, 2009) was used to assess silent reading efficiency in grade 2. Respondents were given 3 min to read 60 sentences and verify the truthfulness of as many sentences as possible. In grade 6, a similar task was used, the Salzburg Sentence Reading Test (Landerl et al., 1997, translated into Finnish by Sini Huemer; Pichler and Wimmer, 2006). Respondents were given 2 min to read 69 sentences and verify the truthfulness of as many sentences as possible. The sum score was based on the number of correct items. In our sample, the Pearson correlation coefficient between grades 2 and 6 was 0.67.

Reading Comprehension

A group-administered subtest of a nationally normed reading test battery (ALLU; Lindeman, 2000) was used to assess reading comprehension. The children silently read a fiction story and then answered 11 multiple-choice questions and one question in which they had to arrange five statements in the correct sequence based on information gathered from the text. The text contained 114 words in grade 2. The child received 1 point for each correct answer (max = 12). Each child completed the task at his or her own pace, but the maximum time allotted was 45 min. Lindeman (2000) reported Kuder–Richardson reliability coefficients of 0.80 in grade 2 and 0.74 in grade 6.

Analysis Description

First, we identified the RD groups using cut-off points on the observed data in a similar fashion as previous RD stability studies. We first calculated z-scores for reading comprehension and z-scores for the three reading fluency tasks in grades 2 and 6 separately. Based on the z-scores for reading fluency, we calculated mean composite scores for each grade. We used the 10th percentile as the cut-off value and dichotomized the reading fluency and reading comprehension z-score variables accordingly. The four variables were coded at each time point as 0 = typical reader (above the 10th percentile) and 1 = RD (below the 10th percentile) for each case (**Table 3**).

The simulation analysis on the effects of measurement error on the RD grouping started by building a SEM for the development of reading fluency and reading comprehension from grades 2 to 6 (see **Supplementary Appendix A**). The model was constructed using four latent variables consisting of separate factors for reading fluency and reading comprehension in grades 2 and 6 (see **Figure 1**). The use of latent variables for reading fluency (composed of three measures) and reading comprehension provide reading measures that do not include measurement error. Because reading comprehension had only one measure at each time point, we calculated the correction of attenuation using the Kuder–Richardson reliability estimates for reading comprehension in each grade from the test manual (Lindeman, 2000). In this way, we can set measurement error also

TABLE 1 | Descriptive statistics for reading fluency and reading comprehension measures in grades 2 and 6.

	N	Minimum	Maximum	Mean	SD	Skewness	Kurtosis
Grade 2							
Word reading fluency task	1,458	3	58	24.75	7.39	0.46	0.11
Word-chain task	1,458	0	35	11.59	5.93	0.61	0.31
Sentence reading task	1,453	3	60	30.57	8.15	0.02	0.26
Reading comprehension	1,436	0	12	8.73	2.58	−0.79	−0.14
Grade 6							
Word reading fluency task	1,820	10	80	47.22	10.94	0.00	0.00
Word-chain task	1,820	1	40	21.63	7.44	0.09	−0.36
Sentence reading task	1,822	4	62	30.61	7.38	0.15	0.28
Reading comprehension	1,821	0	12	7.15	2.55	−0.20	−0.59

in reading comprehension. The model included the stability paths within the reading constructs, the cross-lagged paths between reading fluency and comprehension across time, and correlations between reading fluency and comprehension at both time points.

The structural equation model analyses were carried out using Mplus 7.4 software. Maximum likelihood estimation with robust standard errors (MLR) was used for the analysis. To evaluate model fit, chi-square values and a set of fit indexes were used as follows: (a) the comparative fit index (CFI); (b) Tucker–Lewis index (TLI); and (c) the root-mean-square error of approximation (RMSEA). Good model fit is indicated by a small, preferably non-significant χ^2, CFI > 0.95, TLI > 0.95, and RMSEA < 0.06 (Hu and Bentler, 1999). Since the chi-square test depends on sample size and is sensitive to a large sample size, the chi-square statistics were not regarded as conclusive.

Next, we estimated model parameters of the SEM model to produce two simulated datasets with 200,000 cases. The first dataset was simulated using parameters related to latent factors, hence corresponding to true scores without measurement error (see **Supplementary Appendix B**). The second dataset used all the parameters in the model to produce data that include also the measurement error (see **Supplementary Appendix C**). In this dataset, the error covariances were also included in the simulation equations to correspond to the observed situation. The simulations were produced using the Statistical Package for the Social Sciences (SPSS). We used the estimates for the stability paths, the cross-lagged paths, and the correlations between reading fluency and reading comprehension variables. For each simulation, we entered four discrete variables into the model: fluency in grade 2, fluency in grade 6, comprehension in grade 2, and comprehension in grade 6. The scores for each of the four variables were coded at each time point as 0 = typical reader (highest 90%) and 1 = RD (lowest 10%) for each case. Next, we calculated the frequencies of RD groups in each simulated dataset. By comparing the percentages of RD groups in the simulated samples without and with measurement error, we can examine the effect of measurement error on RD identification and RD identification stability.

The final step of our analysis was to examine RD stability with the use of a buffer zone. We used the same procedure described above, but the four discrete variables that were entered into the model were coded at each time point as 0 = typical reader (highest 75%), 1 = RD (lowest 10%), and 2 = borderline score (lowest 10–25%) for each case. Those with scores in the lowest 10% of the reading fluency and/or reading comprehension distribution were identified as manifesting reading fluency and/or reading comprehension difficulties. Those with scores in the lowest 10–25% range were identified as borderline readers, and those with scores above 25% were identified as having no RD (**Tables 7**, **8**). In this analysis, we focused specifically on the late-emerging and the resolving RD in order to examine how far the scores are from the cut-off and how distinct the two groups are that were identified using the single cut-off.

RESULTS

Descriptive Statistics

See **Table 1** for descriptive statistics and **Table 2** for the correlations between reading fluency and reading comprehension measures in grades 2 and 6. For the reading measures, the stability correlations between grades 2 and 6 ranged from moderate to high (0.49–0.67).

Identification of the RD Groups With Observed Data

First, RD groups in reading fluency and reading comprehension in grades 2 and 6 were calculated from observed data using the 10th percentile as the cut-off value (**Table 3**). There was a

TABLE 2 | Correlations between reading fluency and reading comprehension measures.

	1	2	3	4	5	6	7	8
Grade 2								
(1) Word reading fluency task	1							
(2) Word-chain task	0.48	1						
(3) Sentence reading task	0.66	0.53	1					
(4) Reading comprehension	0.31	0.40	0.43	1				
Grade 6								
(5) Word reading fluency task	0.58	0.38	0.56	0.31	1			
(6) Word-chain task	0.43	0.52	0.53	0.39	0.52	1		
(7) Sentence reading task	0.53	0.44	0.67	0.45	0.63	0.52	1	
(8) Reading comprehension	0.23	0.29	0.33	0.49	0.22	0.32	0.38	1

statistically significant association between RD in grade 2 and RD in grade 6 [$\chi^2(9) = 441.71$, $p < 0.001$].

Approximately 10% of the participants had persistent RD, most of whom had persistent single reading fluency or reading comprehension difficulties. Most of those with either reading fluency or reading comprehension difficulties in grade 2 continued to have the same difficulty in grade 6, or they developed both reading fluency and comprehension difficulties in grade 6. Most of those with both reading fluency and reading comprehension difficulties in grade 2 continued to have only reading fluency difficulties or both reading fluency and reading comprehension difficulties in grade 6. It was rare for those struggling with any type of RD in grade 2 to move to no RD group in grade 6 and the other way around.

The percentage of children with no RD in grade 2 but some type of RD in grade 6 (late-emerging cases) was quite high (12.71%), and two-thirds of these cases had late-emerging reading comprehension difficulties. Only 2.79% of the children had only late-emerging reading fluency difficulties. The percentage of resolving difficulties (some kind of RD in grade 2 but no RD in grade 6) was smaller (6.42%), and the prevalence of resolving reading fluency and resolving reading comprehension difficulties was quite similar. Both resolving and late-emerging profiles were rare among children with both reading fluency RD and comprehension RD.

In order to describe the severity of RD in each group, we calculated means and standard deviations of the RD groups in reading fluency and reading comprehension in grades 2 and 6 (**Table 4**). Those with persistent RD performed 1.5 or more standard deviations below the average level in either fluency, comprehension, or both. Similarly, those with resolving RD performed 1.5 or more standard deviations below the average in grade 2, and those with late-emerging RD performed 1.5 or more standard deviations below the average in grade 6. However, in the resolving RD groups, the grade 6 skill performance was still below average, particularly for reading fluency (as low as −0.88 in the combined group). Similarly, the grade 2 reading levels of late-emerging RD groups were also somewhat below average (as low as −0.76 in the combined group).

Gender was unevenly distributed in the groups, $\chi^2(3) = 15.32$, $p < 0.05$: there were less boys than expected in the group with no RD in both grades (68.38% of boys compared with 77.80% of girls with adjusted standardized residual = −3.43) (**Table 5**). In the RD groups, there were no significant differences. Of the boys, 11.55% belonged to the persistent RD groups, 14.39% to the late-emerging group, and 5.68% to the resolving group with adjusted standardized residuals of 2.69, 2.50, and −0.29, respectively. Among the girls, 6.72% belonged to the persistent RD groups, 9.37% to the late-emerging groups, and 6.11% to the resolving groups.

Simulation-Based Identification of the RD Groups: The Effect of Measurement Error

In order to examine whether measurement error affects the identification and stability of RD across time, we produced one simulation without and another with measurement error.

TABLE 3 | Number and percentage of children in each group based on the observed data for those who were assessed in both grades 2 and 6 using the 10th percentile as a cut-off point.

			Grade 6				
			RF only	RC only	RF + RC	no RD	Total
Grade 2	RF[a] only	Number of children	35	17	9	43	104
		% within grade 2	33.65	16.35	8.65	41.35	100.00
		% of total	2.44	1.19	0.63	3.00	7.26
		Adjusted residual	11.58	1.26	3.05	−9.20	
	RC[b] only	Number of children	7	30	13	38	88
		% within grade 2	7.95	34.09	14.77	43.18	100.00
		% of total	0.49	2.09	0.91	2.65	6.15
		Adjusted residual	0.54	6.36	6.05	−7.99	
	RF + RC	Number of children	12	2	14	11	39
		% within grade 2	30.77	5.13	35.90	28.21	100.00
		% of total	0.84	0.14	0.98	0.77	2.72
		Adjusted residual	6.19	−1.40	11.31	−7.50	
	no RD[c]	Number of children	40	129	13	1,019	1,201
		% within grade 2	3.33	10.74	1.08	84.85	100.00
		% of total	2.79	9.01	0.91	71.16	83.87
		Adjusted residual	−11.27	−4.42	−11.10	15.03	
Total		Count	94	178	49	1,111	1,432
		% within grade 2	6.56	12.43	3.42	77.58	100.00
		% of total	6.56	12.43	3.42	77.58	100.00

[a]*RF, reading fluency;* [b]*RC, reading comprehension;* [c]*no RD, no reading difficulties.*

TABLE 4 | Descriptive statistics for the groups based on the observed data.

	N	Reading fluency, grade 2^z		Reading comprehension, grade 2^z		Reading fluency, grade 6^z		Reading comprehension, grade 6^z	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Persistent									
RC[a] → RC	30	−0.37	0.60	−2.15	0.31	−0.37	0.70	−1.58	0.35
RC → RF[b]	7	−0.78	0.41	−1.99	0.30	−1.49	0.16	−0.71	0.31
RC → RF + RC	13	−0.85	0.24	−1.92	0.17	−1.66	0.21	−1.60	0.35
RF → RC	17	−1.47	0.19	−0.78	0.64	−0.63	0.34	−1.54	0.33
RF → RF	35	−1.58	0.33	−0.34	0.66	−1.75	0.40	−0.14	0.78
RF → RF + RC	9	−1.81	0.40	−0.67	0.64	−2.04	0.38	−1.74	0.38
RF + RC → RC	2	−1.66	0.38	−1.83	0.00	−1.19	0.10	−1.66	0.00
RF + RC → RF	12	−1.95	0.36	−2.37	0.42	−1.89	0.41	−0.29	0.56
RF + RC → RF + RC	14	−1.88	0.46	−2.24	0.44	−1.95	0.50	−1.88	0.40
Late emerging									
no RD → RC	129	−0.06	0.75	−0.21	0.79	−0.05	0.67	−1.54	0.36
no RD → RF	40	−0.64	0.47	−0.28	0.77	−1.56	0.19	−0.01	0.63
no RD → RF + RC	13	−0.76	0.32	−0.70	0.46	−1.60	0.22	−1.72	0.39
Resolving									
RC → no RD[c]	38	−0.28	0.54	−2.07	0.33	−0.24	0.63	−0.28	0.57
RF → no RD	43	−1.49	0.24	−0.22	0.74	−0.69	0.48	0.09	0.73
RF + RC → no RD	11	−1.55	0.22	−2.14	0.38	−0.88	0.27	−0.34	0.59
No RD									
no RD → no RD	1, 019	0.31	0.88	0.35	0.72	0.32	0.86	0.38	0.76
Total	1, 432								

Subscript z refers to standardized score. The reading fluency measures are composite scores calculated as sum scores of the standardized scores for the three reading fluency tasks. [a]RC, reading comprehension; [b]RF, reading fluency; [c]no RD, no reading difficulties. On the left side of the arrow is the reading difficulty status in grade 2 [only reading fluency difficulties (RF), only reading comprehension difficulties (RC), both reading fluency and reading comprehension difficulties (RF + RC), or no reading difficulties (no RD)]. On the right side of the arrow is the reading difficulty status in grade 6. Reading fluency and/or reading comprehension difficulties were calculated using the 10th percentile cut-off value.

SEM Model

Latent factors for reading fluency and reading comprehension were built in grades 2 and 6 (see **Figure 1**). In addition to the regression paths across time (grades 2 and 6) and across constructs (reading fluency and reading comprehension), residual covariances for each measure of reading fluency across time were added to the model based on the inspection of modification indices. The model showed good fit with the data, $\chi^2(14) = 117.163$, $p < 0.001$, RMSEA = 0.072, CFI = 0.977, TLI = 0.954, standardized root mean square (SRMR) = 0.032. The model indicated that reading fluency was very stable across time while reading comprehension was less stable. Of grade 6 reading fluency variance, 60.8% was explained by grade 2 reading fluency and an additional 2% by grade 2 reading comprehension. Of grade 6 reading comprehension variance, 32.5% was explained by grade 2 reading comprehension and an additional 1.2% by grade 2 reading fluency.

Next, based on the model, we produced a simulated dataset without measurement error. We simulated 200,000 cases and identified cases with RD as the lowest 10% of the reading fluency and/or reading comprehension distribution (**Table 6**). In the simulated data without measurement error, 86.45% were in the stable groups (76.52% no RD; 9.93% persistent RD) and 13.55% demonstrated instability in RD across grades. Of the cases demonstrating instability, 6.50% had resolving RD and 7.05% were identified as manifesting late-emerging RD; most of these manifested late-emerging reading comprehension difficulties.

Next, we produced a simulated dataset with 200,000 cases with measurement error. Overall, 82.53% of the cases were in stable groups from grades 2 to 6 (73.92% no RD; 8.61% persistent RD) (**Table 6**). The remaining 17.48% of the cases demonstrated instability in RD across grades. Of the cases demonstrating instability, 8.52% had resolving RD, and 8.96% were identified as manifesting late-emerging RD; most of these manifested difficulties in reading comprehension.

Finally, the group sizes produced by the two simulations were compared in order to examine the effects of measurement error. For the simulation without measurement error, the percentage of the stable groups was slightly higher, and the percentage of the cases with instability in RD was somewhat lower. The results, thus, suggest that 1.32% of the children would have not been identified as having persistent RD due to the inclusion of measurement error in the analysis. In other words, in the observed data of 1,432 children, approximately 19 children would be wrongly classified. Similarly, 3.93% of the children would have changed groups due to the inclusion of measurement error in the analysis, which means that in the observed data, 56 out of 1,432 children would be wrongly classified. More specifically, 28

TABLE 5 | Prevalence of boys and girls in each group based on observed data.

Group		Gender		Total
		Male	Female	
Persistent	Count	65	33	98
	% with the groups	66.33	33.67	100
	% within gender	11.55	6.72	9.30
	% of total	6.17	3.13	9.30
	Adjusted residual	2.69	−2.69	
Late emerging	Count	81	46	127
	% with the groups	63.78	36.22	100
	% within gender	14.39	9.37	12.05
	% of total	7.69	4.36	12.05
	Adjusted residual	2.50	−2.50	
Resolving	Count	32	30	62
	% with the groups	51.61	48.39	100
	% within gender	5.68	6.11	5.88
	% of total	3.04	2.85	5.88
	Adjusted residual	−0.29	0.29	
no RD	Count	385	382	767
	% with the groups	50.20	49.80	100
	% within gender	68.38	77.80	72.77
	% of total	36.53	36.24	72.77
	Adjusted residual	−3.43	3.43	
Total	Count	563	491	1,054
	% with the groups	53.42	46.58	100
	% within gender	100	100	100
	% of total	53.42	46.58	100

children (1.91%) would have been misclassified as having late-emerging RD and 28 children (2.02%) misclassified as having resolved their RD.

The Effect of the Use of a Single Cut-Off

The final step of our analysis was to examine RD stability with the use of a buffer zone. In this analysis, we focused particularly on the late-emerging RD (**Table 7**) and resolving RD groups (**Table 8**). Similar to the use of a single cut-off, we produced two simulated datasets, one without and one with measurement error. We simulated 200,000 cases for each dataset and identified RD for those cases located in the lowest 10% of the reading fluency and/or reading comprehension distribution. Cases with scores in the lowest 10–25% range were identified as having borderline scores, and cases with scores above 25% were identified as having no RD.

The results from the use of the buffer zone suggest that most of the cases from the late-emerging and resolving RD groups actually land in the buffer zone. For the late-emerging group, the simulation without measurement error and with the buffer zone showed that only 33 cases would be identified as having late-emerging RD compared with the 101 that were identified with the use of a single cut-off. Of the 33 cases, 28 were identified with late-emerging reading comprehension difficulties and five with late-emerging reading fluency difficulties (**Table 7**). For the resolving group, the simulation without measurement error and with the buffer zone showed that 28 cases would be identified as manifesting resolving RD compared with the 94 identified with the use of a single cut-off. Of the 28 cases, 19 would be identified with resolving reading comprehension difficulties, 8 with resolving reading fluency difficulties, and 1 with resolving reading fluency and comprehension difficulties (**Table 8**).

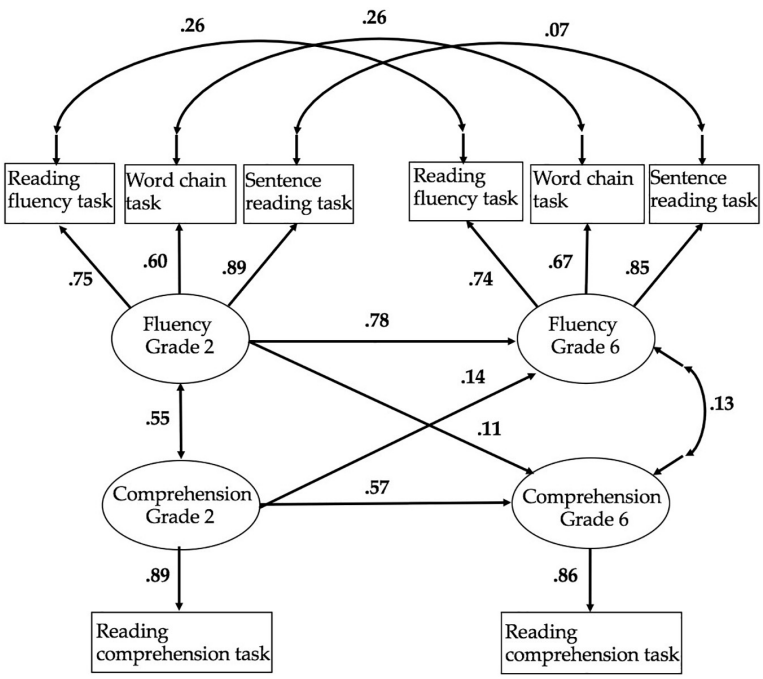


FIGURE 1 | Standardized path estimates of the cross-lagged model. The figure presents standardized estimates using listwise analysis.

TABLE 6 | Percentages and number of individuals in reading difficulty (RD) groups identified with the use of a single cut-off.

Group	Subgroup	Simulation without measurement error (group %)	Simulation with measurement error (group %)
Persistent	RC[a] → RC	1.92 (27)	1.57 (22)
	RC → RF[b]	0.54 (8)	0.73 (10)
	RC → RF + RC	0.47 (7)	0.44 (6)
	RF → RC	0.33 (5)	0.54 (8)
	RF → RF	2.93 (42)	2.56 (37)
	RF → RF + RC	0.68 (10)	0.61 (9)
	RF + RC → RC	0.36 (5)	0.35 (5)
	RF + RC → RF	1.28 (18)	1.02 (15)
	RF + RC → RF + RC	1.42 (20)	0.79 (11)
Persistent total		9.93 (142)	8.61 (123)
Late emerging	no RD[c] → RC	4.38 (63)	5.10 (73)
	no RD → RF	2.23 (32)	3.26 (47)
	no RD → RF + RC	0.44 (6)	0.60 (9)
Late-emerging total		7.05 (101)	8.96 (129)
Resolving	RC → no RD	3.50 (50)	4.38 (63)
	RF → no RD	2.48 (36)	3.41 (49)
	RF + RC → no RD	0.52 (8)	0.73 (10)
Resolving total		6.50 (94)	8.52 (122)
no RD	no RD	76.52 (1, 095)	73.92 (1, 058)

[a]RC, reading comprehension, [b]RF, reading fluency, [c]no RD, no reading difficulties. The number in parentheses shows the number of individuals based on the percentage for N = 1,432. On the left side of the arrow is the RD status in grade 2 [only reading fluency difficulties (RF), only reading comprehension difficulties (RC), both reading fluency and reading comprehension difficulties (RF + RC), or no reading difficulties (no RD)]. On the right side of the arrow is the RD status in grade 6. Reading fluency and/or reading comprehension difficulties were calculated using the lowest 10%.

DISCUSSION

The main focus of the present study was to examine the longitudinal stability of RD identification across two time points, grades 2 and 6 including both reading fluency and reading comprehension. We examined whether RD identification was stable over time even if we control for measurement error and the use of single cut-offs. Our results showed that for some children RD are not stable but also revealed a clear effect of measurement error and the selection of the cut-off in the identification of RD. The findings highlight that measurement error affects the accurate identification of children with RD by causing misclassifications and that the simplicity of the single cut-offs can contribute to false impressions on instability of RD over time.

All the previous studies, except the study conducted by Catts et al. (2012), examining the stability of RD identification, used cut-off points on raw variables with measurement error included (Leach et al., 2003; Lipka et al., 2006; Torppa et al., 2015; Etmanskie et al., 2016). This means that conclusions may be biased by false-positive or false-negative cases due to measurement error causing misclassifications. In this study, we used simulations with models that do and do not include measurement error to estimate the magnitude of this problem. Our findings comparing the results of the simulations with and without measurement error showed the impact of measurement error but still aligned with prior findings suggesting RD instability for a group of children (Leach et al., 2003; Lipka et al., 2006; Catts et al., 2012; Torppa et al., 2015; Etmanskie et al., 2016). In our study, 74–77% of participants were typical readers across time, and each RD group consisted of 7–10% of the participants (depending on the group and the model). Most of the participants (86.45% in the simulation without measurement error and 82.53% in the simulation with measurement error) demonstrated stability in their RD status (no RD and persistent RD). The simulation without measurement error revealed, however, larger proportions of the stable groups (persistent RD and no RD) and smaller proportions of late-emerging and resolving RD compared with the simulation with measurement error. These differences were expected because the results from the simulation with measurement error include false-positive or false-negative cases because of the effect measurement error, which affects the reliability of the estimation of the prevalence of each RD group. Although the differences in the prevalence of the groups were small, they show that measurement error has an effect on the longitudinal stability of RD identification. Overall, though, the findings supported that the unstable groups exist, even if we control for measurement error, which has been a problem in most previous studies.

Catts et al. (2012) study is the only previous study comparable to the present simulation analysis as they used LTA, which relied on multiple indicators for each reading class, and their findings were thus less affected by measurement error. Our results from the analyses using single cut-offs were in line with Catts et al. (2012) in that there was a higher prevalence of late-emerging reading comprehension than late-emerging reading fluency difficulties, but the findings show differences in the proportion of children in each RD group. More specifically, the present study identified a smaller proportion of late-emerging cases (7.05% compared with 13.40%) and more resolving cases (6.50% compared with 1.90%). Of the children with RD in the Catts et al. (2012) study, 42% had late-emerging RD and 6% had resolving RD; in our study, the percentages were 30 and 28%, respectively.

It is possible that differences in orthography, in assessment ages of children, or differences in the criteria used for the identification of RD, explain the differences between the present study and that by Catts et al. (2012). Finnish orthography is highly transparent, with one-on-one correspondence between phonemes and graphemes (Seymour et al., 2003; Lyytinen et al., 2015; Aro, 2017). Due to the transparency of Finnish orthography, most Finnish children learn to read accurately after

TABLE 7 | Percentage and number of individuals in late-emerging groups identified with the use of a buffer zone (bz).

Late-emerging groups identified using a single cut-off	Late-emerging groups identified using a buffer zone	Simulation without measurement error, group % (*N*)	Simulation with measurement error, group % (*N*)
no RD → RC	no RD → RC	1.69 (24)	2.30 (33)
	no RD → RF(bz) + RC	0.28 (4)	0.42 (6)
	RC(bz) → RC	1.10 (16)	0.98 (14)
	RC(bz) → RF(bz) + RC	0.29 (4)	0.32 (5)
	RF(bz) → RC	0.20 (3)	0.36 (5)
	RF(bz) → RF(bz) + RC	0.28 (4)	0.25 (4)
	RF(bz) + RC(bz) → RC	0.21 (3)	0.24 (3)
	RF(bz) + RC(bz) → RF(bz) + RC	0.33 (5)	0.22 (3)
		4.38 (63)	**5.09 (73)**
no RD → RF	no RD → RF	0.29 (4)	0.85 (12)
	no RD → RF + RC(bz)	0.09 (1)	0.22 (3)
	RC(bz) → RF	0.09 (1)	0.26 (4)
	RC(bz) → RF + RC(bz)	0.06 (1)	0.12 (2)
	RF(bz) → RF	0.82 (12)	0.92 (13)
	RF(bz) → RF + RC(bz)	0.26 (4)	0.31 (5)
	RF(bz) + RC(bz) → RF	0.37 (5)	0.37 (5)
	RF(bz) + RC(bz) → RF + RC(bz)	0.25 (4)	0.20 (3)
		2.23 (32)	**3.25 (47)**
no RD → RF + RC	no RD → RF + RC	0.04 (0)	0.13 (2)
	RC(bz) → RF + RC	0.06 (1)	0.11 (1)
	RF(bz) → RF + RC	0.14 (2)	0.18 (3)
	RF(bz) + RC(bz) → RF + RC	0.20 (3)	0.18 (3)
		0.44 (6)	**0.60 (9)**

The number in parentheses shows the number of individuals based on the percentage for N = 1,432. The bold values show the total number of late-emerging cases.

TABLE 8 | Percentages and number of individuals of the resolving groups identified with the use of a buffer zone.

Resolving groups identified using a single cut-off	Resolving groups identified using a buffer zone	Simulation without measurement error (ME), group % (*N*)	Simulation with ME, group % (*N*)
RC → no RD	RC → no RD	1.04 (15)	1.65 (24)
	RC → RC(bz)	0.77 (11)	0.82 (12)
	RC → RF(bz)	0.25 (4)	0.41 (6)
	RC → RF(bz) + RC(bz)	0.25 (4)	0.26 (4)
	RF(bz) + RC → no RD	0.27 (4)	0.42 (6)
	RF(bz) + RC → RC(bz)	0.21 (3)	0.24 (3)
	RF(bz) + RC → RF(bz)	0.35 (5)	0.36 (5)
	RF(bz) + RC → RF(bz) + RC(bz)	0.35 (5)	0.22 (3)
		3.49 (50)	**4.38 (63)**
RF → no RD	RF → no RD	0.39 (6)	0.93 (13)
	RF → RC(bz)	0.07 (1)	0.19 (3)
	RF → RF(bz)	0.97 (14)	1.06 (15)
	RF → RF(bz) + RC(bz)	0.24 (3)	0.26 (4)
	RF + RC(bz) → no RD	0.11 (2)	0.26 (4)
	RF + RC(bz) → RC(bz)	0.06 (1)	0.11 (2)
	RF + RC(bz) → RF(bz)	0.41 (6)	0.42 (6)
	RF + RC(bz) → RF(bz) + RC(bz)	0.22 (3)	0.19 (3)
		2.47 (36)	**3.42 (49)**
RF + RC → no RD	RF + RC → no RD	0.05 (1)	0.15 (2)
	RF + RC → RC(bz)	0.04 (1)	0.10 (1)
	RF + RC → RF(bz)	0.22 (3)	0.28 (4)
	RF + RC → RF(bz) + RC(bz)	0.21 (3)	0.20 (3)
		0.52 (8)	**0.73 (10)**

The number in parentheses shows the number of individuals based on the percentage for N = 1,432. The bold values show the total number of resolving cases.

few months of reading instruction in grade 1 (Lerkkanen et al., 2004), and by the time of grade 2 assessment [which was also the first assessment time point in the Catts et al. (2012) study], most are fluent readers and have a good command of reading comprehension skills (Lerkkanen et al., 2010). It is possible, then, that orthographic transparency could explain why there seem to be fewer cases of late-emerging RD among children learning to read in a context of transparent orthography as the differences in higher-level reading skills are visible already in grade 2. It is also likely that in a transparent orthography, it is possible to develop a resolving pathway more often despite early learning difficulties, as decoding task is cognitively less demanding. Additionally, Catts et al. (2012) followed the children until grade 8, whereas the present analysis extended only to grade 6. The longer gap between assessments may also increase the number of unstable RD cases. Finally, in our study, a somewhat stricter cut-off was used (10th percentile), while Catts et al. (2012) used the criterion of 1 standard deviation below the weighted sample mean (approximately 16th percentile).

Although the use of cut-offs is likely to lead to uncertainties in research findings because of measurement error (Francis et al., 2005; Branum-Martin et al., 2013), it is also a practical tool for the identification of children with RD. However, where and how we set the cut-off affects the identification of the RD groups and could possibly contribute to false impressions about the distinctness of the groups and about the risk of children for developing RD. Therefore, we need to find a way to use cut-offs without the evident problems accompanying them. One precaution against biased conclusions is the use of buffer zones around the single cut-offs (Shankweiler et al., 1999). The use of the buffer zone in the present study revealed a more complex picture than the one of the use of single cut-offs. For instance, the simulation without measurement error and without a buffer zone suggested that 4.38% of the children had late-emerging reading comprehension difficulties. However, when we use the buffer zone, we see that many of the children identified with late-emerging RD actually had borderline skills (lowest 10–25%) in reading fluency and/or reading comprehension already in grade 2. In other words, although many children passed the strict RD criterion, they nevertheless were still at the lower end of the skill distribution. Similarly, the simulation without the measurement error or buffer zone suggested that 2.48% of the children had resolving reading fluency difficulties, leading to the impression that these children were fluent readers in grade 6; in fact, most of them still scored in the borderline zone, just above the strict cut-off. The use of cut-off points is a practical tool for the identification of children with RD, but they can be problematic. Their use would be rational if we did not have a normal skill distribution, but this is not the case in reading achievement (Francis et al., 2005). Consequently, setting an arbitrary cut-off on the continuous distribution of reading achievement can lead to false or biased estimations of the prevalence of instability of the RD groups.

There are certain limitations in this study that need to be considered. First, we used only one measure for the assessment of reading comprehension in grades 2 and 6. Although we calculated the correction of attenuation in each grade in order to control measurement error, having more measures for the assessment of reading comprehension would have increased the strength of our model. Also, more measures and several time points would have allowed a more thorough and reliable assessment of reading comprehension.

In conclusion, this study shows that the use of measures with measurement error and the use of single cut-offs affect the longitudinal stability of RD identification across two time points. Comparing the prevalence of the groups arose from the use of single cut-off and those from the use of the buffer zone, it is evident that the use of single cut-off contributes to false impressions, such as how distinct the RD groups are. However, even after controlling for measurement error and using the buffer zone, our results suggest that RD are not stable over time for all children. Although many children manifest RD in the beginning of their school life and continue to have difficulties across grades, some children do not demonstrate difficulties until mid-primary school, and others may resolve their earlier difficulties by the end of primary school. Given that reading fluency and reading comprehension were not stable over time, the question arises about which additional factors affect the development of reading fluency and especially the development of reading comprehension, which was less stable. Further studies are needed to better understand the factors that could lead to late-emerging RD, either in reading fluency or in reading comprehension, as well as the factors that help children resolve their RD. Closer examination of the resolving cases could provide important information on the mechanisms that trigger protective factors. These insights could be used for the development of support systems and intervention programs, which will help children at risk for RD.

These results raise several clinical and practical implications. First, it seems that a change in the child's RD status can occur both because of the effect of measurement error and because of the instability of RD identification. Consequently, because of the presence of measurement error in every assessment tool, it is not sufficient to diagnose RD based on only one assessment. Although some children may pass the strict cut-off, the results of the buffer zone show that they may still be in jeopardy for RD. Second, the use of a buffer zone along with continuous follow-ups of children's reading development could facilitate more accurate identification of the children with RD. This is especially important in education systems in which access to remedial support or special needs interventions depends on an official diagnosis. Third, the accumulation of evidence for the instability of RD classification from this study and prior literature suggests a need for careful consideration of practices and permanency of diagnosing of RD. This is needed especially in cases where diagnostic practices deprive children with late-emerging RD of interventions or support and where individuals with resolving RD may continue to carry an inaccurate label or perception of one's skills.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

MP and MT drafted the first version of the current manuscript. AT contributed to data analysis. M-KL and A-MP were responsible for data collection and commented on the manuscript. All authors contributed to the manuscript drafting, and read and approved the submitted version.

FUNDING

This research was supported by the Academy of Finland (Grant Numbers #213486, #268586, and #292466). MP was supported by the Alli Paasikiven Foundation. MT was supported by Academy of Finland (Grant Numbers #276239, #284439, and #313768).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02841/full#supplementary-material

REFERENCES

American Psychiatric Association, (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Washington, DC: American Psychiatric Association.

Aro, M. (2017). "Learning to read finnish," in *Reading Acquisition Across Languages and Writing Systems: An International Handbook* 1st Edn, eds L. T. W. Verhoeven, and C. A. Perfetti, (Cambridge: Cambridge University Press), 416–436. doi: 10.1017/9781316155752.017

Aro, M., and Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Appl. Psycholinguist.* 24, 621–635. doi: 10.1017/s0142716403000316

Björn, P. M., Aro, M. T., Koponen, T. K., Fuchs, L. S., and Fuchs, D. H. (2016). The many faces of special education within RTI frameworks in the United States and Finland. *Learn. Disabil. Q.* 39, 58–66. doi: 10.1177/0731948715594787

Branum-Martin, L., Fletcher, J. M., and Stuebing, K. K. (2013). Classification and identification of reading and math disabilities: the special case of comorbidity. *J. Learn. Disabil.* 46, 490–499. doi: 10.1177/0022219412468767

Catts, H. W., Adlof, S. M., and Ellis Weismer, S. (2006). Language deficits in poor comprehenders: a case for the simple view of reading. *J. Speech Lang. Hear. Res.* 49, 278–293. doi: 10.1037/0708-5591.49.2.125

Catts, H. W., Compton, D., Tomblin, J. B., and Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *J. Educ. Psychol.* 104, 166–181. doi: 10.1037/a0025323

Catts, H. W., Hogan, T. P., and Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. *J. Learn. Disabil.* 36, 151–164. doi: 10.1177/002221940303600208

Etmanskie, J. M., Partanen, M., and Siegel, L. S. (2016). A longitudinal examination of the persistence of late-emerging reading disabilities. *J. Learn. Disabil.* 49, 21–35. doi: 10.1177/0022219414522706

Florit, E., and Cain, K. (2011). The simple view of reading: is it valid for different types of alphabetic orthographies? *Educ. Psychol. Rev.* 23, 553–576. doi: 10.1007/s10648-011-9175-6

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Lyon, G. R., Shaywitz, B. A., and Shaywitz, S. E. (2005). Psychometric approaches to the identification of LD IQ and achievement scores are not sufficient. *J. Learn. Disabil.* 38, 98–108. doi: 10.1177/00222194050380020101

García, J. R., and Cain, K. (2014). Decoding and reading comprehension: a meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Rev. Educ. Res.* 84, 74–111. doi: 10.3102/0034654313499616

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remed. Spec. Educ.* 7, 6–10. doi: 10.1177/074193258600700104

Holopainen, L. K., Kiuru, N. H., Mäkihonko, M. K., and Lerkkanen, M. K. (2018). The role of part-time special education supporting students with reading and spelling difficulties from grade 1 to grade 2 in Finland. *Eur. J. Spec. Needs Educ.* 33, 316–333. doi: 10.1080/08856257.2017.1312798

Hoover, W. A., and Gough, P. B. (1990). The simple view of reading. *Read. Writ.* 2, 127–160. doi: 10.1007/BF00401799

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struc. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Landerl, K., and Wimmer, H. (2008). Development of word reading fluency and spelling in a con-sistent orthography: an 8-year follow-up. *J. Educ. Psychol.* 100, 150. doi: 10.1037/0022-0663.100.1.150

Landerl, K., Wimmer, H., and Moser, E. (1997). *Salzburger Lese- und rechtschreibtest [Salzburg Reading and Spelling Test]*. Bern: Huber.

Leach, J. M., Scarborough, H. S., and Rescorla, L. (2003). Late-emerging reading disabilities. *J. Educ. Psychol.* 95, 211–224. doi: 10.1037/0022-0663.95.2.211

Lerkkanen, M.-K. (2007). "The beginning phases of reading literacy instruction in Finland," in *Finnish Reading Literacy. When Quality and Equity Meet*, eds P. Linnakylä, and I. Arffman, (Jyväskylä: University of Jyväskylä), 155–174. doi: 10.1108/s0735-004x(2010)0000023008

Lerkkanen, M.-K. (2018). "The influence of instruction on reading motivation in Finland," in *Reading achievement and Motivation in Boys and Girls: Field Studies and Methodological Approaches. Literacy Studies, 15*, eds P. O. Garcia, and P. B. Lind, (Cham: Springer), 65–78. doi: 10.1007/978-3-319-75948-7_4

Lerkkanen, M.-K. (2019). "Early language and literacy development in the Finnish context," in *The Sage Handbook of Developmental Psychology and Early Childhood Education*, eds D. Whitebread, V. Grau, K. Kumpulainen, M. M. McClelland, N. E. Perry, and D. Pino-Pasternak, (London: Sage), 403–417. doi: 10.4135/9781526470393.n23

Lerkkanen, M.-K., Niemi, P., Poikkeus, A.-M., Poskiparta, M., Siekkinen, M., and Nurmi, J.-E. (2006). *The First Steps Study (Ongoing)*. Finland: University of Jyväskylä.

Lerkkanen, M.-K., and Poikkeus, A.-M. (2009). *Lausetasoinen luetun ymmärtäminen ja sujuvuus: TOSREC-testin adaptoitu ja lyhennetty versio: Alkuportaat-tutkimuksen testimateriaalia [Sentence reading efficiency and comprehension: Adapted and shortened version of the TOSREC test: Test material of the First Steps study] [Unpublished test material]*. Finland: University of Jyväskylä.

Lerkkanen, M.-K., Poikkeus, A.-M., Ahonen, T., Siekkinen, M., Niemi, P., and Nurmi, J.-E. (2010). Luku- ja kirjoitustaidon sekä motivaation kehitys esi- ja alkuopetusvuosina [The development of reading and spelling skills from kindergarten to Grade 2]. *Kasvatus* 41, 116–128.

Lerkkanen, M.-K., Rasku-Puttonen, H., Aunola, K., and Nurmi, J.-E. (2004). Predicting reading performance during the first and the second year of primary school. *Br. Educ. Res. J.* 30, 67–92. doi: 10.1080/0141192031000162997 4

Lindeman, J. (1998). *ALLU–Ala-Asteen Lukutesti [ALLU–Reading Test for Primary School]*. Turku: University of Turku.

Lindeman, J. (2000). *Ala-asteen lukutesti: Tekniset tiedot* ([2. p.*].). [ALLU Reading Test for Primary School: Technical Information]*. Turku: University of Turku.

Lipka, O., Lesaux, N., and Siegel, L. (2006). Retrospective analyses of the reading development of Grade 4 students with reading disabilities: risk status and profiles over 5 years. *J. Learn. Disabil.* 39, 364–378. doi: 10.1177/00222194060390040901

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *J. Am. Stat. Assoc.* 83, 1198–1202. doi: 10.1080/01621459.1988.10478722

Lyytinen, H., Aro, M., Richardson, U., Erskine, J., Banff, A., Li, U. H., et al. (2015). "Reading skills, acquisition of: cultural, environmental, and developmental impediments," in *International Encyclopedia of the Social & Behavioral Sciences*, 2nd Edn, ed. J. D. Wright, (Amsterdam: Elsevier), 5–11. doi: 10.1016/B978-0-08-097086-8.23111-5

Moll, K., Kunze, S., Neuhoff, N., Bruder, J., and Schulte-Körne, G. (2014). Specific learning disorder: prevalence and gender differences. *PLoS One* 9:e103537. doi: 10.1371/journal.pone.0103537

Nevala, J., and Lyytinen, H. (2000). *Sanaketjutesti [Differentiate Word Test]*. Jyväskylä: Niilo Mäki Instituutti.

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition* 101, 385–413. doi: 10.1016/j.cognition.2006.04.008

Perfetti, C. A. (1985). *Reading Ability*. Oxford: Oxford University Press.

Pichler, C., and Wimmer, L. (2006). Das salzburger Lesescreening 2-9. Handreichnung für Lehrerinnen und Lehrer. Based on Mayringer, H., & Wimmer, H. (2003). Salzburger Lesescreening für die Klassenstufen 1-4 and Auer, M., Gruber, G., Mayringer, H., & Wimmer, H. (2005). *Salzburger Lesescreening für die Klassenstufen*. 5–8.

Santos, S., Cadime, I., Viana, F. L., and Ribeiro, I. (2019). Cross-lagged relations among linguistic skills in european portuguese: a longitudinal study. *Read. Res. Q.* doi: 10.1002/rrq.261

Schatschneider, C., Wagner, R. K., Hart, S. A., and Tighe, E. L. (2016). Using simulations to investigate the longitudinal stability of alternative schemes for classifying and identifying children with reading disabilities. *Sci. Stud. Read.* 20, 34–48. doi: 10.1080/10888438.2015.1107072

Seymour, P. H. K., Aro, M., and Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *Br. J. Psychol.* 94, 143–174. doi: 10.1348/000712603321661859

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., et al. (1999). Comprehension and decoding: patterns of association in children with reading difficulties. *Sci. Stud. Read.* 3, 69–94. doi: 10.1207/s1532799xssr0301_4

Snowling, M. J., and Hulme, C. (2012). Annual Research Review: the nature and classification of reading disorders–a commentary on proposals for DSM-5. *J. Child Psychol. Psychiatry* 53, 593–607. doi: 10.1111/j.1469-7610.2011.02495.x

Soodla, P., Lerkkanen, M. K., Niemi, P., Kikas, E., Silinskas, G., and Nurmi, J. E. (2015). Does early reading instruction promote the rate of acquisition? A comparison of two transparent orthographies. *Learn. Instruc.* 38, 14–23. doi: 10.1016/j.learninstruc.2015.02.002

Statistics Finland (2005). *Erityisopetusta Saavien Määrän Kasvu Jatkui [Number of Students Receiving Special Education has Decreased]*. Available at: http://www.stat.fi/til/erop/2004/erop_2004_2005-06-15_tie_001.html (accessed September 25, 2018).

Statistics Finland (2007). *Statistical Databases*. Available at: http://www.stat.fi/tup/tilastotietokannat/index_en.html (accessed September 25, 2018).

Torppa, M., Eklund, K., van Bergen, E., and Lyytinen, H. (2015). Late-emerging and resolving dyslexia: a follow-up study from age 3 to 14. *J. Abnorm. Child Psychol.* 43, 1389–1401. doi: 10.1007/s10802-015-0003-1

Torppa, M., Georgiou, G. K., Lerkkanen, M. K., Niemi, P., Poikkeus, A. M., and Nurmi, J. E. (2016). Examining the simple view of reading in a transparent orthography: a longitudinal study from kindergarten to grade 3. *Merrill-Palmer Q.* 62, 179–206. doi: 10.13110/merrpalmquar1982.62.2.0179

Torppa, M., Tolvanen, A., Poikkeus, A. M., Eklund, K., Lerkkanen, M. K., Leskinen, E., et al. (2007). Reading development subtypes and their early characteristics. *Ann. Dyslexia* 57, 3–32. doi: 10.1007/s11881-007-0003-0

Tunmer, W. E., and Chapman, J. W. (2012). The simple view of reading redux: vocabulary knowledge and the independent components hypothesis. *J. Learn. Disabil.* 45, 453–466. doi: 10.1177/0022219411432685

van Bergen, E., van der Leij, A., and de Jong, P. F. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Front. Hum. Neurosci.* 8:346. doi: 10.3389/fnhum.2014.00346

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., and Pearson, N. A. (2009). *TOSREC: Test of Sentence Reading Efficiency and Comprehension*. Austin, TX: Pro-Ed.

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.

Copyright © 2020 Psyridou, Tolvanen, Lerkkanen, Poikkeus and Torppa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 22 January 2020
doi: 10.3389/fpsyg.2019.03075

Academic Achievement, Self-Concept, Personality and Emotional Intelligence in Primary Education. Analysis by Gender and Cultural Group

Lucía Herrera[1], Mohamed Al-Lal[2] and Laila Mohamed[1]*

[1] Department of Developmental and Educational Psychology, University of Granada, Melilla, Spain, [2] Early Childhood and Primary Education School "Pedro de Estopiñán", Melilla, Spain

OPEN ACCESS

Edited by:
Juan Luis Castejon,
University of Alicante, Spain

Reviewed by:
Antonio Valle,
University of A Coruña, Spain
Manuel Jesús De La Torre-Cruz,
University of Jaén, Spain
Edgar Galindo,
University of Evora, Portugal
Mercedes Novo,
University of Santiago
de Compostela, Spain

***Correspondence:**
Lucía Herrera
luciaht@ugr.es

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

Received: 12 September 2019
Accepted: 28 December 2019
Published: 22 January 2020

Citation:
Herrera L, Al-Lal M and
Mohamed L (2020) Academic
Achievement, Self-Concept,
Personality and Emotional Intelligence
in Primary Education. Analysis by
Gender and Cultural Group.
Front. Psychol. 10:3075.
doi: 10.3389/fpsyg.2019.03075

A review of the scientific literature shows that many studies have analyzed the relationship between academic achievement and different psychological constructs, such as self-concept, personality, and emotional intelligence. The present work has two main objectives. First, to analyze the academic achievement, as well as the self-concept, personality and emotional intelligence, according to gender and cultural origin of the participants (European vs. Amazigh). Secondly, to identify what dimensions of self-concept, personality and emotional intelligence predict academic achievement. For this, a final sample consisting of 407 students enrolled in the last 2 years of Primary Education were utilized for the study. By gender, 192 were boys (47.2%) and 215 girls (52.8%), with an average age of 10.74 years old. By cultural group, 142 were of European origin (34.9%) and 265 of Amazigh origin (65.1%). The academic achievements were evaluated from the grades obtained in three school subjects: Natural Sciences, Spanish Language and Literature, and Mathematics, and the instruments used for data collection of the psychological constructs analyzed were the Self-Concept Test-Form 5, the Short-Form Big Five Questionnaire for Children, and the BarOn Emotional Quotient Inventory: Youth Version-Short. Based on the objectives set, first, the grades in the subject of Spanish Language and Literature varied depending on the gender of the students. Likewise, differences were found in self-concept, personality, and emotional intelligence according to gender. Also, the physical self-concept varied according to the cultural group. Regarding the second objective, in the predictive analysis for each of the subjects of the curriculum of Primary Education, the academic self-concept showed a greater predictive value. However, so did other dimensions of self-concept, personality and emotional intelligence. The need to carry out a comprehensive education in schools that addresses the promotion of not only academic but also personal and social competences is discussed. Also, that the study of the variables that affect gender differences must be deepened.

Keywords: academic achievement, self-concept, personality, emotional intelligence, gender, cultural group

INTRODUCTION

A review of the scientific literature has shown that many studies have analyzed the relationship between academic achievement and different psychological constructs such as self-concept (Susperreguy et al., 2018; Wolff et al., 2018; Sewasew and Schroeders, 2019), personality (Janošević and Petrović, 2019; Perret et al., 2019; Smith-Woolley et al., 2019), and emotional intelligence (Corcoran et al., 2018; Deighton et al., 2019; Piqueras et al., 2019). In this work, these psychological constructs are analyzed together with primary school children by gender and cultural group. Gender has been a highly studied variable since there are differences between boys and girls in academic performance as well as in the psychological constructs mentioned above (Chrisler and McCreary, 2010; Voyer and Voyer, 2014; Carvalho, 2016; Herrera et al., 2017; Janošević and Petrović, 2019). There are also studies that analyze the possible differences that may exist in the school context between children from different cultures (Schmitt et al., 2007; Strayhorn, 2010; Cvencek et al., 2018; Min et al., 2018). In this sense, there is a disadvantage in the school context for children of minority culture. The present study has been developed in Melilla, a Spanish city located in North Africa, close to Morocco. In their schools, children of European culture and children of Amazigh culture (also known as Berber) have been together from early childhood education. In addition, the predictive value of each of the dimensions that integrate self-concept, personality and emotional intelligence regarding the grades in three subjects of the Primary Education curriculum are analyzed. The psychological constructs analyzed in the present study are described below.

Self-Concept

Many research studies have highlighted that the psychological construction of a positive self-concept by the students, during their academic stage, leads to success in educational environments and social and emotional situations (Eccles, 2009; Harter, 2012; Nasir and Lin, 2012; Chen et al., 2013). Therefore, the positive self-concept acquired in the formative years could help in the development of the strategies and skills needed for confronting life challenges (Huang, 2011). It has also been found that self-concept is positively associated with different factors such as the individual experiencing greater happiness (Hunagund and Hangal, 2014); a greater and better academic performance (Salami and Ogundokun, 2009); greater and more pro-social behaviors (Schwarzer and Fuchs, 2009); and lastly, an overall greater well-being (Mamata and Sharma, 2013).

Among the different models that link self-concept and academic performance, we found the Reciprocal Effects Model (REM), with a theoretical, methodological and empirical review conducted by Marsh and Martin (2011). This model argues that academic self-concept and performance mutually re-enforce themselves, with one producing advances in the other.

Starting with the evolution perspective, the Developmental Equilibrium Hypothesis has also been highlighted. The objective of this hypothesis is centered on achieving equilibrium between two factors that are directly related: self-concept and academic performance (Marsh et al., 2016a,b). Hence, achieving a state of equilibrium has important implications for the development of the individual, but it cannot be ignored that each individual's development of self-concept is different depending on the personal, emotional, and social characteristics surrounding them (Eccles, 2009; Murayama et al., 2013; Paramanik et al., 2014).

The studies that relate self-concept with school or academic performance are exhaustive in the first educational stages as well as higher education (Guay et al., 2010; Möller et al., 2011; Skaalvik and Skjaalvik, 2013). The student's self-concept, and the academic self-concept within it, has a strong influence on student self-efficacy (Ferla et al., 2009). Additionally, academic self-concept significantly correlates with school adjustment in Primary Education (Wosu, 2013; Mensah, 2014) and predicts academic achievement (Marsh and Martin, 2011; Guo et al., 2016). Therefore, in this research it is expected to find such predictive value.

The results from cross-cultural studies have shown that a negative self-concept had detrimental effects on the academic performance of the students from the different samples and countries (Marsh and Hau, 2003; Seaton et al., 2010; Nagengast and Marsh, 2012). Cvencek et al. (2018), when analyzing primary school students from a minority group and a majority group in North America, found that the academic performance, as well as the academic self-concept of the children from the minority group, were lower as compared to those from majority group. Similar results that show the disadvantage of minority groups in schools are found in other studies (Strayhorn, 2010). According to these results, it would be expected that in the present study children of Amazigh cultural origin obtained lower scores than those of European cultural origin in their academic performance and academic self-concept.

Another variable that has been analyzed along with self-concept and academic performance has been gender (Chrisler and McCreary, 2010; DiPrete and Jennings, 2012). Thus, in the meta-analysis study by Voyer and Voyer (2014), it was shown that a certain advantage in school performance existed in women, with their results showing differences in favor of the women for the Language subject. Differences according to gender were also found in self-concept (Nagy et al., 2010). Huang (2013), in a meta-analysis study, identified that the women had a greater self-concept in the subject matter or courses related to language, as well as the arts as compared to the men. Therefore, in this study we expect to find that girls obtain higher grades than boys in Spanish Language and Literature as well as academic self-concept.

Personality

In general terms, personality and self-concept predict satisfaction with life (Parker et al., 2008). Also, personality moderates the effects of the frame of reference that are central for the shaping of self-concept (Jonkmann et al., 2012).

Within the models of personality, the Five Factor Model (McCrae and Costa, 1997) has been the most developed (Herrera et al., 2018), and it represents the dominant conceptualization of the structure of personality in current literature. It postulates that the five great factors of personality (emotional instability, extraversion, intellect/imagination,

agreeableness, and conscientiousness) are found at the highest level in the hierarchy of personality.

Among the strongest arguments utilized to show that the measurements of personality, based on the Big-Five Factor Structure (Goldberg, 1990, 1992), correlate with academic performance, we find the evidence that supports the importance of the personality factors to predict behaviors that are socially valued and the recognition of personality as a component of the individual's will (Chamorro-Premuzic et al., 2006). In this respect, the scientific literature shows studies that relate personality, through the five-factor model, with academic performance. Thus, agreeableness, and intellect/imagination (also known as openness) are related to academic performance (Poropat, 2009; Smith-Woolley et al., 2019). Specifically, conscientiousness predicts academic achievement (O′Connor and Paunonen, 2007), which is expected to be found in the present study.

Personality has been analyzed in different cultures (Allik et al., 2012). A good example of a broad study, which included 56 countries, is the one conducted by Schmitt et al. (2007). Among the main results, it was found that the five-factor structure of personality was robust among the main regions of the world. Also, the inhabitants from South America and East Asia were significantly different in their intellect/imagination characteristics as compared to the rest of the world regions. Thus, while the South American and European countries tended to occupy a higher position in openness, the cultures from East Asia were found in lower positions. This is attributed, among other factors, in that the Asian cultures are more collective, so that the openness dimension could be difficult to clearly identify, as proposed in the starting theoretical model. Based on these results, differences in personality dimensions are expected to be found among children of European and Amazigh cultural origin.

As for gender, differences have also been found. For example, the academic achievement in Primary Education is related to a higher conscientiousness in girls than in boys (Janošević and Petrović, 2019).

Emotional Intelligence

Another factor that should be taken into account, as related to the academic achievements and school adjustment, is the emotional intelligence (EI). The models or theoretical approaches of EI are different (Cherniss, 2010; Herrera et al., 2017). On the one hand, models have been identified that are based on the processing of emotional information, which are focused on basic emotional abilities (Brackett et al., 2011). On the other hand, mixed models of EI have also been identified, which involve both intellectual and personality factors. The socio-emotional competence model by Bar-On (2006) forms part of the second group. In it, different dimensions are identified: intrapersonal, interpersonal, stress management, adaptability, and general mood.

Numerous research studies have examined the relationship between EI and academic performance (Pulido and Herrera, 2017). They have also analyzed their relationship with other variables such as adjustment and permanence in the school context (Hogan et al., 2010; Szczygieł and Mikolajczak, 2017), coping styles (MacCann et al., 2011), the degree of social competence (Franco et al., 2017), and school motivation (Usán and Salavera, 2018).

Emotional intelligence has also been analyzed in groups with different ethnic or cultural origins (Dewi et al., 2017; Min et al., 2018), and according to gender, differences were found in EI as well. Thus, for example, Herrera et al. (2017) obtained results that showed that girls in primary schools in Colombia exceeded the boys in the interpersonal dimension, while the boys stood out in the adaptability dimension. Similarly, Ferrándiz et al. (2012) identified that Spanish girls had higher scores in the interpersonal dimensions and the boys had higher scores in adaptability and general mood. Accordingly, we expect to find differences in emotional intelligence based on the cultural origin and gender of primary school children in this study.

As a function of what has been described until now, the present work has two main objectives. Firstly, to analyze the academic performance, as well as self-concept, personality and emotional intelligence, as a function of gender and cultural origin (European vs. Amazigh) of the participants. It is important to mention that the research study took place in the autonomous city of Melilla, a Spanish city that neighbors Morocco, with unique social, cultural and economic characteristics. In it, people from different cultures co-habit: European, Amazigh (also known as Berber, and who come from the Moroccan Rif), Sephardic and Hindu, although the majority of the population is of European and Amazigh descent and culture. The children with an Amazigh culture origin cohabit live and grow between their maternal culture, which counts with the Tamazight (a dialect that is orally transmitted) as a means of communication, and the European culture, with Spanish being the language employed at school and administrative environments of the city (Herrera et al., 2011).

Secondly, to identify which dimensions of self-concept, personality and emotional intelligence predict academic performance.

In addition, different hypotheses are raised based on the results found in the scientific literature that addresses the research topics described above.

Hypothesis 1. Academic grades differ depending on the gender and cultural origin of students. Thus, for example, as indicated by Voyer and Voyer (2014), girls will achieve higher grades than boys in the subject of Spanish Language and Literature. Likewise, children of cultural origin different from the school (i.e., children of Amazigh culture) will obtain lower grades than Spanish children (Strayhorn, 2010).

Hypothesis 2. The psychology constructs evaluated (self-concept, personality and emotional intelligence) differ according to gender and cultural origin. Among other issues, it is expected to find that girls have a higher academic self-concept than boys (Chrisler and McCreary, 2010), higher scores in the personality dimension of conscientiousness (Janošević and Petrović, 2019) as well as in the interpersonal EI dimension (Ferrándiz et al., 2012; Herrera et al., 2017). Likewise, children of European cultural origin are expected to obtain higher scores than

those of Amazigh cultural origin in academic self-concept (Cvencek et al., 2018), intellect/imagination (Schmitt et al., 2007) and in the intrapersonal and interpersonal EI dimensions (Dewi et al., 2017; Min et al., 2018).

Hypothesis 3. Academic self-concept (Marsh and Martin, 2011; Guo et al., 2016), conscientiousness (O'Connor and Paunonen, 2007) and adaptability (Hogan et al., 2010) predict academic achievement.

MATERIALS AND METHODS

Participants

A non-probabilistic sampling was used. Initially, 422 Primary school students were included in the research study. Nevertheless, once the non-valid cases were eliminated, defined as those who did not complete the evaluation instruments, or whose scores did not comply to what was set, the final sample was comprised of 407 students. These students were enrolled in eight of the twelve public early childhood and primary education centers in the autonomous city of Melilla, Spain (see **Table 1**), with a minimum age of 10 and a maximum of 12 years old. The description of the participants according to cultural origin, gender, grade and age is presented in **Table 2**.

The children of European cultural origin are mainly of Spanish nationality and Catholic religion. They were born in the autonomous city of Melilla or elsewhere in the Iberian Peninsula. Their parents were born in Melilla or have changed their residence to this city for professional reasons (mainly to work in public administration or in the army). Children of Amazigh cultural origin were born in the autonomous city of Melilla, so their nationality is Spanish, or they reside in that city. Many of them are Muslims and have family in Morocco so, given the short distance away, they usually travel at weekends or holidays to Moroccan cities close to Melilla. Rearing practices of children in families of each cultural group developed, among other things, based on cultural values and identities that define them. Thus, for example, the raising of children of Amazigh cultural origin is similar to that of children in the Rif region of Morocco. However, these same children socialize not only with children of their own cultural group but also with children of European cultural origin in a Spanish city, that is, the autonomous city of Melilla. The same can be indicated for children of European cultural origin.

TABLE 1 | Distribution of participants according to the center of early childhood and primary education.

Centers	*N*	%
Center 1	16	3.9
Center 2	19	4.7
Center 3	12	2.9
Center 4	117	28.7
Center 5	18	4.4
Center 6	41	10.1
Center 7	100	24.6
Center 8	84	20.6
Total	407	100.0

TABLE 2 | Distribution of participants according to cultural origin, gender, grade, and age.

Variables	Cultural origin						Age	
	European		Amazigh		Total			
	N	%	*N*	%	*N*	%	*Mean*	*SD*
Gender								
Boy	69	35.9	123	64.1	192	47.2	10.77	0.70
Girl	73	34.0	142	66.0	215	52.8	10.70	0.63
Grade								
Fifth	59	30.4	135	69.6	194	47.7	10.23	0.47
Sixth	83	39.0	130	61.0	213	52.3	11.20	0.45
Total	142	34.9	265	65.1	407	100.0		
Age	*Mean* = 10.78 *SD* = 0.67		*Mean* = 10.71 *SD* = 0.67		*Mean* = 10.74 *SD* = 0.67			

Instruments

Academic Achievement

The final grades of the students of the school subjects Natural Sciences, Spanish Language and Literature, and Mathematics were obtained through a registry, provided by the student's teachers. These were classified as insufficient (0–4.9 points), sufficient (5–5.9 points), good (6–6.9 points), notable (7–8.9 points) and outstanding (9–10 points).

Self-Concept

A Self-Concept Test-Form 5 (AF-5, García and Musitu, 2001) was utilized. It is composed of 30 items that evaluate the self-concept of an individual in academic (e.g., "I do my homework well"), social (e.g., "I make friends easily"), emotional (e.g., "I am afraid of some things"), family (e.g., "I feel that my parents love me") and physical (e.g., "I take good care of my physical health") contexts. This form has to be answered according to an attributive scale ranging from 1 to 99, according to how the item adjusts to what the individual evaluated thinks of it. For example, if a phrase indicates "music helps human well-being" and the student strongly agrees, he/she would answer with a high number, such as 94. But if the student disagreed, he/she would choose a low number, for example 9. Esnaola et al. (2011), when analyzing the psychometric properties of this test in the Spanish population from 12 to 84 years old, indicated that its total reliability was $\alpha = 0.74$. The index of internal consistency, *Cronbach's alpha*, calculated for the present work, had a value of $\alpha = 0.795$. Also, its factorial or construct validity was corroborated in other research works (Elosua and Muñiz, 2010; Malo et al., 2011).

Personality

For the evaluation of personality, the Short-Form Big Five Questionnaire for Children (S-BFQ-C, Beatton and Frijters, 2012) was utilized. It is based on the model of personality structured by five factors (Big-Five Factor Structure), formulated by Goldberg (1990, 1992). These factors are denominated as

emotional instability (e.g., "I am often sad"), extraversion (e.g., "I make friends easily"), intellect/imagination (e.g., "When the teacher explains something, I understand immediately"), agreeableness (e.g., "I share my things with other people") and conscientiousness (e.g., "During class I concentrate on the things I do"), creating the Big Five Questionnaire-Children (BFQ-C). This questionnaire, is directed at children aged between 9 to 15 years old, and was designed and validated by Barbaranelli et al. (2003). In its initial version, its psychometric properties were analyzed with Italian children, although there are studies that have analyzed them in other populations such as for example the German (Muris et al., 2005), Spanish (Carrasco et al., 2005) or Argentinian (Cupani and Ruarte, 2008) populations. Nevertheless, one of the problems of this instrument is its length, given that is composed by 65 items, 13 for each scale. This is the reason why Beatton and Frijters (2012), in a broader study that sought to measure the effects of personality and satisfaction with life on the happiness of Australian youth aged from 9 to 14 years old, reduced the BFQ-C to a shorter version. This shorter version, named S-BFQ-C, is composed by 30 items, so that each of the scales is composed by 6 items. In this version, the questions have to be answered using a *Likert*-type scale with 5 response options (1 = Almost never; 5 = Almost always). The reliability, measured with *Cronbach's Alpha*, was found to be between 0.60 and 0.80 for each of the five scales. For the present study, the total reliability found was $\alpha = 0.783$.

Emotional Intelligence

The BarOn Emotional Quotient Inventory: Youth Version-Short (EQ-i: YV-S, Bar-On and Parker, 2000) was used. It is directed at children aged from 7 to 18 years old, and is composed of 30 items which have to be answered with a *Likert* scale with four possible responses (1 = Very seldom or Not true of me, 4 = Very often or True of me). Six items shape each of the following scales: intrapersonal (e.g., "It is easy to tell people how I feel"), interpersonal (e.g., "I care what happens to other people"), adaptability (e.g., "I can come up with good answers to hard questions"), stress management (e.g., "I can stay calm when I am upset"), and positive impression (e.g., "I like everyone I meet"). This last scale is useful for eliminating the cases of high social desirability. The sum of the first four scales provides the total EQ.

The reliability or internal consistency of the EQ-i YV-S scale oscillates between 0.65 and 0.87 (Bar-On and Parker, 2000). For this study, the total reliability was $\alpha = 0.745$. Its internal structure was confirmed in Spanish (Esnaola et al., 2016), Hungarian (Kun et al., 2012), Mexican (Esnaola et al., 2018b), English (Davis and Wigelsworth, 2018) and Chinese (Esnaola et al., 2018a) populations.

Procedure

Information Collection

In the first place, the participation of the management teams of the 12 early childhood and primary school education centers in Melilla was solicited. Of these, eight centers answered affirmatively. Afterward, within each center, the professor-tutor from each class or classes interested were contacted. A group meeting was conducted with the parents from each group-class, where information was provided about the objectives of the research study. The authorization of the children's parents for the exclusive use of the results obtained, for educational and scientific purposes, was requested.

Once this process was finished, a document was provided to the teachers-tutors of each participating class which explained how to access the web program utilized for the management of the student's grades in order to download this information in pdf format. Once this information was downloaded, they were asked to write down, in a double-entry table provided for each student, the final grades obtained in the subjects of Natural Sciences, Spanish Language and Literature, and Mathematics, using the scoring system of insufficient, sufficient, good, notable or outstanding. Teachers provided students' grades to researchers at the end of the academic year.

The AF-5, the S-BFQ-C and the EQ-i: YV-S questionnaires were administered in the first school term to the students in fifth and sixth grade of Primary Education, collectively according to group-class. The maximum time provided for this was 55 min. Previously, the students were told that there were no right or wrong answers, and that they should answer with total sincerity, given that the test was anonymous. Also, that they should not write their name; and that what they were about to answer did not have any relation with the school grades; and lastly, that they should read the questions, and if they had any doubts (for example, not understanding a term), they should raise their hand so that the question could be resolved.

In order to be able to relate the results of the evaluation of the different psychological constructs and the academic grades, the teacher of each class assigned a number to each student. This number was recorded both in the grades provided by him/her and on the first page of each of the questionnaires administered.

Statistical Analysis of the Data

Before proceeding with the statistical analysis, from the 422 students who participated, it was determined if there were students who had not completed the three evaluation tests, and also if they obtained high scores in the positive impression scale of the EQ-i: YV-S. This resulted in the elimination of 15 individuals, resulting in a final sample of 407 students.

The statistical program *IBM SPSS Statistics 23* was used to carry out the statistical analysis. Descriptive statistics were utilized to describe the data (frequencies, percentages, mean and standard deviation). In other words, to answer the first research objective and the first two hypotheses, two Analysis of variance (ANOVA) were performed in which the Academic achievement was used as the dependent variable in one case, and self-concept, personality and EI as dependent variables in the other. In both cases, the independent variables were gender (boy or girl) and cultural group (European vs. Amazigh). The effect size was calculated with the partial eta-squared as the *post hoc* test, through the use of the *Bonferroni* test.

To address the second objective and the third hypothesis, three multiple linear regression analysis (with the enter method) were conducted, in which each subject was introduced as the dependent variable, with the predictive variables being the different dimensions which comprised the self-concept,

personality and EI constructs. To justify the method used, the non-autocorrelation of the data was determined, using the Durbin Watson test, and the non-existence of multicollinearity, through the Variance Inflation Factor.

RESULTS

Academic Achievement by Gender and Cultural Group

All the subjects had a maximum of five points, and were scored as: 1 = Insufficient, 2 = Sufficient, 3 = Good, 4 = Notable, 5 = Outstanding. The mean grade in Natural Sciences was 3.26 (SD = 1.33), for Spanish Language and Literature it was 3.33 (SD = 1.24) and in Mathematics, it was 3.19 (SD = 1.25).

Academic achievement as a function of the student's gender and cultural group is presented in **Table 3**. The analysis of variance performed as a function of gender and cultural group showed that there were differences according to gender for the subject Spanish Language and Literature, F = 5.812, p = 0.016, *Eta2p* = 0.014, so that the girls obtained higher grades than the boys, t = 0.313, p = 0.016. No differences were found neither in Nature Sciences, F = 0.763, p = 0.383, *Eta2p* = 0.002, nor Mathematics, F = 1.692, p = 0.194, *Eta2p* = 0.004. On their part, no differences were found as a function of the cultural group, $F_{\text{Natural Sciences}}$ = 0.376, p = 0.540, *Eta2p* = 0.001; $F_{\text{Language and Literature}}$ = 0.565, p = 0.453, *Eta2p* = 0.001; $F_{\text{Mathematics}}$ = 0.576, p = 0.448, *Eta2p* = 0.001.

TABLE 3 | Academic achievement by gender and cultural group.

Subjects	Gender	Cultural group	*Mean*	*SD*	*N*
Natural sciences	Boy	Amazigh	3.10	1.41	123
		European	3.31	1.11	69
		Total	3.17	1.31	192
	Girl	Amazigh	3.35	1.26	142
		European	3.31	1.50	73
		Total	3.33	1.34	215
	Total	Amazigh	3.23	1.33	265
		European	3.31	1.32	142
		Total	3.26	1.33	407
Spanish language and literature	Boy	Amazigh	3.13	1.29	123
		European	3.24	1.24	69
		Total	3.17	1.27	192
	Girl	Amazigh	3.45	1.16	142
		European	3.54	1.27	73
		Total	3.48	1.20	215
	Total	Amazigh	3.30	1.23	265
		European	3.39	1.26	142
		Total	3.33	1.24	407
Mathematics	Boy	Amazigh	3.09	1.27	123
		European	3.15	1.20	69
		Total	3.11	1.24	192
	Girl	Amazigh	3.22	1.22	142
		European	3.36	1.33	73
		Total	3.27	1.25	215
	Total	Amazigh	3.16	1.24	265
		European	3.26	1.27	142
		Total	3.19	1.25	407

Self-Concept, Personality and EI by Gender and Cultural Group

The analysis of variance results (see **Supplementary Table S1**) showed that there were significant differences as a function of gender for self-concept, more specifically in academic self-concept, with the girls achieving higher grades in *post hoc* comparisons using the *Bonferroni* test, t = 0.667, p = 0.007, and self-esteem, t = 1.139, p < 0.001, where the boys stood out. Likewise, differences were found in personality in favor of the girls within the conscientiousness, t = 1.136, p = 0.018, and agreeableness dimensions, t = 1.641, p = 0.001. Also, with respect to the EI, the girls had a higher score in the interpersonal scale, t = 1.016, p = 0.007, while the boys had a higher score in the stress management, t = 1.513, p < 0.001, and adaptability, t = 1.110, p = 0.008. Lastly, with respect to the analysis according to cultural group, there were only significant differences in the physical self-concept, with higher scores reached by the children of Amazigh cultural origin, t = 0.420, p = 0.036.

Predictive Value of the Different Dimensions Evaluated With Respect to Academic Achievement

In first place, a linear regression analysis was conducted, where the dependent variable was the subject Natural Sciences and the predictive variables were the five dimensions of the self-concept, the five dimensions from personality and the four dimensions from EI (see **Table 4**). The model was significant with values F = 11.003, p < 0.001. Likewise, the coefficient of determination was R^2 = 0.311 (adjusted R^2 = 0.282). Durbin–Watson's d test showed that there was no auto-correlation in the data (d = 1.583). Values of the Durbin Watson test between 1.5 and 2.5 indicate that the data are not correlated (Durbin and Watson, 1951). Also, the Variance Inflation Factor (VIF) obtained values lower than 5, so multicollinearity was not present (Berry and Feldman, 1985; Belsley, 1991).

In the order from greater to lesser predictive value, the dimensions were: academic self-concept, physical self-concept, intrapersonal, intellect/imagination, and family self-concept. The physical self-concept, as well as intrapersonal intelligence, negatively predicted the grades in Natural Sciences.

In second place, as related to the subject Spanish Language and Literature (see **Table 5**), the model was significant with values of F = 10.442, p < 0.001 and with a coefficient of determination of R^2 = 0.299, adjusted R^2 = 0.271. The data was not correlated (d = 1.672) and no multicollinearity was found.

Once again, the academic self-concept dimension had the greatest predictive value, followed by the physical self-concept, intrapersonal intelligence, and intellect/imagination dimensions. The negative predictions remained the same.

In third and last place, for the subject of Mathematics (see **Table 6**), the model had a statistical significance, as shown by

TABLE 4 | Regression analysis of the different dimensions analyzed with respect to the natural sciences subject.

Dimensions	Non-standardized coefficients		Standardized coefficients	*t*	*p*	Collinearity tests	
	B	*Standard error*	*Beta*			TI	VIF
Academic self-concept	0.297	0.037	0.490	7.940***	0.001	0.530	1.888
Social self-concept	0.018	0.052	0.020	0.344	0.731	0.592	1.689
Self-esteem	0.032	0.030	0.054	1.053	0.293	0.771	1.297
Family self-concept	0.110	0.055	0.108	1.993*	0.047	0.681	1.469
Physical self-concept	−0.235	0.040	−0.318	−5.810***	0.001	0.672	1.488
Conscientiousness	6.452E−5	0.020	0.000	0.003	0.997	0.505	1.979
Agreeableness	−0.028	0.019	−0.092	−1.450	0.148	0.496	2.017
Emotional instability	−0.004	0.016	−0.014	−0.252	0.801	0.691	1.447
Intellect/Imagination	0.037	0.018	0.120	2.025*	0.044	0.577	1.733
Extraversion	0.003	0.020	0.010	0.169	0.866	0.592	1.690
Intrapersonal	0.040	0.019	0.107	2.165*	0.031	0.822	1.217
Interpersonal	0.003	0.023	0.007	0.120	0.904	0.590	1.695
Stress management	0.010	0.020	0.029	0.497	0.619	0.580	1.723
Adaptability	0.029	0.021	0.082	1.377	0.170	0.562	1.781

****p < 0.001, *p < 0.05. TI, tolerance index; VIF, variance inflation factor.*

TABLE 5 | Regression analysis of the different dimensions analyzed with respect to the Spanish language and literature subject.

Dimensions	Non-standardized coefficients		Standardized coefficients	*t*	*p*	Collinearity tests	
	B	*Standard error*	*Beta*			TI	VIF
Academic self-concept	0.266	0.035	0.470	7.556***	0.001	0.530	1.888
Social self-concept	0.024	0.049	0.029	0.496	0.621	0.592	1.689
Self-esteem	0.014	0.028	0.025	0.481	0.631	0.771	1.297
Family self-concept	0.089	0.052	0.094	1.711	0.088	0.681	1.469
Physical self-concept	−0.170	0.038	−0.247	−4.471***	0.001	0.672	1.488
Conscientiousness	0.009	0.018	0.031	0.480	0.632	0.505	1.979
Agreeableness	−0.021	0.018	−0.075	−1.164	0.245	0.496	2.017
Emotional instability	−0.006	0.015	−0.020	−0.370	0.711	0.691	1.447
Intellect/Imagination	0.036	0.017	0.123	2.060*	0.040	0.577	1.733
Extraversion	0.021	0.018	0.067	1.144	0.253	0.592	1.690
Intrapersonal	0.037	0.017	0.107	2.138*	0.033	0.822	1.217
Interpersonal	0.000	0.022	0.001	0.016	0.987	0.590	1.695
Stress management	0.000	0.019	0.001	0.010	0.992	0.580	1.723
Adaptability	0.009	0.020	0.026	0.435	0.664	0.562	1.781

****p < 0.001, *p < 0.05. TI, tolerance index; VIF, variance inflation factor.*

$F = 10.790$, $p < 0.001$. The coefficient of determination obtained was $R^2 = 0.306$, adjusted $R^2 = 0.278$. The data was not correlated ($d = 1.600$) and multicollinearity was not present.

The predictive dimensions were academic self-concept, physical self-concept (in a negative manner), adaptability, intellect/imagination, and conscientiousness.

DISCUSSION

Based on the hypotheses set, first, the grades of the Spanish Language and Literature school subject varied depending on the gender of the students, which coincided with the results from other studies, which highlighted the girls' higher grades (Huang, 2013; Voyer and Voyer, 2014). In this regard, it could be argued that academic and social expectations are different depending on gender (Voyer and Voyer, 2014). Likewise, the influence of socialization on the formation of gender behaviors must be taken into account in accordance with the cultural norms of masculinity and femininity (Gibb et al., 2008). Gender differences in academic achievement remain between different countries, regardless of their political, economic or social equality (Stoet and Geary, 2015). However, it is noteworthy that in adulthood women occupy fewer representations of political, economic and academic leadership than men.

Contrary to expectations (Strayhorn, 2010; Whaley and Noël, 2012), children of Amazigh origin did not obtain lower grades than those of European origin. These results may be due to the fact that in the city of Melilla children of both cultures are educated from early childhood education in schools where

TABLE 6 | Regression analysis of the different dimensions analyzed with respect to the mathematics subject.

Dimensions	Non-standardized coefficients		Standardized coefficients	*t*	*p*	Collinearity tests	
	B	*Standard error*	*Beta*			TI	VIF
Academic self-concept	0.273	0.035	0.480	7.759***	0.001	0.530	1.888
Social self-concept	−0.004	0.049	−0.005	−0.078	0.938	0.592	1.689
Self-esteem	−0.021	0.028	−0.038	−0.744	0.458	0.771	1.297
Family self-concept	0.069	0.052	0.073	1.330	0.185	0.681	1.469
Physical self-concept	−0.189	0.038	−0.273	−4.968***	0.001	0.672	1.488
Conscientiousness	0.039	0.018	0.133	2.105*	0.036	0.505	1.979
Agreeableness	0.013	0.018	0.046	0.714	0.476	0.496	2.017
Emotional instability	−0.018	0.015	−0.067	−1.240	0.216	0.691	1.447
Intellect/Imagination	0.045	0.017	0.153	2.575*	0.010	0.577	1.733
Extraversion	0.012	0.018	0.038	0.647	0.518	0.592	1.690
Intrapersonal	0.009	0.017	0.026	0.529	0.597	0.822	1.217
Interpersonal	0.018	0.022	0.049	0.839	0.402	0.590	1.695
Stress management	0.002	0.019	0.006	0.104	0.917	0.580	1.723
Adaptability	0.054	0.020	0.162	2.693**	0.007	0.562	1.781

****p < 0.001, **p < 0.01, *p < 0.05. TI, tolerance index; VIF, variance inflation factor.*

the language used is Spanish. Thus, the academic performance at the end of Primary Education does not differ depending on the cultural origin of the students. However, it is necessary to show that early childhood teachers dedicate great efforts so that children of Amazigh cultural origin develop the linguistic skills necessary for the correct learning and use of the Spanish language (Herrera et al., 2011). Therefore, hypothesis 1 is partially confirmed. That is, the results found indicate that academic achievement varies according to gender but not the cultural origin of the students.

Likewise, differences were found according to gender in self-concept, specifically in the academic self-concept and self-esteem; for personality, within the factors of conscientiousness and agreeableness; in addition to emotional intelligence, particularly in the interpersonal, stress management and adaptability scales. As for the differences found for self-concept according to gender (Nagy et al., 2010), the results found for academic self-concept showed differences in favor of the girls (Malo et al., 2011). Nevertheless, other factors should be taken into account, such as the academic responsibilities associated to school success and failure, given that, for example, the boys in Compulsory Secondary Education attribute their academic success to their skills, while the girls attribute them to their effort (Inglés et al., 2012). As for emotional self-concept or self-esteem, the boys exceeded the girls (Xie et al., 2019). Cross-cultural studies show that differences in self-esteem according to gender are maintained in different countries, although their magnitude differ according to the cultural differences found in the socioeconomic, sociodemographic, gender equality and cultural value indicators (Bleidorn et al., 2016). In this respect, the emotion literacy programs, based on the development of emotional intelligence, could be a useful tool for the development of self-esteem (Cheung et al., 2014).

As for the differences in the personality dimensions conscientiousness and agreeableness in favor of the girls, the results were in agreement with previous studies (Rahafar et al., 2017; Janošević and Petrović, 2019). Within the differences in EI according to gender, the girls scored higher in the interpersonal scale, while the boys did so in stress management and adaptability (Ferrándiz et al., 2012; Herrera et al., 2017). In this way, the girls showed competencies and skills that were higher than the boys in empathy, social responsibility, and interpersonal relationships. On the contrary, the boys stood out in stress tolerance and impulse control (stress management), as well as in reality-testing, flexibility, and problem-solving (adaptability). These differences, as a function of gender, could be due to cultural factors and family rearing practices differentiated as a function of gender (Joseph and Newman, 2010).

Also, the physical self-concept varied according to the cultural origin, where children from the Amazigh culture obtained higher scores than children of European culture origin. This may be due to the influence of cultural values (their own, meaning Amazigh, as well as the context in which they live in, given that the children are socialized in a European context), with respect to body image and physical self-concept (Marsh et al., 2007).

Based on the results found, the second hypothesis is partially confirmed. The three psychological constructs evaluated differ according to gender in the expected direction but only in the self-concept are differences found according to the cultural origin. Although it was expected to find differences in favor of children of European cultural origin in academic self-concept (Cvencek et al., 2018), they have been found in physical self-concept in favor of children of Amazigh cultural origin. As previously indicated, children of European and Amazigh culture develop in the same school contexts from the early educational stages. Thus, educational policies developed in schools may have contributed to eliminating the possible socio-cultural disadvantages of children of Amazigh cultural origin. This implies, therefore, that there are no differences depending on the cultural group in the academic self-concept.

In the predictive analysis developed for each of the school subjects of the curriculum of Primary Education, with the aim of answering the second objective and the third hypothesis of the study, the academic self-concept showed a greater predictive value (Marsh and Martin, 2011; Jansen et al., 2015; Guo et al., 2016; Lösch et al., 2017; Susperreguy et al., 2018). This result confirms the third hypothesis. That is, the relevance of academic self-concept in school performance. However, so did other dimensions of self-concept. More specifically, the physical self-concept negatively predicted the academic results in the three subjects evaluated (Lohbeck et al., 2016). Children who participated in the study are in the process of transition from childhood to adolescence. Biological changes in their bodies due to this stage of evolutionary development as well as greater attention to appearance and physical abilities may interfere at the end of Primary Education in their academic performance. Furthermore, the family self-concept predicted the grades of the Natural Sciences school subject. This last result points to the influence of the family on self-concept as well as academic results (Corrás et al., 2017; Mortimer et al., 2017; Häfner et al., 2018).

Personality also predicted the academic results in the three school subjects from the Primary Education curriculum analyzed (O′Connor and Paunonen, 2007; Spengler et al., 2016; Bergold and Steinmayr, 2018), i.e., the intellect/imagination dimension for the three subjects and conscientiousness for Mathematics. In the first case, it may be because intellect/imagination or openness is a personality dimension that reflects cognitive exploration (DeYoung, 2015). It refers to the ability and tendency to find, understand and use complex patterns of both sensory and abstract information. Therefore, those children who score higher in intellect/imagination will get better academic results than those with lower scores. In the second case, conscientiousness relates to responsibility, persistence, trustworthiness, and being purposeful (Conrad and Patry, 2012). Children with high conscientiousness can develop a variety of effective learning strategies, which may be associated with higher academic performance in Mathematics.

Likewise, EI predicted academic achievement in every case (Salami and Ogundokun, 2009; Hogan et al., 2010; Brackett et al., 2011; MacCann et al., 2011). More specifically, the intrapersonal scale predicted it for the subjects of Natural Sciences and Spanish Language and Literature. Intrapersonal intelligence involves the knowledge and labeling of one's own feelings. This ability may contribute to achieving better grades in both subjects of the curriculum. For example, in the subject of Spanish Language and Literature it can facilitate the communicative skills related to the reading of different kinds of texts, their reflection and their understanding. On the other hand, in the subject of Nature Sciences it can contribute to interpret reality in order to address the solution to the different problems that arise, as well as to explain and predict natural phenomena and to face the need to develop critical attitudes before the consequences that result from scientific advances. In the case of the Mathematics subject, the adaptability scale predicted the academic achievement. Adaptability implies abilities such as being able to adjust one's emotions and behaviors to changing situations or conditions, which is closely related to mathematical thinking.

In general, scientific literature shows that academic achievement is related to self-concept (Susperreguy et al., 2018; Wolff et al., 2018; Sewasew and Schroeders, 2019), personality (Perret et al., 2019; Smith-Woolley et al., 2019), and EI (Corcoran et al., 2018; Deighton et al., 2019; Piqueras et al., 2019). Also, that within these construct, academic self-concept (Ferla et al., 2009; Guay et al., 2010; Chen et al., 2013; Marsh et al., 2014), intellect/imagination (Poropat, 2009; Smith-Woolley et al., 2019), and adaptability (MacCann et al., 2011; Szczygieł and Mikolajczak, 2017) correlate significantly with academic achievement. In this research the predictive value of the dimensions of self-concept, personality and EI regarding the academic grades obtained in three subjects of the Primary Education curriculum has been established. One of its strengths is that it analyzes the predictive value of these psychological constructs together, not separately as in other studies.

In addition, the study has been developed in a multicultural context where children of European and Amazigh cultural origin coexist. Children of Amazigh cultural origin usually have access to early childhood education centers with a lower knowledge of the Spanish language than children of European cultural origin (Herrera et al., 2011). Although studies carried out with groups of cultural minorities show differences in their school performance (Strayhorn, 2010; Whaley and Noël, 2012), in the present study they are not at the end of Primary Education. This fact may be due to the linguistic policy developed in Melilla educational centers, which means that the mother language of children of Amazigh origin does not represent a disadvantage for academic achievement.

Further, gender differences found in the study seem to be more relevant than cultural differences. In fact, they are only in the physical self-concept in the latter case. Personality can mediate in adapting to school demands, so that girls are more conscientiousness than boys and follow norms in a more adaptive way (Carvalho, 2016). Moreover, since girls excel in their academic self-concept, their self-efficacy may also be superior to that of boys, which contributes to a better school adjustment (Ferla et al., 2009). Girls also have greater interpersonal intelligence, indicating better empathy, social responsibility and interpersonal relationships (Ferrándiz et al., 2012). Such non-cognitive abilities can stimulate the development of positive interpersonal relationships in the classroom with both the teachers and their peers. These individual differences may be due to family and social influences where, for example, girls are expected to be more emotionally expressive than boys (Meshkat and Nejati, 2017). In this same direction it could explain why children have greater self-esteem and stress management that girls.

Practical Implications for Education

In light of the results obtained in the present research study, the need to carry out a comprehensive education in schools that addresses the promotion of not only academic but also personal, social and emotional competences, are underlined (Cherniss, 2010; Hunagund and Hangal, 2014; Herrera et al., 2017; Szczygieł and Mikolajczak, 2017; Corcoran et al., 2018; Cvencek et al., 2018). For this, the application of the principles

derived from Positive Psychology in the education field would be an adequate strategy (Suldo et al., 2015; Chodkiewicz and Boyle, 2017; Domitrovich et al., 2017; Shoshani and Slone, 2017). Thus, intellectual, procedural and emotional aspects have to be worked on in learning, the latter being clear drivers of learning. The pleasant emotions experienced by children in educational settings will allow greater happiness and emotional well-being in them (Gil and Martínez, 2016). For it, teachers must be trained in good teaching practices that allow the interest of students to learn as well as guide them in the emotional domain (Castillo et al., 2013; Oberle et al., 2016; Conners-Burrow et al., 2017).

Likewise, schools must respond to the gender and cultural differences of students (Chrisler and McCreary, 2010; DiPrete and Jennings, 2012), particularly the first based on the results of this study. Thus, for example, the development of greater self-esteem in girls (Bleidorn et al., 2016; Xie et al., 2019) should be encouraged. As indicated by Cheung et al. (2014), emotional literacy programs that are based on emotional intelligence are an appropriate strategy for promoting self-esteem. Similarly, gender differences must be taken into account in response to other factors such as cultural group, family beliefs and parenting practices (Chrisler and McCreary, 2010; Joseph and Newman, 2010; Nagy et al., 2010; Allik et al., 2012; Marsh et al., 2015).

Study Limitations and Proposal for Future Research

The present study has been developed taking into account only the last two school years of the education stage of Primary Education, just before the transition to Compulsory Secondary Education. Given that the scientific literature shows evolutionary changes in the development of the constructs analyzed (Huang, 2011; Murayama et al., 2013; Marsh et al., 2015; Bleidorn et al., 2016), longitudinal studies could be conducted in future research studies from Primary Education to Compulsory Secondary Education in order to determine the magnitude and direction of these changes.

On the other hand, all the instruments for data collection used to evaluate the psychological constructs analyzed in the present study are based on self-report measures. Different types of measuring instruments (self-report measures and performance measures) should be combined in future studies (Petrides et al., 2010; Mayer et al., 2012).

Gender differences in academic achievement as well as the psychological constructs analyzed have been revealed. However, it has to deepen the analysis of personal variables, family, social and cultural factors that contribute to that, even though women get better scores on their school performance across the different educational stages, at adulthood that reach fewer representations than men in leadership positions (Stoet and Geary, 2015).

Finally, given the cultural diversity in schools it is necessary to develop studies that analyze academic achievement as well as its relationship with different psychological variables in students of different cultural groups. Cross-cultural studies comparing different countries are necessary (Marsh and Hau, 2003; Nagengast and Marsh, 2012; Bleidorn et al., 2016; Min et al., 2018) but teachers have to know how to deal with coexistence and cultural diversity within the classrooms.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Commission, Faculty of Educational Sciences and Sports, University of Granada, Melilla, Spain. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

LH, MA-L, and LM shared conception, design, and the final version of the work, were jointly accountable for the content of the work, ensured that all aspects related to accuracy or integrity of the study were investigated and resolved in an appropriate way, and shared the internal consistency of the manuscript. MA-L and LM contributions were mainly in the theoretical part and in revising it critically. LH contribution was mainly in methodological question and data analysis.

FUNDING

This research was co-financed by the Research Group Development, Education, Diversity, and Culture: Interdisciplinary Analysis (HUM-742).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.03075/full#supplementary-material

REFERENCES

Allik, J., Massoudi, K., Realo, A., and Rossier, J. (2012). Personality and culture. cross-cultural psychology at the next crossroads. *Swiss J. Psychol.* 71, 5–12. doi: 10.1024/1421-0185/a000069

Barbaranelli, C., Caprara, G. V., Rabasca, A., and Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Pers. Individ. Differ.* 34, 645–654. doi: 10.1016/S0191-8869(02)00051-X

Bar-On, R. (2006). The bar-on model of emotional-social intelligence (ESI). *Psicothema* 18, 13–25.

Bar-On, R., and Parker, J. (2000). *BarOn Emotional Quotient Inventory: Youth Version (BarOn EQi:YV), Technical Manual*. Toronto: Multi-Health Systems Inc.

Beatton, T., and Frijters, P. (2012). Unhappy Young Australians: a domain approach to explain life satisfaction change in children. *QUT School of Economics and Finance (Working/Discussion Paper # 289 August)*, Indooroopilly.

Belsley, D. A. (1991). *Conditioning Diagnostics: Collinearity and Weak Data in Regression*, 1st Edn. New York, NY: John Wiley & Sons.

Bergold, S., and Steinmayr, R. (2018). Personality and intelligence interact in the prediction of academic achievement. *J. Intell.* 6:27. doi: 10.3390/jintelligence6020027

Berry, W. D., and Feldman, S. (1985). *Multiple Regression in Practice*. Thousand Oaks, CA: SAGE Publications.

Bleidorn, W., Arslan, R. C., Denissen, J. J. A., Rentfrow, P. J., Gebauer, J. E., Potter, J., et al. (2016). Age and gender differences in self-esteem—A cross-cultural window. *J. Pers. Soc. Psychol.* 111, 396–410. doi: 10.1037/pspp0000078

Brackett, M., Rivers, S. E., and Salovey, P. (2011). Emotional intelligence: implications for personal, social, academic, and workplace success. *Soc. Personal. Psychol. Compass* 5, 88–103. doi: 10.1111/j.1751-9004.2010.00334.x

Carrasco, M. A., Holgado, F. P., and Del Barrio, M. V. (2005). Dimensionalidad del cuestionario de los cinco grandes (BFQ-N) en población infantil española [Big five questionnaire dimensions in Spanish children (BFQ-C)]. *Psicothema* 17, 286–291.

Carvalho, R. G. (2016). Gender differences in academic achievement: the mediating role of personality. *Pers. Individ. Differ.* 94, 54–58. doi: 10.1016/j.paid.2016.01.011

Castillo, R., Fernández-Berrocal, P., and Brackett, M. A. (2013). Enhancing teacher effectiveness in spain: a pilot study of the RULER approach to social and emotional learning. *J. Educ. Train. Stud.* 1, 263–272. doi: 10.11114/jets.v1i2.203

Chamorro-Premuzic, T., Furnham, A., and Ackerman, P. L. (2006). Incremental validity of the typical intellectual engagement scale as predictor of different academic performance measures. *J. Pers. Assess.* 87, 261–268. doi: 10.1207/s15327752jpa8703_07

Chen, S., Yeh, Y., Hwang, F., and Lin, S. S. J. (2013). The relationship between academic self-concept and achievement: a multicohort- multioccasion study. *Learn. Individ. Differ.* 23, 172–178. doi: 10.1016/j.lindif.2012.07.021

Cherniss, C. (2010). Emotional intelligence: toward clarification of a concept. *Ind. Organ. Psychol.* 3, 110–126. doi: 10.1111/j.1754-9434.2010.01231.x

Cheung, C. K., Cheung, H. Y., and Hue, M. T. (2014). Emotional intelligence as a basis for self-esteem in young adults. *J. Psychol.* 149, 63–84. doi: 10.1080/00223980.2013.838540

Chodkiewicz, A. R., and Boyle, C. (2017). Positive psychology school-based interventions: a reflection on current success and future directions. *Rev. Educ.* 5, 60–86. doi: 10.1002/rev3.3080

Chrisler, J. C., and McCreary, D. R. (2010). *Handbook of Gender Research in Psychology. Volume 1: Gender Research in General and Experimental Psychology*. New York, NY: Springer.

Conners-Burrow, N. A., Patrick, T., Kyzer, A., and McKelvey, L. (2017). A preliminary evaluation of REACH: training early childhood teachers to support children's social and emotional development. *Early Child. Educ. J.* 45, 187–199. doi: 10.1007/s10643-016-0781-2

Conrad, N., and Patry, M. W. (2012). Conscientiousness and academic performance: a mediational analysis. *Int. J. Scholarsh. Teach. Learn.* 6:8. doi: 10.20429/ijsotl.2012.060108

Corcoran, R. P., Cheung, A. C. K., Kim, E., and Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: a systematic review and meta-analysis of 50 years of research. *Educ. Res. Rev.* 25, 56–72. doi: 10.1016/j.edurev.2017.12.001

Corrás, T., Seijo, D., Fariña, F., Novo, M., Arce, R., and Cabanach, R. G. (2017). What and how much do children lose in academic settings owing to parental separation? *Front. Psychol.* 8:1545. doi: 10.3389/fpsyg.2017.01545

Cupani, M., and Ruarte, M. (2008). Propiedades psicométricas del Cuestionario de los Cinco Factores para Niños (BFQ-C) en una muestra de adolescentes argentinos [Psychometric properties of the Big Five Questionnaire for Children (BFQ-C) in an Argentine sample of young adolescents]. *Estud. Psicol.* 29, 351–364. doi: 10.1174/021093908786145421

Cvencek, D., Fryberg, S. A., Covarrubias, R., and Meltzoff, A. N. (2018). Self-concepts, self-esteem, and academic achievement of minority and majority north american elementary school children. *Child Dev.* 89, 1099–1109. doi: 10.1111/cdev.12802

Davis, S. K., and Wigelsworth, M. (2018). Structural and predictive properties of the emotional quotient inventory youth version–short form (EQ-i:YV[S]). *J. Pers. Assess.* 100, 197–206. doi: 10.1080/00223891.2017.1280502

Deighton, K., Hudson, J., Manley, A. J., Kaiseler, M., Patterson, L. B., Rutherford, Z. H., et al. (2019). Effects of emotional intelligence and supportive text messages on academic outcomes in first-year undergraduates. *J. Furth. High. Educ.* 43, 494–507. doi: 10.1080/0309877X.2017.1377161

Dewi, Z. L., Halim, M. S., and Derksen, J. (2017). Emotional intelligence competences of three different ethnic groups in Indonesia. *Asian Ethn.* 19, 36–58. doi: 10.1080/14631369.2017.1310615

DeYoung, C. G. (2015). "Openness/intellect: a dimension of personality reflecting cognitive exploration," in *APA Handbook of Personality and Social Psychology: Personality Processes and Individual Differences*, Vol. 4, eds M. L. Cooper, and R. J. Larsen (Washington, DC: American Psychological Association), 369–399. doi: 10.1037/14343-017

DiPrete, T. A., and Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Soc. Sci. Res.* 41, 1–15. doi: 10.1016/j.ssresearch.2011.09.001

Domitrovich, C. E., Durlak, J. A., Staley, K. C., and Weissberg, R. P. (2017). Social-emotional competence: an essential factor for promoting positive adjustment and reducing risk in school children. *Child Dev.* 88, 408–416. doi: 10.1111/cdev.12739

Durbin, J., and Watson, G. S. (1951). Testing for serial correlation in least squares regression, II. *Biometrika* 38, 159–178. doi: 10.1093/biomet/38.1-2.159

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

Elosua, P., and Muñiz, J. (2010). Exploring the factorial structure of the self-concept. a sequential approach using CFA, MIMIC, and MACS models, across gender and two languages. *Eur. Psychol.* 15, 58–67. doi: 10.1027/1016-9040/a000006

Esnaola, I., Arias, V. B., Freeman, J., Wang, Y., and Arias, B. (2018a). Validity evidence based on internal structure of scores of the emotional quotient inventory: youth version short (EQ-i: YV-S) in a Chinese sample. *J. Psychoeduc. Assess.* 36, 576–587. doi: 10.1177/0734282916689439

Esnaola, I., Azpiazu, L., Antonio-Aguirre, I., Sarasa, M., and Ballina, E. (2018b). Validity evidence of emotional quotient inventory: youth version (Short) in a sample of Mexican adolescents. *Estud. Psicol.* 39, 127–153. doi: 10.1080/02109395.2017.1407905

Esnaola, I., Freeman, J., Sarasa, M., Fernández-Zabala, A., and Axpe, I. (2016). Validity evidence based on internal structure of scores of the emotional quotient-inventory: youth version short (EQ-i: YV-S) in a Spanish sample. *Span. J. Psychol.* 19:E12. doi: 10.1017/sjp.2016.12

Esnaola, I., Rodríguez, A., and Goñi, E. (2011). Propiedades psicométricas del cuestionario de Autoconcepto AF5 [Psychometric properties of AF5 self-concept questionnaire]. *An. Psicol.* 27, 109–117.

Ferla, J., Valcke, M., and Cai, Y. (2009). Academic self-efficacy and academic self-concept: reconsidering structural relationships. *Learn. Individ. Differ.* 19, 499–505. doi: 10.1016/j.lindif.2009.05.004

Ferrándiz, C., Hernández, D., Bermejo, R., Ferrando, M., and Sáinz, M. (2012). Social and emotional intelligence in childhood and adolescence: Spanish validation of a measurement instrument. *Rev. Psicodidact.* 17, 309–338. doi: 10.1387/Rev.Psicodidact.2814

Franco, M. D. G., Beja, M. J., Candeias, A., and Santos, N. (2017). Emotion understanding, social competence and school achievement in children from primary school in Portugal. *Front. Psychol.* 8:1376. doi: 10.3389/fpsyg.2017.01376

García, F., and Musitu, G. (2001). *Autoconcepto Forma 5. AF5 [Self-Concept Test-Form 5. AF-5]*. Madrid: TEA.

Gibb, S. J., Fergusson, D. M., and Horwood, L. J. (2008). Gender differences in educational achievement to age 25. *Aust. J. Educ.* 52, 63–80. doi: 10.1177/000494410805200105

Gil, P., and Martínez, M. (2016). Emociones percibidas, por alumnos y maestros, en Educación Física en 6° curso de Primaria [Perceived emotions, students and teachers, in Physical Education in the 6th grade of Primary]. *Educación XX1* 19, 179–204. doi: 10.5944/educXX1.1423

Goldberg, L. R. (1990). An alternative "description of personality": the Big-Five factor structure. *J. Pers. Soc. Psychol.* 59, 1215–1229. doi: 10.1037/0022-3514.59.6.1216

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychol. Assess.* 4, 26–42. doi: 10.1037/1040-3590.4.1.26

Guay, F., Ratelle, C. F., Doy, A., and Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: mediating and additive effects. *Learn. Individ. Differ.* 20, 644–653. doi: 10.1016/j.lindif.2010.08.001

Guo, J., Nagengast, B., Marsh, H. W., Kelava, A., Gaspard, H., Brandt, H., et al. (2016). Probing the unique contributions of self-concept, task values, and their interactions using multiple value facets and multiple academic outcomes. *AERA Open* 2, 1–20. doi: 10.1177/2332858415626884

Häfner, I., Flunger, B., Dicke, A. L., Gaspard, H., Brisson, B. M., Nagengast, B., et al. (2018). The role of family characteristics for students' academic outcomes: a person-centered approach. *Child Dev.* 89, 1405–1422. doi: 10.1111/cdev.12809

Harter, S. (2012). "New directions in self-development: resurrecting the I-self," in *Theory Driving Research: New Wave Perspectives on Self-Processes and Human Development*, eds D. McInerney, H. W. Marsh, and R. G. Craven (New York, NY: Information Age Publishing), 1–30.

Herrera, L., Buitrago, R. E., and Cepero, S. (2017). Emotional intelligence in colombian primary school children. location and gender analysis. *Univ. Psychol.* 16, 1–10. doi: 10.11144/Javeriana.upsy16-3.eips

Herrera, L., Lorenzo, O., Defior, S., Fernández-Smith, G., and Costa-Giomi, E. (2011). Effects of phonological and musical training on spanish- and tamazight-speaking children's reading readiness. *Psychol. Music* 39, 68–81. doi: 10.1177/0305735610361995

Herrera, L., Soares-Quadros, J. F. Jr., and Lorenzo, O. (2018). Music preferences and personality in Brazilians. *Front. Psychol.* 9:1488. doi: 10.3389/fpsyg.2018.01488

Hogan, M. J., Parker, J. D., Wiener, J., Watters, C., Wood, L. M., and Oke, A. (2010). Academic success in adolescence: relationships among verbal IQ, social support and emotional intelligence. *Aust. J. Psychol.* 62, 30–41. doi: 10.1080/00049530903312881

Huang, C. (2011). Self-concept and academic achievement: a meta-analysis of longitudinal relations. *J. Sch. Psychol.* 49, 505–528. doi: 10.1016/j.jsp.2011.07.001

Huang, C. (2013). Gender differences in academic self-efficacy: a meta-analysis. *Eur. J. Psychol. Educ.* 28, 1–35. doi: 10.1007/s10212-011-0097-y

Hunagund, D. L., and Hangal, S. J. (2014). Self-efficacy and happiness in youth. *J. Indian Acad. Appl. Psychol.* 40, 70–73.

Inglés, C. J., Díaz, A., García, J. M., Ruiz, C., Delgado, B., and Martínez, M. C. (2012). Auto-atribuciones académicas: diferencias de género y curso en estudiantes de educación secundaria [Academic self-attributions: gender and grade differences in students of secondary education]. *Rev. Latinoam. Psicol.* 44, 57–68. doi: 10.14349/rlp.v44i3.1151

Janošević, M., and Petrović, B. (2019). Effects of personality traits and social status on academic achievement: gender differences. *Psychol. Sch.* 56, 497–509. doi: 10.1002/pits.22215

Jansen, M., Scherer, R., and Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: differential relations to antecedents and educational outcomes. *Contemp. Educ. Psychol.* 41, 13–24. doi: 10.1016/j.cedpsych.2014.11.002

Jonkmann, K., Becker, M., Marsh, H. W., Lüdtke, O., and Trautwein, U. (2012). Personality traits moderate the Big-Fish–Little-Pond effect of academic self-concept. *Learn. Individ. Differ.* 22, 736–746. doi: 10.1016/j.lindif.2012.07.020

Joseph, D. L., and Newman, D. A. (2010). Emotional intelligence: an integrative meta-analysis and cascading model. *J. Appl. Psychol.* 95, 54–78. doi: 10.1037/a0017286

Kun, B., Urbán, R., Paksi, B., Csóbor, L. V., Oláh, A., and Demetrovics, Z. (2012). Psychometric characteristics of the emotional quotient inventory, youth version, short form, in Hungarian high school students. *Psychol. Assess.* 24, 518–523. doi: 10.1037/a0026013

Lohbeck, A., Tietjens, M., and Bund, A. (2016). Physical self-concept and physical activity enjoyment in elementary school children. *Early Child Dev. Care* 186, 1792–1801. doi: 10.1080/03004430.2015.1132708

Lösch, T., Lüdtke, O., Robitzsch, A., Kelava, A., Nagengast, B., and Trautwein, U. (2017). A well-rounded view: using an interpersonal approach to predict achievement by academic self-concept and peer ratings of competence. *Contemp. Educ. Psychol.* 51, 198–208. doi: 10.1016/j.cedpsych.2017.07.003

MacCann, C., Fogarty, G. J., Zeidner, M., and Roberts, R. D. (2011). Coping mediates the relationship between emotional intelligence (EI) and academic achievement. *Contemp. Educ. Psychol.* 36, 60–70. doi: 10.1016/j.cedpsych.2010.11.002

Malo, S., Bataller, S., Casas, F., Gras, M. E., and González, M. (2011). Análisis psicométrico de la escala multidimensional de autoconcepto AF5 en una muestra de adolescentes y adultos de Cataluña [Psychometric analysis of the AF5 multidimensional scale of self-concept in a sample of adolescents and adults in Catalonia]. *Psicothema* 23, 871–878.

Mamata, N., and Sharma, N. R. (2013). Resilience and self-efficacy as correlates of wellbeing among the elderly persons. *J. Indian Acad. Appl. Psychol.* 39, 281–288.

Marsh, H. W., Abduljabbar, A. S., Morin, A. J. S., Parker, P., Abdelfattah, F., Nagengast, B., et al. (2015). The big-fish-little-pond effect: generalizability of social comparison processes over two age cohorts from western, Asian, and middle eastern Islamic countries. *J. Educ. Psychol.* 107, 258–271. doi: 10.1037/a0037485

Marsh, H. W., Craven, R. G., Parker, P. D., Parada, R. H., Guo, J., Dicke, T., et al. (2016a). Temporal ordering effects of adolescent depression, relational aggression, and victimization over six waves: fully latent reciprocal effects models. *Dev. Psychol.* 52, 1994–2009. doi: 10.1037/dev0000241

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., et al. (2016b). Long-term positive effects of repeating a year in school: six-year longitudinal study of self-beliefs, anxiety, social relations, school grades, and test scores. *J. Educ. Psychol.* 109, 425–430. doi: 10.1037/edu0000144

Marsh, H. W., and Hau, K. T. (2003). Big-Fish–Little-Pond effect on academic self-concept: a cross-cultural (26-country) test of the negative effects of academically selective schools. *Am. Psychol.* 58, 364–376. doi: 10.1037/0003-066X.58.5.364

Marsh, H. W., Hau, K. T., Sung, R. Y. T., and Yu, C. W. (2007). Childhood obesity, gender, actual-ideal body image discrepancies, and physical self-concept in Hong Kong children: cultural differences in the value of moderation. *Dev. Psychol.* 43, 647–662. doi: 10.1037/0012-1649.43.3.647

Marsh, H. W., Kuyper, H., Seaton, M., Parker, P. D., Morin, A. J. S., Möller, J., et al. (2014). Dimensional comparison theory: an extension of the internal/external frame of reference effect on academic self-concept formation. *Contemp. Educ. Psychol.* 39, 326–341. doi: 10.1016/j.cedpsych.2014.08.003

Marsh, H. W., and Martin, A. J. (2011). Academic self-concept and academic achievement: relations and causal ordering. *Br. J. Educ. Psychol.* 81, 59–77. doi: 10.1348/000709910X503501

Mayer, J. D., Salovy, P., and Carusso, D. R. (2012). The validity of the MSCEIT: additional analyses and evidence. *Emot. Rev.* 4, 403–408. doi: 10.1177/1754073912445815

McCrae, R. R., and Costa, P. T. (1997). Personality trait structure as a human universal. *Am. Psychol.* 52, 509–516. doi: 10.1037/0003-066X.52.5.509

Mensah, T. (2014). Relationship between selfconcept and school adjustment among Primary school pupils. *J. Educ. Rev.* 13, 205–212.

Meshkat, M., and Nejati, R. (2017). Does emotional intelligence depend on gender? A study on undergraduate english majors of three Iranian Universities. *SAGE Open* 7, 1–8. doi: 10.1177/2158244017725796

Min, M. C., Islam, M. N., Wang, L., and Takai, J. (2018). Cross-cultural comparison of university students' emotional competence in Asia. *Curr. Psychol.* doi: 10.1007/s12144-018-9918-3

Möller, J., Retelsdorf, J., Köller, O., and Marsh, H. W. (2011). The reciprocal internal/external frame of reference model. an integration of models of relations between academic achievement and self-concept. *Am. Educ. Res. J.* 48, 1315–1346. doi: 10.3102/0002831211419649

Mortimer, J. T., Zhang, L., Wu, C. Y., Hussemann, J., and Johnson, M. (2017). Familial transmission of educational plans and the academic self-concept: a three-generation longitudinal study. *Soc. Psychol. Q.* 80, 85–107. doi: 10.1177/0190272516670582

Murayama, K., Pekrun, R., Lichtenfeld, S., and Vom Hofe, R. (2013). Predicting long-term growth in students' mathematics achievement: the unique contributions of motivation and cognitive strategies. *Child Dev.* 84, 1475–1490. doi: 10.1111/cdev.12036

Muris, P., Meesters, C., and Diederen, R. (2005). Psychometric properties of the Big Five Questionnaire for Children (BFQ-C) in a Dutch sample of young adolescents. *Pers. Individ. Differ.* 38, 1757–1769. doi: 10.1016/j.paid.2004.11.018

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

Nagy, G., Watt, H., Eccles, J. S., Trautwein, U., Lüdtke, O., and Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: different countries, different trajectories. *J. Res. Adolesc.* 20, 482–506. doi: 10.1111/j.1532-7795.2010.00644.x

Nasir, R., and Lin, L. S. (2012). The relationship between self-concept and career awareness amongst students. *Asian Soc. Sci.* 9, 193–201. doi: 10.5539/ass.v9n1p193

Oberle, E., Domitrovich, C. E., Meyers, D. C., and Weissberg, R. P. (2016). Establishing systemic social and emotional learning approaches in schools: a framework for schoolwide implementation. *Camb. J. Educ.* 46, 277–297. doi: 10.1080/0305764X.2015.1125450

O'Connor, M. C., and Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Pers. Individ. Differ.* 43, 971–990. doi: 10.1016/j.paid.2007.03.017

Paramanik, J., Saha, B., and Modal, B. C. (2014). Adjustment of secondary school students with respect to gender and residence. *Am. J. Educ. Res.* 2, 1138–1143. doi: 10.12691/education-2-12-2

Parker, P. D., Martin, A. J., and Marsh, H. W. (2008). Factors predicting life satisfaction: a process model of personality, multidimensional self-concept, and life satisfaction. *J. Psychol. Couns. Sch.* 18, 15–29. doi: 10.1375/ajgc.18.1.15

Perret, P., Ayad, M., Dauvier, B., and Congard, A. (2019). Self- and parent-rated measures of personality are related to different dimensions of school adjustment. *Learn. Individ. Differ.* 70, 182–189. doi: 10.1016/j.lindif.2019.02.007

Petrides, K. V., Vernon, P. A., Schermer, J. A., Lightart, L., Boomsma, D. I., and Veselka, L. (2010). Relationships between trait emotional intelligence and the Big Five in the Netherlands. *Pers. Individ. Differ.* 48, 906–910. doi: 10.1016/j.paid.2010.02.019

Piqueras, J. A., Mateu-Martínez, O., Cejudo, J., and Pérez-González, J. C. (2019). Pathways into psychosocial adjustment in children: modeling the effects of trait emotional intelligence, social-emotional problems, and gender. *Front. Psychol.* 10:507. doi: 10.3389/fpsyg.2019.00507

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychol. Bull.* 135, 322–338. doi: 10.1037/a0014996

Pulido, F., and Herrera, F. (2017). La influencia de las emociones sobre el rendimiento académico [The influence of emotions on academic performance]. *Cienc. Psicol.* 11, 29–39. doi: 10.22235/cp.v11i2.1344

Rahafar, A., Castellana, I., Randler, C., and Antúnez, J. M. (2017). Conscientiousness but not agreeableness mediates females' tendency toward being a morning person. *Scand. J. Psychol.* 58, 249–253. doi: 10.1111/sjop.12362

Salami, S. O., and Ogundokun, M. O. (2009). Emotional intelligence and self-efficacy as predictors of academic performance. *Perspect. Educ.* 25, 175–185.

Schmitt, D. P., Allik, J., McCrae, R. R., and Benet-Martinez, V. (2007). The geographic distribution of Big Five personality traits: patterns and profiles of human self-description across 56 nations. *J. Cross Cult. Psychol.* 38, 173–212. doi: 10.1177/0022022106297299

Schwarzer, R., and Fuchs, R. (2009). "Self-efficacy and healthy behaviours," in *Predicting Health Behaviour: Research and Practice in Social Cognition Models*, 2nd Edn, eds M. Conner, and P. Norman (Buckingham: Open University Press), 163–196.

Seaton, M., Marsh, H. W., and Craven, R. G. (2010). Big-Fish-Little-Pond effect: generalizability and moderation-Two sides of the same coin. *Am. Educ. Res. J.* 47, 390–434. doi: 10.3102/0002831209350493

Sewasew, D., and Schroeders, U. (2019). The developmental interplay of academic self-concept and achievement within and across domains among primary school students. *Contemp. Educ. Psychol.* 58, 204–212. doi: 10.1016/j.cedpsych.2019.03.009

Shoshani, A., and Slone, M. (2017). Positive education for young children: effects of a positive psychology intervention for preschool children on subjective well being and learning behaviors. *Front. Psychol.* 8:1866. doi: 10.3389/fpsyg.2017.01866

Skaalvik, E. M., and Skjaalvik, S. (2013). School goal structure: associations with students' perceptions of their teachers as emotionally supportive, academic self-concept, intrinsic motivation, effort, and help seeking behavior. *Int. J. Educ. Res.* 61, 5–14. doi: 10.1016/j.ijer.2013.03.007

Smith-Woolley, E., Selzam, S., and Plomin, R. (2019). Polygenic score for educational attainment captures DNA variants shared between personality traits and educational achievement. *J. Pers. Soc. Psychol.* 117, 1145–1163. doi: 10.1037/pspp0000241

Spengler, M., Brunner, M., Martin, R., and Lüdtke, O. (2016). The role of personality in predicting (change in) students' academic success across four years of secondary school. *Eur. J. Psychol. Assess.* 32, 95–103. doi: 10.1027/1015-5759/a000330

Stoet, G., and Geary, D. C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence* 48, 137–151. doi: 10.1016/j.intell.2014.11.006

Strayhorn, T. L. (2010). When race and gender collide: social and cultural capital's influence on the academic achievement of african american and latino males. *Rev. High. Educ.* 33, 307–332. doi: 10.1353/rhe.0.0147

Suldo, S. M., Hearon, B. V., Bander, B., McCullough, M., Garofano, J., Roth, R. A., et al. (2015). Increasing elementary school students' subjective well-being through a classwide positive psychology intervention: results of a pilot study. *Contemp. Sch. Psychol.* 19, 300–311. doi: 10.1007/s40688-015-0061-y

Susperreguy, M. I., Davis-Kean, P. E., Duckworht, K., and Chen, M. (2018). Self-concept predicts academic achievement across levels of the achievement distribution: domain specificity for math and reading. *Child Dev.* 89, 2196–2214. doi: 10.1111/cdev.12924

Szczygieł, D., and Mikolajczak, M. (2017). Why are people high in emotional intelligence happier? They make the most of their positive emotions. *Pers. Individ. Differ.* 117, 177–181. doi: 10.1016/j.paid.2017.05.051

Usán, P., and Salavera, C. (2018). School motivation, emotional intelligence and academic performance in students of secondary education. *Actual. Psicol.* 32, 95–112. doi: 10.15517/ap.v32i125.32123

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

Whaley, A. L., and Noël, T. (2012). Sociocultural theories, academic achievement, and African American adolescents in a multicultural context: a review of the cultural compatibility perspective. *J. Negro Educ.* 81, 25–38. doi: 10.7709/jnegroeducation.81.1.0025

Wolff, F., Nagy, N., Helm, F., and Möller, J. (2018). Testing the internal/external frame of reference model of academic achievement and academic self-concept with open self-concept reports. *Learn. Instr.* 55, 58–66. doi: 10.1016/j.learninstruc.2017.09.006

Wosu, T. M. (2013). Self-concept as a correlate of social adjustment. *Int. J. Educ. Rev.* 7, 211–219.

Xie, F., Xin, Z., Chem, X., and Zhang, L. (2019). Gender difference of chinese high school students' math anxiety: the effects of self-esteem, test anxiety and general anxiety. *Sex Roles* 81, 235–244. doi: 10.1007/s11199-018-0982-9

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.

Copyright © 2020 Herrera, Al-Lal and Mohamed. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 29 January 2020
doi: 10.3389/fpsyg.2019.03050

Gender Differences in Mathematics Motivation: Differential Effects on Performance in Primary Education

Susana Rodríguez[1], Bibiana Regueiro[2], Isabel Piñeiro[1], Iris Estévez[3] and Antonio Valle[1]*

[1] Department of Psychology, University of A Coruña, A Coruña, Spain, [2] Department of Pedagogy and Didactics, University of Santiago de Compostela, Santiago de Compostela, Spain, [3] Department of Pedagogy and Didactics, University of A Coruña, A Coruña, Spain

In addition to attempting to verify gender differences, this study aims to examine the explanatory potential of boys' and girls' attitudes toward mathematics on their performance. The sample comprised 897 students in the 5th and 6th years of primary education (450 boys and 447 girls). The results confirm what previous research has suggested, that girls tended to exhibit less positive attitudes about mathematics than their male classmates, in particular lower motivation, worse perception of competence, and higher rates of anxiety, although in all cases the effect sizes were small. Even though there no significant gender differences in academic performance, as expected, the explanatory power of attitudes toward mathematics was clearly more significant in boys than in girls (R^2 = 0.194 and R^2 = 0.103, respectively). The results of the regression analysis for each sample reinforce the well-known positive impact of perceived self-efficacy on mathematics performance and introduce the effect of achievement emotions of academic performance. Test anxiety in mathematics seems to only have a negative effect on boys' grades, as this variable does not appear in the regression equation when explaining girls' performance. In the light of control-value theory, we discuss the contingency of perceived competence and its involvement in anxiety and academic performance. Boys' results could be affected by the levels of anxiety inasmuch as they tend to be confident in their abilities, motivated to stand out, and interested in mathematics. Whereas despite girls reporting high rates of anxiety, what may have a negative impact on their results might have more to do with a higher value placed on mathematics, as their perception of control may be low.

Keywords: gender differences, mathematics, academic motivation, academic performance, primary education

OPEN ACCESS

Edited by:
Juan Luis Castejon, University of Alicante, Spain

Reviewed by:
Jose I. Navarro, University of Cádiz, Spain
Celestino Rodríguez, Universidad de Oviedo Mieres, Spain

****Correspondence:***
Susana Rodríguez
susana.rodriguez1@udc.es

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

Received: *18 October 2019*
Accepted: *23 December 2019*
Published: *29 January 2020*

Citation:
Rodríguez S, Regueiro B, Piñeiro I, Estévez I and Valle A (2020) Gender Differences in Mathematics Motivation: Differential Effects on Performance in Primary Education. Front. Psychol. 10:3050. doi: 10.3389/fpsyg.2019.03050

INTRODUCTION

Spanish students' scores in the mathematics tests in Programme for International Student Assessment (PISA) 2015 (Ministerio de Educación Cultura y Deporte [MECD], 2016) do not help us to be optimistic about the teaching and learning in this subject in our country. The score in the mathematics test from May 2015, which evaluated 37,205 of the 414,276 15-year-old students in Spain, was 486 points, significantly lower than the OECD average of 490 points. In a ranking of the 36 member countries, Spain is in 26th position, and only 7.2% of Spanish students

reached high levels of achievement (5 and 6) in mathematics, which is 3.5 percentage points less than the OECD average (10.7%).

These results seem to reflect a problem that continually calls into question the effectiveness of teaching–learning processes for mathematics content. Motivational and emotional variables, which involve beliefs, emotions, and attitudes, seem to be important when it comes to fully understanding and explaining the results. In fact, everything points toward math achievement being related to variables such as perceived competence and self-efficacy (Randhawa et al., 1993; Pajares and Graham, 1999; Fast et al., 2010; Williams and Williams, 2010; Parker et al., 2014), interest (Köller et al., 2001; Lee et al., 2014; Jansen et al., 2016) and anxiety (Ashcraft, 2002; Ashcraft and Ridley, 2005; Lyons and Beilock, 2012; Ramirez et al., 2016).

In this context, when we look at the results of PISA 2015 (Ministerio de Educación Cultura y Deporte [MECD], 2016), the difference in mathematics performance between boys and girls aged 15 is an average of 8 points in OECD countries and 11 points in the European Union (EU) overall, with boys scoring higher than girls. In Spain, boys scored 16 points higher than girls, much higher than the OECD average. Various meta-analyses in the 1970s and 1980s demonstrated a slight male advantage in mathematics in secondary students (Hyde et al., 1990; Hedges and Nowell, 1995), at least in complex problem-solving tasks. However, this varies by country (Else-Quest et al., 2010; Mullis et al., 2012; Ministerio de Educación Cultura y Deporte [MECD], 2016), and more recent data suggest that this gender gap may be disappearing (Hyde et al., 2008; Hyde and Mertz, 2009; Lindberg et al., 2010).

It is an established fact that beliefs and attitudes can have a significant impact on the decision to choose a professional career related to mathematics (Colbeck et al., 2001; Ceci and Williams, 2011; Sadler et al., 2012; Kanny et al., 2014; Legewie and DiPrete, 2014; Wang and Degol, 2017). In addition, women are often underrepresented in STEM programs (science, technology, engineering, and mathematics programs), particularly in engineering, physical sciences, and computational sciences in western universities (e.g., Larivière et al., 2013; Hyde, 2014). With that in mind, it is essential to include this topic in the research agenda of educational psychology in order to thoroughly understand the interaction between motivational constructs and each group's performance (girls and boys). In addition, research suggests that gender differences in mathematics ability are minimal during early childhood and do not consistently arise until middle to late adolescence (e.g., Lindberg et al., 2010; Robinson and Lubienski, 2011). The current study addresses the need to delve into the affective-motivational dimension of students in the final years of primary education, as girls may be losing their motivation for mathematics as they advance through this particular educational stage.

Academic Motivation and Performance in Mathematics

Motivational research in mathematics has addressed constructs such as self-efficacy, which indicates students' judgments of their abilities to perform specific mathematical tasks, and interest in mathematics, as positive affective experiences in activities related to mathematics. As laid out in *expectancy-value* theory, both expectations of success and subjective values related to the task directly influence performance, attainment, effort, and persistence (Wigfield and Eccles, 2000). Various studies have shown that perceived self-efficacy positively predicts academic achievement in mathematics (Randhawa et al., 1993; Pajares and Graham, 1999; Stevens et al., 2006; Fast et al., 2010; Williams and Williams, 2010; Parker et al., 2014). Similarly, students' interest in mathematics is associated with a strong preference for mathematics content, which translates to sustained commitment over time and better performance (Köller et al., 2001; Hidi and Renninger, 2006; Lee et al., 2014; Jansen et al., 2016) in both childhood and adolescence (Lepper et al., 2005; Aunola et al., 2006; Denissen et al., 2007; Viljaranta et al., 2009).

There is also evidence that math anxiety is negatively related to performance, leading to avoidance and diminishing working memory resources needed to deal with mathematical tasks (Ashcraft, 2002; Ashcraft and Ridley, 2005; Lyons and Beilock, 2012; Ramirez et al., 2016). The nuances of the research must be understood when it comes to operationalizing the measure of math anxiety, as must the fact that the full spectrum of emotional reactions may be associated with it (Goetz and Hall, 2013; Pekrun and Linnenbrink-Garcia, 2014) and affect academic performance (Pekrun et al., 2014, 2017). For that reason, in this study, we address negative emotions associated with mathematics and math anxiety separately. In accordance with the *control-value theory* of achievement emotions (Pekrun, 2006), students who enjoy mathematics are assumed to focus their attention on the tasks, making better use of deep learning strategies and, therefore, getting better results. Students who are, for example, bored in mathematics classes pay less attention and make less use of learning strategies or use more superficial strategies, which leads to them having lower achievement (Pekrun, 2006; Pekrun et al., 2017; Putwain et al., 2018).

Gender Differences in Mathematics Motivation

Research has confirmed gender differences, even in primary education, in mathematics self-concept, self-efficacy, and interest, suggesting that boys generally have better motivational profiles in mathematics than have girls (Eccles et al., 1993; Kurtz-Costes et al., 2008). The study of gender differences in mathematics motivation in these first educational phases is of particular interest. In this way, the most could be made of female students' potential, improving the gender balance of participation in future STEM courses, which have been recognized as a critical filter for highly qualified and highly paid jobs.

According to previous research (Fredricks and Eccles, 2002; Jacobs et al., 2002; Preckel et al., 2008; Else-Quest et al., 2010; Frenzel et al., 2010; Guo et al., 2015; Ganley and Lubienski, 2016), girls report lower levels of individual interest and perceived mathematics competence. The most significant differences are in secondary school and university students rather than students in lower educational levels.

Demonstrating the need to explore gender differences in academic motivation, previous research has concluded that, for example, mathematics self-concept could be positively linked to achievement in boys but could even have a negative effect on girls' achievement (Yoon et al., 1996). It has also demonstrated that the impact of interest on mathematics achievement may be slightly more important for girls than for boys (Ganley and Lubienski, 2016). On this point, although expectation-value theory does not develop a theoretical framework to address gender differences in particular, it may be used to facilitate the interpretation of the differential impact of self-beliefs and values (Wigfield et al., 1991, 1997; Eccles et al., 1993).

In terms of gender differences related to emotions provoked by mathematics, among which studies on math anxiety stand out, research indicates (albeit with small effect sizes) the existence of greater rates of anxiety in girls than boys during tasks involving mathematical reasoning (Hyde et al., 1990; Else-Quest et al., 2010). Assuming the possibility that these differences start as early as primary school (Yüksel-şahin, 2008; Griggs et al., 2013), research suggests the existence of a higher rate of math anxiety, and also more negative feelings and attitudes in boys than girls (Hyde et al., 1990; Nagy et al., 2008; Goetz and Hall, 2013; Goetz et al., 2013; Bieg et al., 2015). According to the principles of control-value theory, studies addressing the different impact suggest that positive emotions associated with mathematics could have a more pronounced effect on girls' dedication (Pinxten et al., 2014) and that the rates of anxiety and negative emotions may not affect girls' achievement as negatively as might have been expected (Goetz et al., 2013).

Based on these considerations, and apart from attempting to verify gender differences, the primary purpose of this study is to analyze the possible differential impact of the variables used to explore mathematical motivation on the academic achievement of boys and girls in primary education. Firstly, we hypothesize that there will be statistically significant differences in mathematical motivation between boys and girls. We expect the boys' motivational pattern, in terms of perceived competence and intrinsic motivation for mathematics, to be more positive than the girls' (Fredricks and Eccles, 2002; Jacobs et al., 2002; Preckel et al., 2008; Else-Quest et al., 2010; Frenzel et al., 2010; Guo et al., 2015; Ganley and Lubienski, 2016), and the girls to exhibit more negative feelings and greater anxiety toward mathematics than do the boys (Hyde et al., 1990; Nagy et al., 2008; Goetz et al., 2013; Bieg et al., 2015). We will examine the explanatory potential and differential incidence of competency beliefs and intrinsic motivation, along with negative feelings and anxiety over mathematics performance in boys and girls, assuming that perceived competence will have an impact on academic achievement in both cases (Randhawa et al., 1993; Pajares and Graham, 1999; Stevens et al., 2006; Fast et al., 2010; Williams and Williams, 2010; Parker et al., 2014). We will also explore the possibility that the impact of intrinsic motivation on performance will be more significant when it comes to explaining girls' achievement rather than boys' (e.g., Yoon et al., 1996; Ganley and Lubienski, 2016). Finally, we will also look at whether anxiety is more important when explaining boys,' compared with girls,' mathematics achievement (e.g., Goetz et al., 2013).

MATERIALS AND METHODS

Participants

The sample was composed of 897 students from 13 public primary schools in the Spanish province of A Coruña. Half (50.2%) were boys and half (49.8%) were girls. They were aged between 9 and 13 years (M = 10.77; SD = 0.69). Out of the total, 437 were in the 5th year of primary education (223 boys and 213 girls) and 460 were in the 6th year of primary education (227 boys and 233 girls).

Instruments

To measure students' attitudes toward mathematics, we used the IAM (*Inventario de Actitudes hacia las Matemáticas/Inventory of Attitudes Toward Mathematics*). This instrument is the result of expanding the *Fennema–Sherman Mathematics Attitude Scale* (FSS) by Fennema and Sherman (1976). It is an extended version of the scale, with some modifications, adapted into Spanish and including new dimensions aimed at measuring more accurately students' attitudes and motivation for mathematics (Silva, 2005; González-Pienda et al., 2012; Cueli et al., 2014). In this study, we used the following IAM dimensions:

- *Intrinsic motivation toward mathematics* (four items; α = 0.72): this evaluates motivation toward learning and understanding mathematics content for the pleasure and personal satisfaction that comes from working with this type of content (example item: "I find mathematics enjoyable and stimulating").
- *Perceived competence in mathematics* (four items; α = 0.75): this evaluates the level of confidence in oneself for learning and getting good mathematics results (example item: "I think I can do even more difficult math tasks").
- *Negative feelings caused by mathematics* (three items; α = 0.71): this evaluates the presence and intensity of sadness and unease caused by studying, homework, or attending math classes (example item: "In math class I am sad and unhappy").
- *Math anxiety* (three items; α = 0.77): this evaluates the level of students' fear and nervousness with math tests and tasks (example item: "I feel uncomfortable and nervous about mathematics").

The items in each dimension were in a Likert-type format with five response options from 1 (completely false) to 5 (completely true).

The evaluation of *academic performance in mathematics* was obtained via the final school grades that the participating students received in this subject. The following grades were used: 1 = poor, 2 = satisfactory, 3 = good, 4 = very good, and 5 = outstanding.

Procedure

The data were collected during school hours by personnel external to the school with the prior consent of the school management and the students' teachers. Prior to participating in the study, the teachers, students, and parents (depending on school regulations) were informed of the study content and procedure. Before data collection, which was done on a single

occasion, the participants were reminded of the importance of answering the various questions honestly.

Data about the target variables were collected in accordance with the recommendations of the ethical standards established in the Research and Teaching Ethics Committee of the University of A Coruña and the Declaration of Helsinki. Confidentiality of data was ensured, and participation was voluntarily such that withdrawal from the study was possible at any time.

Data Analysis

In addition to descriptive and correlational analyses, analysis of variance (ANOVA) was also used to examine gender differences in mathematics motivation. A stepwise linear regression analysis was performed to examine the predictive capacity of the motivational variables on achievement in mathematics between boys and girls. In both cases, the predictor variables were *intrinsic motivation*, *perceived competence*, *negative feelings*, and *math anxiety*, with the criterion variable being the students' final academic grade in mathematics. Firstly, the analysis aimed to understand the contribution of each of the variables added to the regression equation when it came to predicting mathematics performance. In addition, we wanted to ascertain the weight and specific significance of the predictor variables in each sample. Effect sizes were calculated according to the criteria in Cohen (1988) classic work: $d < 0.20$ = non-significant effect; $d \geq 0.20$ and $d < 0.50$ = small effect; $d \geq 0.50$ and $d < 0.80$ = moderate effect; and $d \geq 0.80$ = large effect.

RESULTS

The analysis of results first looks at determining whether there are significant differences in *intrinsic motivation*, *perceived competence*, *negative feelings*, and *math anxiety* by *gender*. Following that, we examine the contribution of this set of affective-motivational variables on students' academic performance.

Table 1 shows the correlation coefficients as well as descriptive statistics of central tendency, distribution, and dispersion for the variables in this study. Considering the correlations, academic achievement in mathematics had a positive, statistically significant relationship to *intrinsic motivation* toward mathematics ($r = 0.19$, $p < 0.01$) and *perceived competence* ($r = 0.28$, $p < 0.01$). *Anxiety* and *negative feelings* toward mathematics had a negative association with performance ($r = -0.29$, $p < 0.01$, and $r = -0.33$, $p < 0.01$, respectively).

The relationship between perceived competence and both negative emotions and math anxiety was negative and significant ($r = -0.43$, $p < 0.01$, and $r = -0.49$, $p < 0.01$, respectively). Similarly, negative feelings toward mathematics and anxiety were both significantly, negatively correlated with intrinsic motivation toward mathematics ($r = -0.50$, $p < 0.01$, and $r = -0.49$, $p < 0.01$, respectively). As one might expect, there was a significant, positive association between negative feelings toward mathematics and math anxiety ($r = 0.47$, $p < 0.01$).

TABLE 1 | Means, standard deviations, asymmetry, kurtosis, and correlation matrix.

	1	2	3	4	5
1. Perceived competence	–				
2. Intrinsic motivation	0.68[a]	–			
3. Negative feelings	−0.43[a]	−0.50[a]	–		
4. Anxiety	−0.49[a]	−0.44[a]	0.47[a]	–	
5. Academic performance	0.28[a]	0.19[a]	−0.33[a]	−0.29[a]	–
M	4.04	3.71	1.77	2.10	3.41
SD	0.75	0.86	0.87	1.07	1.27
Asymmetry	−0.89	−0.50	1.27	0.93	−0.43
Kurtosis	0.79	−0.18	1.44	0.10	−0.91

[a]$p < 0.01$.

Differences in Mathematics Motivation Between Boys and Girls

We found statistically significant differences in mathematics motivation between boys and girls in primary education, with a moderate effect size, even though the differences in mathematics performance were not significant [$F_{(1,895)} = 1.174$, $p = 0.279$]. Girls reported lower levels of *self-efficacy* [$F_{(1,985)} = 11.227$; $p < 0.01$; $\eta_p^2 = 0.012$; $d = 0.71$] and *intrinsic motivation* [$F_{(1,895)} = 6.522$; $p < 0.05$; $\eta_p^2 = 0.007$; $d = 0.61$] for mathematics than did boys. On the other hand, whereas both boys and girls had similar levels of *negative feelings* toward mathematics [$F_{(1,895)} = 1.272$, $p = 0.260$], the girls exhibited higher rates of *math anxiety* than did their male classmates [$F_{(1,895)} = 11.018$; $p < 0.01$; $\eta_p^2 = 0.012$; $d = 0.70$].

The Impact of Motivation on Mathematics Performance in Girls and Boys

We performed two identical stepwise regression analyses, one for the sample of boys ($n = 450$; $M_{age} = 10.79$; $SD = 0.70$) and one for the girls ($n = 447$; $M_{age} = 10.75$; $SD = 0.68$), with the aim of determining the predictive value of the motivational variables on mathematics performance in each sample.

As **Table 2** shows, in the boys' group, both negative feelings and anxiety associated with mathematics, along with perceived competence in the subject, contributed significantly to the prediction of academic performance [$F_{(3,446)} = 36.914$; $p < 0.001$], explaining almost 20% of the variance ($R^2 = 0.194$) in the criterion variable.

In the girls' group, intrinsic motivation, negative feelings toward mathematics, and perceived competence in the subject were also predictor variables of mathematics performance [$F_{(3,443)} = 18.093$; $p < 0.001$], although in this case the percentage of explained variance was half of that in the boys' sample ($R^2 = 0.103$).

The results of the analyses confirm the explanatory potential of both *negative feelings* toward mathematics ($\beta = -0.207$; $t = -4.156$; $p < 0.001$ for the boys, and $\beta = -0.267$; $t = -5.087$; $p < 0.001$ for the girls) and *perceived competence* ($\beta = 0.162$; $t = 3.301$; $p < 0.001$ for the boys, and $\beta = 0.225$; $t = 3.625$; $p < 0.001$ for the girls) in mathematics performance (see **Table 2**).

TABLE 2 | Explained variance (R^2), change in R^2 (ΔR^2), regression coefficients (β), and associated statistical significance [$t(p <)$] in the prediction of boys' and girls' mathematics performance.

		R^2_{adjust}	ΔR^2	β	$t(p <)$
Boys					
Model 1		0.131	0.133		
	Negative feelings			−0.365	−8.299***
Model 2		0.176	0.046		
	Negative feelings			−0.253	−5.228***
	Anxiety			−0.242	−5.011***
Model 3		0.194	0.020		
	Negative feelings			−0.207	−4.156***
	Anxiety			−0.192	−3.819***
	Competence			0.162	3.301***
Girls					
Model 1		0.080	0.082		
	Negative feelings			−0.287	−6.322***
Model 2		0.093	0.015		
	Negative feelings			−0.228	−4.525***
	Competence			0.134	2.670***
Model 3		0.103	0.012		
	Negative feelings			−0.267	−5.087***
	Competence			0.225	3.625***
	Intrinsic Motivation			−0.159	−2.667**

**$p < 0.01$ and ***$p < 0.001$.

In both samples, negative feelings were associated with poor mathematics performance, whereas the positive regression coefficient for perceived competence in the subject indicates, as previous research has noted, the positive link between these types of beliefs and performance.

The results also indicate some differential aspects that should be underscored. Firstly, the contribution of math anxiety was only an explanatory factor for the boys ($\beta = -0.192$; $t = -3.819$; $p < 0.001$), whereas intrinsic motivation appeared as a significant variable in the explanation of girls' performance ($\beta = -0.159$; $t = -2.667$; $p < 0.05$). The negative regression coefficient of this variable indicates that, for girls, commitment to mathematics for intrinsic reasons may negatively affect their performance in this subject. We discuss this result in light of Pekrun's (2006) control-value theory.

DISCUSSION AND CONCLUSION

In line with previous research (Hyde et al., 1990; Else-Quest et al., 2010; Goetz et al., 2013; Bieg et al., 2015), we confirm that girls in the 5th and 6th years of primary education report higher levels of math anxiety than do boys, in this case, with moderate effect sizes. However, these gender differences in math anxiety may not be transferable to other negative feelings or emotions such as sadness or boredom, as previous studies have suggested (Hyde et al., 1990; Nagy et al., 2008; Goetz and Hall, 2013; Goetz et al., 2013; Bieg et al., 2015). Our results also replicate prior studies, showing that boys report greater intrinsic motivation and perceived competence for mathematics than do girls (Fredricks and Eccles, 2002; Jacobs et al., 2002; Preckel et al., 2008; Else-Quest et al., 2010; Frenzel et al., 2010; Louis and Mistele, 2012; Guo et al., 2015; Ganley and Lubienski, 2016), even though their performance in mathematics is similar (Skaalvik and Skaalvik, 2004; Else-Quest et al., 2010; Lindberg et al., 2010).

In general terms, as expected, our results confirm the well-known positive impact of perceived self-efficacy on mathematics performance at school (Zimmerman, 1995; Pajares and Graham, 1999; Valentine et al., 2004; Fast et al., 2010; Williams and Williams, 2010; Parker et al., 2014), found in both standardized tests and, as in our case, the grades awarded in class (Randhawa et al., 1993; Fast et al., 2010; Grigg et al., 2018).

In addition, as various perspectives have suggested, our results highlight the importance and influence of students' emotions on their learning and performance. The explanatory power of negative emotions associated with mathematics has been shown to be even more important than the oft-quoted math anxiety when it comes to explaining performance in this subject. Research on emotions in mathematics has repeatedly focused on math anxiety, very probably ignoring a good number of different emotions such as sadness and boredom, which as "hot cognition," may mark the path of learning and success. Even though studies indicate that the relationships between emotions and performance generally tend to have small or moderate effect sizes, it is possible that academic emotions end up having a significant cumulative long-term impact on performance. For this reason, we should not lose sight of their role in health, subjective well-being, choice and continuation of study, or in lifelong learning (Pekrun, 2006).

The differences in percentages of variance in academic performance explained by the motivational variables that we examined in boys and girls indicate the need to encourage research into gender differences in academic emotions and motivations in general and in mathematics in particular. Whereas negative feelings, anxiety, and perceived competence together explain almost 20% of the variance in boys' performance in mathematics, the explanatory power of negative feelings, competence, and intrinsic motivation in girls is practically half of that. Producing a body of knowledge that would let us characterize differential motivational profiles, in contexts of achievement, would allow better adjustment of affective-motivational interventions. The gender gap in cognitive abilities could be reduced in the long term with the promotion of specific educational experiences in these first stages of formal education (Ganley and Lubienski, 2016).

In this study, the performance of the boys, who were more confident in their abilities, more motivated, and more interested in mathematics, was seen to diminish in relation to the appearance of negative feelings about mathematics, and there were high rates of anxiety. Nevertheless, and in line with observations by Goetz et al. (2013), our results seem to demonstrate that although girls reported higher levels of anxiety, their performance in mathematics tasks and tests was not affected as negatively as might have been expected. The state-trait discrepancy that research recognizes in these and other academic emotions (e.g., Porter et al., 2000; Frenzel et al., 2007;

Bieg et al., 2013, 2014) suggests that gender differences found in math anxiety may not be reflected in day-to-day school learning processes. Emotional valuations may be strongly influenced by subjective beliefs such as expectations or attributional tendencies. The commonly accepted idea that in mathematics girls believe themselves to be less competent may be behind an erroneous anticipatory evaluation of their math anxiety (e.g., Skaalvik and Skaalvik, 2004; Ganley and Lubienski, 2016). These girls' beliefs about their competence in mathematics, together with gender stereotypes (Wigfield and Eccles, 2000; Pekrun, 2006; Thoman et al., 2013; Bieg et al., 2015), contribute to an emotional evaluation that differs from the emotions really associated with the specific tasks or situations. This is why despite girls exhibiting higher levels of anxiety about mathematics than do boys, this variable is not an explanatory factor in girls' performance.

With respect to the premises of the control-value theory, negative feelings such as sadness, unhappiness, and despair associated with mathematics tasks are surely the result of a pattern of low control with high value placed on success (Frenzel et al., 2007; Bieg et al., 2015). The inclusion of intrinsic motivation as a negative predictor of girls' mathematics performance must be understood in this context, characterized by low perceived competence for the subject. Just as Pekrun's (2006) theory determines, if negative emotions are the result of a pattern of low control together with high value placed on success, involvement for intrinsic reasons, indicating high value placed on mathematics, could end up decreasing girls' performance inasmuch as it could contribute to negative feelings in this area.

EDUCATIONAL IMPLICATIONS

In terms of the potential of emotions associated with the classroom, it is worth suggesting longitudinal, experimental, and interventional research to examine the assumed causal relationships between discrete emotions, both negative (as we have done in this study) and positive (enjoyment, hope, and gratitude among others), and performance. Apart from incorporating knowledge about moderating and mediating variables of academic emotions and their background into the teacher training syllabus, teachers should also consider the impact of educational styles on their students' academic emotions, and their role in encouraging positive emotions and reducing negative emotions associated with the classroom.

In terms of our results, and in line with Ganley and Lubienski (2016), mathematics interventions for girls should start early and deal specifically with perceptions of confidence and control. On the other hand, management of emotions, particularly anxiety, could be extremely important for mathematics interventions for boys. The identification and development of instruction strategies and intervention plans to improve the affective-emotional experiences associated with the learning process should be on the educational and research agenda.

LIMITATIONS

Apart from the inherent limitations of the research design, we must recognize the use of a self-report to measure emotions and the fact that boys and girls surely differ in their abilities and dispositions when reporting their emotions (Bryant et al., 1996). Information about emotions may be especially vulnerable to social desirability or to stereotyping in a context such as mathematics, which is perceived as a male domain.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

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

AUTHOR CONTRIBUTIONS

SR and IE contributed to collect the data, data analysis, and writing the manuscript. BR and AV contributed to data analysis and writing the manuscript. IP wrote the manuscript.

FUNDING

This work was carried out with financing from the research projects EDU2013-44062-P (MINECO) and EDU2017-82984-P (MEIC).

REFERENCES

Ashcraft, M. H. (2002). Math anxiety: personal, educational, and cognitive consequences. *Curr. Dir. Psychol. Sci.* 11, 181–185. doi: 10.1111/1467-8721.00196

Ashcraft, M. H., and Ridley, K. S. (2005). "Math anxiety and its cognitive consequences," in *Handbook of Mathematical Cognition*, ed. J. I. D. Campbell, (New York, NY: Psychology Press), 315–327.

Aunola, K., Leskinen, E., and Nurmi, J. E. (2006). Developmental dynamics between mathematical performance, task motivation, and teachers' goals during the transition to primary school. *Br. J. Educ. Psychol.* 76, 21–40. doi: 10.1348/000709905X51608

Bieg, M., Goetz, T., and Hubbard, K. (2013). Can I master it and does it matter? An intraindividual analysis on control–value antecedents of trait and state academic emotions. *Learn. Individ. Differ.* 28, 102–108. doi: 10.1016/j.lindif.2013.09.006

Bieg, M., Goetz, T., and Lipnevich, A. A. (2014). What students think they feel differs from what they really feel–academic self-concept moderates the discrepancy between students' trait and state emotional self-reports. *PLos One* 9:e92563. doi: 10.1371/journal.pone.0092563

Bieg, M., Goetz, T., Wolter, I., and Hall, N. C. (2015). Gender stereotype endorsement differentially predicts girls' and boys' trait-state discrepancy in math anxiety. *Front. Psychol.* 6:1404. doi: 10.3389/fpsyg.2015.01404

Bryant, F. B., Yarnold, P. R., and Grimm, L. G. (1996). Toward a measurement model of the affect intensity measure: a three-factor structure. *J. Res. Pers.* 30, 223–247. doi: 10.1006/jrpe.1996.0015

Ceci, S. J., and Williams, W. M. (2011). Understanding current causes of women's underrepresentation in science. *Proc. Natl. Acad. Sci.* 108, 3157–3162. doi: 10.1073/pnas.1014871108

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. New Jersey: Lawrence Erlbaum.

Colbeck, C. L., Cabrera, A. F., and Terenzini, P. T. (2001). Learning professional confidence: linking teaching practices, students' self-perceptions, and gender. *Rev. High. Educ.* 24, 173–191. doi: 10.1353/rhe.2000.0028

Cueli, M., González-Castro, P., Álvarez, L., García, T., and González-Pienda, J. A. (2014). Variables afectivo-motivacionales y rendimiento en matemáticas: un análisis bidireccional. *Rev. Mex. Psicol.* 31, 153–163.

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

Eccles, J., Wigfield, A., Harold, R. D., and Blumenfeld, P. (1993). Age and gender differences in children's self-and task perceptions during elementary school. *Child Dev.* 64, 830–847. doi: 10.1111/j.1467-8624.1993.tb02946.x

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

Fast, L. A., Lewis, J. L., Bryant, M. J., Bocian, K. A., Cardullo, R. A., Rettig, M., et al. (2010). Does math self-efficacy mediate the effect of the perceived classroom environment on standardized math test performance? *J. Educ. Psychol.* 102, 729–740. doi: 10.1037/a0018863

Fennema, E., and Sherman, J. A. (1976). Fennema-sherman mathematics attitudes 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

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

Frenzel, A. C., Goetz, T., Pekrun, R., and Watt, H. M. (2010). Development of mathematics interest in adolescence: influences of gender, family, and school context. *J. Res. Adolesc* 20, 507–537. doi: 10.1111/j.1532-7795

Frenzel, A. C., Pekrun, R., and Goetz, T. (2007). Perceived learning environment and students' emotional experiences: a multilevel analysis of mathematics classrooms. *Learn. Instr.* 17, 478–493. doi: 10.1016/j.learninstruc.2007.09.001

Ganley, C. M., and Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: examining gender patterns and reciprocal relations. *Learn. Individ. Differ.* 47, 182–193. doi: 10.1016/j.lindif.2016.01.002

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., and Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychol. Sci.* 24, 2079–2287. doi: 10.1177/0956797613486989

Goetz, T., and Hall, N. C. (2013). "Emotion and achievement in the classroom," in *International Guide to Student Achievement*, eds J. Hattie, and E. M. Anderman, (New York, NY: Routledge), 192–195.

González-Pienda, J. A., Fernández–Cueli, M., García, T., Suárez, N., Fernández, E., Tuero–Herrero, E., et al. (2012). Diferencias de género en actitudes hacia las matemáticas en la enseñanza obligatoria. *Rev. Iberoam. Psicol. Salud.* 3, 55–73.

Grigg, S., Perera, H. N., McIlveen, P., and Svetleff, Z. (2018). Relations among math self-efficacy, interest, intentions, and achievement: a social cognitive perspective. *Contemp. Educ. Psychol.* 53, 73–86. doi: 10.1016/j.cedpsych.2018.01.007

Griggs, M. S., Rimm-Kaufman, S. E., Merritt, E. G., and Patton, C. L. (2013). The responsive classroom approach and fifth grade students' math and science anxiety and self-efficacy. *Sch. Psychol. Q.* 28, 360–373. doi: 10.1037/spq0000026

Guo, J., Parker, P. D., Marsh, H. W., and Morin, A. J. (2015). Achievement, motivation, and educational choices: a longitudinal study of expectancy and value using a multiplicative perspective. *Dev. Psychol.* 51, 1163–1176. doi: 10.1037/a0039440

Hedges, L. V., and Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269, 41–45. doi: 10.1126/science.7604277

Hidi, S., and Renninger, K. A. (2006). The four-phase model of interest development. *Educ. Psychol.* 41, 111–127. doi: 10.1207/s15326985ep4102_4

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

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

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., and Williams, C. C. (2008). Gender similarities characterize math performance. *Science* 321, 494–495. doi: 10.1126/science.1160364

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

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

Jansen, M., Lüdtke, O., and Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: examining between-person and within-person variation in five domains. *Contemp. Educ. Psychol.* 46, 116–127. doi: 10.1016/j.cedpsych.2016.05.004

Kanny, M. A., Sax, L. J., and Riggers-Piehl, T. A. (2014). Investigating forty years of STEM research: how explanations for the gender gap have evolved over time. *J. Women Minor. Sci. Eng.* 20, 127–148. doi: 10.1615/JWomenMinorScienEng.2014007246

Köller, O., Baumert, J., and Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *J. Res. Math. Educ.* 32, 448–470. doi: 10.2307/749801

Kurtz-Costes, B., Rowley, S. J., Harris-Britt, A., and Woods, T. A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *MerrillPalmer. Quarterl.* 54, 386–409. doi: 10.1007/s10212-014-0216-7

Larivière, V., Ni, C., Gingras, Y., Cronin, B., and Sugimoto, C. R. (2013). Bibliometrics: global gender disparities in science. *Nat. News* 504, 211–213. doi: 10.1038/504211a

Lee, W., Lee, M. J., and Bong, M. (2014). Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemp. Educ. Psychol.* 39, 86–99. doi: 10.1016/j.cedpsych.2014.02.002

Legewie, J., and DiPrete, T. A. (2014). The high school environment and the gender gap in science and engineering. *Soc. Educ.* 87, 259–280. doi: 10.1177/0038040714547770

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

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

Louis, R. A., and Mistele, J. M. (2012). The differences in scores and self-efficacy by student gender in mathematics and science. *Int. J. Sci. Math. Educ.* 10, 1163–1190. doi: 10.1007/s10763-011-9325-9

Lyons, I. M., and Beilock, S. L. (2012). When math hurts: math anxiety predicts pain network activation in anticipation of doing math. *PLos One* 7:e48076. doi: 10.1371/journal.pone.0048076

Ministerio de Educación Cultura y Deporte [MECD], (2016). *PISA 2015. Programa Para la Evaluación Internacional de los Alumnos. Informe Español*. Madrid: Ministerio de Educación, Cultura y Deporte. .

Mullis, I. V., Martin, M. O., Foy, P., and Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Amsterdam: International Association for the Evaluation of Educational Achievement.

Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., and Eccles, J. (2008). "Gendered high school course selection as a precursor of gendered occupational careers: The mediating role of self-concept and intrinsic value," in *Gendered Occupational Outcomes: Longitudinal Assessments of Individual, Social, and Cultural Influences*, eds H. M. G. Watt, and J. S. Eccles, (Washington, DC: American Psychological Association), 115–143. doi: 10.1037/11706-004

Pajares, F., and Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemp. Educ. Psychol.* 24, 124–139. doi: 10.1006/ceps.1998.0991

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., and Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educ. Psychol.* 34, 29–48. doi: 10.1080/01443410.2013.797339

Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341. doi: 10.1007/s10648-006-9029-9

Pekrun, R., Hall, N. C., Goetz, T., and Perry, R. P. (2014). Boredom and academic achievement: testing a model of reciprocal causation. *J. Educ. Psychol.* 106, 696–710. doi: 10.1037/a0036006

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., and Goetz, T. (2017). Achievement emotions and academic performance: longitudinal models of reciprocal effects. *Child. Dev.* 88, 1653–1670. doi: 10.1111/cdev.12704

Pekrun, R., and Linnenbrink-Garcia, L. (eds) (2014). *International Handbook of Emotions in Education*. New York, NY: Taylor & Francis.

Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., and Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *Br. J. Educ. Psychol.* 84, 152–174. doi: 10.1111/bjep.12028

Porter, L. S., Marco, C. A., Schwartz, J. E., Neale, J. M., Shiffman, S., and Stone, A. A. (2000). Gender differences in coping: a comparison of trait and momentary assessments. *J. Soc. Clin. Psychol.* 19, 480–498. doi: 10.1521/jscp.2000.19.4.480

Preckel, F., Goetz, T., Pekrun, R., and Kleine, M. (2008). Gender differences in gifted and average-ability students: comparing girls' and boys' achievement, self-concept, interest, and motivation in mathematics. *Gift. Child. Q.* 52, 146–159. doi: 10.1177/0016986208315834

Putwain, D. W., Becker, S., Symes, W., and Pekrun, R. (2018). Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. *Learn. Instr.* 54, 73–81. doi: 10.1016/j.learninstruc.2017.08.004

Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., and Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: the role of problem solving strategies. *J. Exp. Child. Psychol.* 141, 83–100. doi: 10.1016/j.jecp.2015.07.014

Randhawa, B. S., Beamer, J. E., and Lundberg, I. (1993). Role of mathematics self-efficacy in the structural model of mathematics achievement. *J. Educ. Psychol.* 85, 41–48. doi: 10.1037/0022-0663.85.1.41

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

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

Silva, E. H. (2005). *Actitudes Hacia el Aprendizaje de las Matemáticas*.Tesis Doctoral no publicada, Universidad de Oviedo, Oviedo.

Skaalvik, E. M., and Skaalvik, S. (2004).). Self-concept and self-efficacy: A test of the internal/external frame of reference model and predictions of subsequent motivation and achievement. *Psychol. Rep.* 95, 1187–1202. doi: 10.2466/pr0.95.3f.1187-1202

Stevens, T., Olivarez, A., and Hamman, D. (2006). The role of cognition, motivation, and emotion in explaining the mathematics achievement gap between Hispanic and white students. *Hispanic J. Behav. Sci.* 28, 161–186. doi: 10.1177/0739986305286103

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

Valentine, J., DuBois, D., and Cooper, H. (2004). The relation between self-beliefs and academic achievement: a meta-analytic review. *Educ. Psychol.* 39, 111–133. doi: 10.1207/s15326985ep3902_3

Viljaranta, J., Lerkkanen, M. K., Poikkeus, A. M., Aunola, K., and Nurmi, J. E. (2009). Cross-lagged relations between task motivation and performance in arithmetic and literacy in kindergarten. *Learn. Instr.* 19, 335–344. doi: 10.1016/j.learninstruc.2008.06.011

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

Wigfield, A., Eccles, J., Mac Iver, D., Reuman, D., and Midgley, C. (1991). Transitions at early adolescence: changes in children's domain-specific self-perceptions and general selfesteem across the transition to junior high school. *Dev. Psychol.* 27, 552–565. doi: 10.1037/0012-1649.27.4.552

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

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A., Freedman-Doan, K., et al. (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: a three-year study. *J. Educ. Psychol.* 89, 451–469. doi: 10.1037/0022-0663.89.3.451

Williams, T., and Williams, K. (2010). Self-efficacy and performance in mathematics: reciprocal determinism in 33 nations. *J. Educ. Psychol.* 102, 453–466. doi: 10.1037/a0017271

Yoon, K. S., Eccles, J. S., and Wigfield, A. (1996). "Self-concept of ability, value, and academic achievement: a test of causal relations," in *Annual Meeting of the American Educational Research Association*, (New York, NY).

Yüksel-şahin, F. (2008). Mathematics anxiety among 4th and 5th grade Turkish elementary school students. *Int. Electron. J. Math. Educ.* 3, 179–192.

Zimmerman, B. J. (1995). "Self-efficacy and educational development," in *Self-Efficacy in Changing Societies*, ed A. Bandura, (New York, NY: Cambridge University Press), 202–231. doi: 10.1017/CBO9780511527692.009

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.

Copyright © 2020 Rodríguez, Regueiro, Piñeiro, Estévez and Valle. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 04 February 2020
doi: 10.3389/fpsyg.2020.00127

University Entrance Examinations in Spain: Using the Construct Comparability Approach to Analyze Standards Quality

Alejandro Veas, Leandro Navas, Teresa Pozo-Rico and Pablo Miñano*

Department of Developmental Psychology and Didactics, University of Alicante, Alicante, Spain

In recent years, important methodological attempts have been made to explore the comparability of examination standards, especially in the context of certifications and university entrance. The present study aimed to explore the use of a construct comparability approach through a comparative analysis of the academic scores on 15 subjects from Spanish University Entrance Examinations in the Valencian Community, with a sample of 22,996 students in the call of June 2018. We employed the Rasch partial credit model as an estimation method, counting each subject as the item of an instrument related to academic achievement. The results confirmed the unidimensionality assumption and the goodness of fit of the model in relation to all subjects, although no discrimination between high and low ability students was detected because of the lack of monotonicity of the score categories. We observed that the level of difficulty of the subjects was appropriate to the students' ability levels. Important conclusions have been drawn for the improvement of the standard qualification process, and future research directions have been proposed.

Keywords: University Entrance Examinations, academic achievement, construct comparability approach, rasch partial credit model, higher education

OPEN ACCESS

Edited by:
José Carlos Núñez,
Universidad de Oviedo Mieres, Spain

Reviewed by:
Purya Baghaei,
Islamic Azad University, Iran
María del Carmen Pérez-Fuentes,
University of Almería, Spain

****Correspondence:***
Alejandro Veas
alejandro.veas@ua.es

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

Received: *31 October 2019*
Accepted: *16 January 2020*
Published: *04 February 2020*

Citation:
Veas A, Navas L, Pozo-Rico T and Miñano P (2020) University Entrance Examinations in Spain: Using the Construct Comparability Approach to Analyze Standards Quality. Front. Psychol. 11:127. doi: 10.3389/fpsyg.2020.00127

INTRODUCTION

Across different countries, standard examinations constitute a formal procedure to select high school students based on academic achievement in different courses. This type of procedure has served as a governance instrument to provide consistent required standards of achievement, objective examination conditions, and grading procedures (Neumann et al., 2011).

The use of improved measures of academic achievement can be considered a positive consequence of the desire to increase economic growth and competitiveness (Sahlberg, 2006). Moreover, there has been a notable research interest in understanding how students' achievement can be improved with analysis of the cognitive, motivational, and contextual variables involved in causal or predictive models (Valle et al., 2008; Dicke et al., 2018). For these reasons, it is also relevant to study how different types of examinations (tests or written exams) use the required psychometric properties according to specific goals determined by educational administrations (Raykov and Pohl, 2013). This article aims to explore the measurement quality of the 2018 University Entrance Examinations in the Spanish territory of the Valencian Community, based on the construct comparability approach (Coe, 2008).

The University Entrance Examinations in Spain

In Spain, the University Entrance Examinations (known as *PAU*) are formal procedures for access to higher education, undertaken by those who have previously obtained the Spanish Baccalaureate certificate (*Bachillerato*); these are based on examination standards of mandatory and modality subjects that have been studied during the previous course. Depending on the subjects, examinations have different formats, such as essays (e.g., History of Spain and Spanish Language), analyzing visual images (Art History), texts on a specific topic of large or short extension (e.g., English Language and Latin Language), or solving problems (e.g., Economy), among others. Moreover, it must be noted that there is no unique examination for the whole country; rather, each community is autonomous and has the objective to design specific examinations for the students living within that community. The mean grade obtained from these examinations is weighted with the Baccalaureate grade, and a final evaluation is obtained to access the chosen undergraduate degree that may be located in any part of the country.

As the Spanish University Entrance Examinations are crucial for the future of thousands of students every year, it is necessary to consider the role of research assessment in the field of education. Within this context, the analysis of the process and the results obtained, as well as the employment of the distinct procedures, are relevant to ensure equality and equity of opportunities in higher education access.

In quantitative research, statistical methods have been applied to investigate the necessary conditions for measuring academic achievement objectively, with the correct design and use of measurement instruments – for example, value added models and multilevel models from a longitudinal perspective (Blanco et al., 2009; López-Martín et al., 2014).

With respect to the PAU, important research was conducted by Gaviria (2005), where different statistical techniques were applied – classic, ordinary least squares, multilevel, and mean and standard deviation equality methods – to match the grade obtained in the Baccalaureate with that obtained in the PAU; the last served as anchor, as the examination was the same for all participants. The results showed that the non-classical method produced worse results than classical methods, improving justice in student selection. Apart from this study, no other relevant research is found beyond quantitative analyses, which refer to group differences in a specific context (Ruiz et al., 2013). For this reason, this study aimed to fulfill the existing limitations by adding new comparability analyses of standard examinations.

Advances and Limitations of Standard Examinations Comparability

Traditionally, standardized achievement tests are considered the most objective procedure, as they reflect a unidimensional construct that is highly dependent upon students' cognitive abilities (Hübner et al., 2019). Different international organizations have clarified the improvements in the design and implementation of international standardized tests such as TIMMS (*Trends in International Mathematics and Science Study*), PIRLS (*Progress in International Reading Literacy Study*), IALS (*International Assessment of Literacy Survey*), and especially PISA (*Programme for International Student Assessment*).

On the other hand, written examinations based on grades are considered a multidimensional construct, in which teachers use different criteria (Guskey, 2006). In this area, multiple studies have claimed the impact of various frames of reference. For example, Westphal et al. (2016) found that teachers' judgments were associated with the socioeconomic composition of the classroom in a sample of 3,285 math fourth graders. Zimmermann et al. (2013) showed, in a longitudinal study of 1,045 students from Grade 5 to Grade 9, that external problems are reflected in teacher-given grades more than in standardized achievement tests.

Given the possible factors associated with grading, there have been several attempts to improve objective grading criteria in Europe, as written examinations are crucial in educational systems, especially for obtaining institutional certificates, or selecting students for higher education (Newton, 2005, 2010). For instance, the implementation of the central Abitur examination in Germany is remarkable (Kuehn, 2012). Although these examinations present differences in procedures or subjects between each German state, their higher level of standardization means that these grades are less affected by factors related to the schools or the teachers (Neumann et al., 2011). On the other hand, important methodological advances have been implemented in England for examinations used in academic qualifications such as the General Certificate of Secondary Education (GCSE, taken by students aged 16) and General Certificate of Education Advanced level (GCE A level, taken by students aged 18) (Coe, 2008; Newton et al., 2007). In this context, special attention has been given to inter-subject comparability using a variety of statistical procedures, including pair analyses, common examinee linear models, and item response theory models (Coe et al., 2008).

Inter-subject comparability of examination standards constitutes an educational need to apply statistical aligns when grades from different subjects are used for specific objectives. When this is possible, academic achievement can be measured as the level of an individual's skill in a specific examination of a certain difficulty. In the context of comparing academic grades, it is also important to notice that we can only compare those measuring a shared construct. For this reason, the concept *construct comparability approach* constitutes a formal theoretical framework in which statistical applications are applied (Coe, 2010).

Use of the Rasch Model Within the Construct Comparability Approach

Different authors have developed advanced psychometric analyses for the comparison of subject examinations. In this context, more specifically, the Rasch measurement model was chosen as the most appropriate, given the theoretical framework and the complexity of data. Coe (2008) implemented it in a sample of nearly 6,000 candidates who took GCSEs in 2004, including the exploration of Differential Item

Functioning (DIF). Recently, He et al. (2018) also applied the Rasch analysis to both the GCSE and GCE A levels over a period of 4 years, in order to establish the consistency of difficulty parameters and grade comparison in the same country. Other countries, such as Tasmania, have approved educational policies based on the formal application of the Rasch model to the alignment of statistical standards (Tasmanian Qualification Authority [TQA], 2006, 2007).

The Rasch (1980) model is regarded as the most renowned of IRT models, providing a method based on the calibration of ordinal data from a shared measurement scale and enabling one to test conditions such as dimensionality, linearity, and monotonicity. This model analyses the difficulty of items and individuals' ability on the same scale, employing a logarithmic function to test the probability of a subject to correctly respond to an item. Use of the same measurement scale established homogeneous intervals, meaning that the same difference between the difficulty parameter of an item and the ability of a subject involves equal probability of success along the entire scale (Preece, 2002).

According to comparability criteria, we started by considering each of the courses as a specific item, with a range of grades from 1 to 10, which implies various degrees or categories of success. In this case, the partial credit model (PCM) (Wright and Masters, 1982) enabled the analysis of the difficulty in achieving a specific score for each of the subjects separately, following the Rasch methodology. In this study, the use of PCM is justified in the fact that, in Spain, the same grades obtained in different examinations are not necessarily related to the same level of effort (He et al., 2018).

The formula of the model is as follows:

$$\mathrm{Ln}(P_{nij}/P_{ni(j-1)}) = B_n - D_i F_{ij} = B_n - D_{ij}$$

where:

P_{nij} is the probability of subject n responding correctly to item i observed in category j;

B_n is the measured ability of subject n;

D_i is the measured difficulty of item i; and

F_{ij} is the calibration measured for item i in category j compared to category j-1, the point at which categories j-1 and j are equally likely compared to the measurement of the item (Masters, 1982).

The Present Study

The use of the Rasch model for analyzing inter-subject comparability has been employed in different countries (Coe, 2008; Korobko et al., 2008). Based on the literature review, the present study aimed to apply the Rasch PCM in the Spanish University Entrance Examinations taken in the Valencian Community, according to the construct comparability approach, which was developed in England over the last decades. Concretely, three main objectives were followed, specifically (1) to analyze the unidimensionality of the measures; (2) to compare the fit statistics and difficulty parameters between the different subjects, and (3) to compare the distribution of difficulty level of the subject grades along the latent trait. Given that no previous IRT analysis has been conducted on these examinations in Spain, there are no directional hypotheses to be determined.

MATERIALS AND METHODS

Sample

The sample was taken from all students in the Valencian Community that participated in the Spanish University Entrance Examinations in the last call of June 2018. The community is located on the east coast of the country and comprises three provinces: Alicante, Valencia, and Castellón. A total of 22,996 students were considered: 10,015 students took the exam in the province of Alicante (43.55% of the total sample), 2248 students in the province of Castellón (9.77% of the total sample), and 10,733 students in the province of Valencia (46.77%). For each province, examinations were taken in different public universities or venues belonging to these universities (extension areas where specific degrees are taught). Approximately 60% were females.

Measures

The Spanish University Entrance Examinations from the call of June 2018 were considered for further analysis. A total of 24 subjects (described in the section "Results") were first considered, accounting for both mandatory and modality subjects. All the examinations have correction standards previously approved by the qualification board. In this sense, corrections criteria are defined and given top scores for each specific question in each exam, together with a qualitative instruction that helps examiners ensure objectivity. For all the examinations, the lowest score is 0 and the highest is 10, with the sum of the grade obtained in each question based on raters' assignments. These qualification criteria are public and available on the website of the Valencian Community Government (2019).

Procedure

Necessary permission was first obtained by the University Regulation Service, an institution belonging to the Valencian Community Government; it provided the grades from all students enrolled in the University Entrance Examinations in the three provinces of the Valencian Community – Alicante, Valencia, and Castellón – at the call of June and July 2018. For the present research, data from June 2018 were taken for the analysis. This study was approved by the Institutional Review Board and complied with the Ethical standards of the 1964 Helsinki Declaration and its later amendments, or comparable ethical standards.

Data Analysis

For the present study, the construct comparability approach was applied based on the assumption that it is possible to compare the qualifications obtained by the students for the subjects involved in the higher education selection process. The software Winsteps version 4.4.0 (Linacre, 2019) was employed to implement Rasch PCM, where a joint maximum likelihood estimation was realized. In this model, each of the included subjects was considered an

item of the same instrument that contributes to the measurement of the construct academic achievement.

First, and according to the Rasch assumptions, unidimensionality was tested with a principal component analysis of residuals. According to Linacre (1998, 2002), the eigenvalues obtained for each contrast comparison should be no more than 2. Moreover, the estimation process of the item difficulty parameters (including their respective categories) and individuals' abilities is iterative, by examining the relation with the probability of obtaining a specific score according to the individual's ability. With this procedure, it is possible to obtain a value that better explains the achievement pattern registered. Simultaneously, it is possible to obtain the ability value for each individual according to the item difficulty pattern. This process was repeated by using the estimations of ability and difficulty until the iteration converged.

In the Rasch analysis, two basic fit statistics are employed: infit and outfit. These are calculated based on room mean squares, depending on the statistical value of Pearson's chi-squared divided by the degrees of freedom, thus forming a scale with values ranging from 0 to infinity. Values below 1 indicate a higher than expected fit of the model, while values greater than 1 indicate a poor fit. Linacre (2002) suggested that those with values higher than 2 imply a bad fit to the model, making the conclusion of a reliable analysis impossible. For this reason, the authors of the present study used this value as a formal cut-off, both in items and subjects which, according to previous research, are also within the construct comparability approach (He et al., 2018). Moreover, the mean of individuals' ability was set to 0 for the different subjects, as to allow the comparison of parameters estimations.

RESULTS

Before the implementation of the Rasch analysis, descriptive statistics of all subjects and participants for each province were observed. As seen in **Table 1**, the mean values are mostly located between 6 and 7.9, which is considered positive in terms of certification aptitude. Some exceptions are Geography and Greek Language, both from the humanities field, with 5.6 and 5.4, respectively.

It may also be noted that the number of participants presented an imbalance, due to the fact that students have to choose specific examinations. For example, the majority of students chose English Language, as it is mandatory for all educational centers. However, other languages, such as German or Italian, are not mandatory and are offered only by a few educational centers.

As the number of participants may affect the calculation parameter accuracy, there was a final selection that included those subjects with at least 1500 participants. For the Rasch

TABLE 1 | Examination grades by province: means and standard deviations.

	Alicante		Valencia		Castellón		Total participants	Total mean
Subject	**Participants**	**Mean (SD)**	**Participants**	**Mean (SD)**	**Participants**	**Mean (SD)**		
German language	1	7 (–)	8	7.63 (1.57)	1	9.3 (–)	10	7.74
Scenic arts	53	7.38 (1.47)	87	7.13 (1.34)	18	6.95	158	7.19
Biology	1745	6.19 (2.18)	2652	6.43 (2.15)	526	5.77 (2.34)	4923	6.28
Spanish language	6123	6.42 (1.65)	9510	6.14 (1.68)	2059	6.34 (1.66)	17692	6.26
Audiovisual culture	217	7.12 (1.56)	388	6.95 (1.67)	43	6.99 (1.26)	648	7.01
Design	187	7.08 (1.43)	328	6.27 (1.55)	45	6.59 (1.15)	560	6.56
Technical drawing	697	6.18 (2.34)	1348	6.50 (2.41)	218	6.83 (2.14)	2263	6.32
Economy	1698	6.47 (2.05)	3012	6.69 (1.91)	617	6.85 (1.87)	5327	6.64
Fundamentals of arts	249	6.78 (2.3)	392	6.78 (2.09)	70	6.11 (2.04)	711	6.71
Physics	1505	6.31 (2.41)	2312	6.35 (2.42)	457	6.43 (2.26)	4274	6.35
French language	161	7.43 (1.51)	231	8.14 (1.30)	73	8.42 (1.17)	465	7.94
Geology	94	5.8 (1.85)	69	5.18 (2.15)	14	6.33 (2.19)	177	5.6
Geography	1622	5.15 (2.41)	2882	5.57 (2.42)	600	5.02 (2.36)	5104	5.4
Greek language	539	6.62 (2.48)	730	6.39 (2.21)	155	6.67 (1.94)	1424	6.5
History of arts	715	5.79 (2.42)	926	6.10 (2.21)	181	6.35 (2.14)	1822	6
History of Spain	6124	7.03 (1.68)	9515	6.89 (1.66)	2060	6.97 (1.47)	17699	6.95
History of philosophy	780	6.38 (2.30)	1155	6.38 (2.16)	226	6.75 (2.33)	2161	6.42
English language	5960	6.71 (1.97)	9267	6.76 (1.92)	1986	6.86 (1.89)	17213	6.76
Italian language	4	7.4 (1.24)	4	9.06 (0.89)	0	–	8	8.23
Latin language	864	6.42 (2.18)	1098	6.31 (2.08)	249	6.91 (1.89)	1611	6.42
Mathematics	3180	7.17 (2.12)	5036	7.21 (2.16)	1074	7.40 (2.09)	9290	7.22
Applied mathematics to social sciences	3180	5.47 (2.28)	3833	5.77 (2.25)	846	5.92 (2.06)	7859	5.69
Chemistry	2182	5.70 (2.30)	3383	5.5 (2.31)	675	6.12 (2.22)	6240	5.77
Valencian language	4747	6.47 (1.54)	9048	6.38 (1.52)	1964	6.42 (1.63)	15759	6.42

N, number of participants; M, mean; SD, standard deviation.

PCM analysis, the subjects of German Language, Scenic Arts, Audiovisual Culture, Design, Fundamentals of Arts, French Language, Geology, Greek Language, and Italian Language were removed to improve accuracy.

The Rasch PCM analysis showed the summary statistics, including person reliability and separation indexes of 0.74 and 1.69, respectively. These values can be considered low, which indicates that the group of subjects was not sensitive enough to appropriately distinguish students with high and low achievement (Bond and Fox, 2007).

With respect to the unidimensionality of the model based on principal component analysis of residual scores, the results show a principal factor that explains 51.3% of the variance of the latent trait. With respect to a hypothetical second factor, it shows a value lower than 2 (Eigenvalue V2 = 1.4), which confirms the unidimensionality of the model.

In **Table 2**, examinations are ordered by their difficulty parameter (from high to low), together with their respective fit indexes. An optimal fit can be observed according to the established criteria. The examinations with a higher difficulty level were Chemistry, Geography, and Physics, whereas those with a lower difficulty level were Mathematics, History of Spain, and Economy.

Within the PCM framework, **Table 3** shows the average of the category parameters that are used to estimate fit statistics, showing nearly perfect infit and outfit values. Moreover, the observed average of the measures – a description of the sample expected to increase with category value, as in this case – is computed and modeled to produce the responses observed in the category. Andrich Thresholds (also called step difficulty, step calibration or Rasch-Andrich threshold) are based on the calibrated measure of the transition from one category below to another adjacent category – the point on the latent variable at which adjacent categories are equally probable to be observed. For this reason, it indicates the difficulty to observe a specific category and not the difficulty to respond to this category (Linacre, 2019, p. 532). Step calibrations show that category ordering is interrupted only in the pair categories 3 and 4 (−0.86). However, this value is strictly influenced by the distribution of frequencies of observations in each category. As the average measures of the persons advance across categories, it can be assumed that the categories support monotonicity (Linacre, 2019, p. 532).

TABLE 2 | Difficulty parameters and fit statistics of the University Entrance Examinations.

Examinations	Difficulty	Infit	Outfit
Chemistry	−0.09	0.94	0.93
Geography	−0.23	1.43	1.42
Physics	−0.30	1.22	1.21
Applied mathematics to social sciences	−0.33	1.25	1.24
Biology	−0.35	0.92	0.90
Technical drawing	−0.41	1.58	1.56
Spanish language	−0.45	0.70	0.72
Art history	−0.47	1.24	1.23
Valencian language	−0.52	0.70	0.72
Latin language	−0.60	1.16	1.14
History of philosophy	−0.64	1.29	1.25
English language	−0.68	1.15	1.14
Economy	−0.73	0.98	0.97
History of Spain	−0.78	0.89	0.91
Mathematics	−0.83	1.28	1.22

TABLE 3 | Summary of category structure.

Score	Observed count	Observed average	Infit	Outfit	Andrich threshold	Category measure
0	411	−0.54	1.35	1.37	–	−3.14
1	1037	−0.50	1.02	1.04	−1.62	−1.90
2	3143	−0.29	1.01	1.01	−1.49	−1.25
3	4889	−0.14	0.98	0.99	−0.65	−0.83
4	10994	0.03	0.99	0.99	−0.86	−0.46
5	15014	0.19	0.93	0.91	−0.19	−0.10
6	23062	0.40	0.97	0.96	−0.13	0.29
7	19242	0.63	0.96	0.96	0.69	0.73
8	22196	0.91	0.97	0.97	0.61	1.28
9	12055	1.22	1.02	1.01	1.66	2.07
10	6855	1.57	1.10	1.06	1.98	3.44

Figure 1 shows the "Wright map," where persons and items are distributed along the ability and difficulty range, respectively. Persons are located on the left side of the graph, whereas examinations are located on the right side. It is noted that the difficulty of the examinations corresponded to persons' abilities between log its 0 and −1. This may be positive, as most persons had sufficient ability to take the examinations. However, it also means that these examinations cannot accurately differentiate persons located at the top of the ability continuum.

DISCUSSION

This study aimed to analyze an empirical estimation of the qualifications obtained in the Spanish University Entrance Examinations with the application of the Rasch PCM, following the theoretical framework of the construct comparability approach. It is assumed that the measurement of academic achievement is a latent construct, allowing the comparison of difficulty parameters obtained for each of the standard scores for the corresponding subjects. This model has been considered useful in the assessment field for access to higher education (Tognolini and Andrich, 1996). The measurement system produced has been employed in different certificate examinations in many countries, including Tasmania (Tasmanian Qualification Authority [TQA], 2007) and England (Coe, 2008).

Following the first objective, the analyses showed the accomplishment of criteria of unidimensionality, which is essential for the application of the Rasch model, and the possibility of using a defined latent construct, namely academic achievement, in the PAU context. However, it must be noted that the establishment of this operative construct cannot be interpreted as the existence of a unique global process. The scientific literature posits that the interpretation of this construct

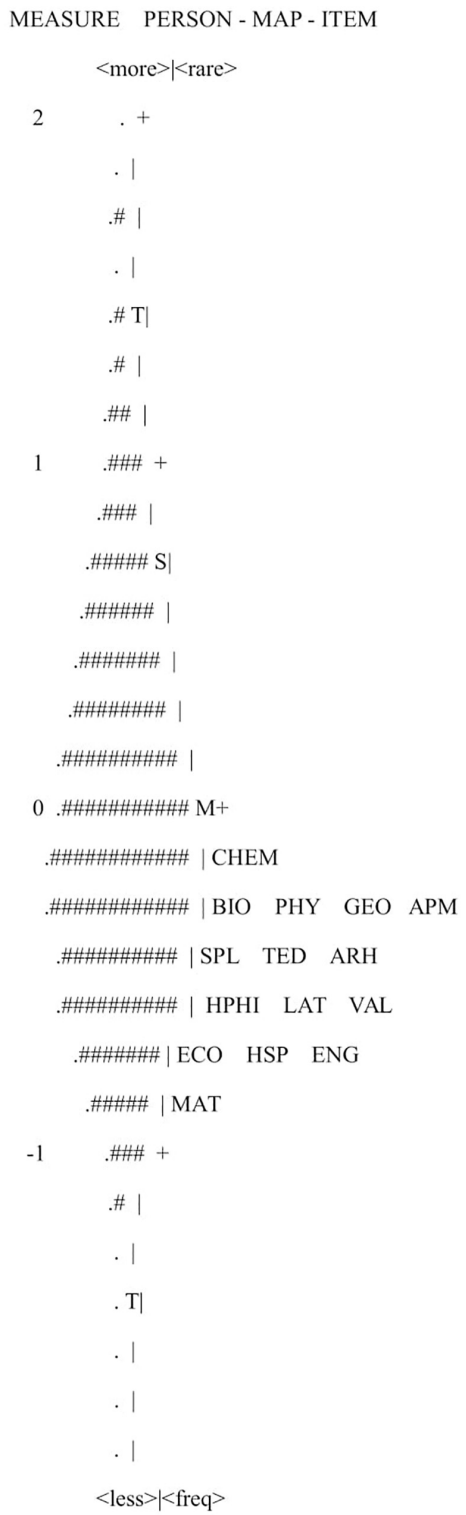


FIGURE 1 | Item-Person map. EACH "#" IS 146: EACH "." IS 1 TO 145. CHEM, chemistry; BIO, biology; PHY, physics; GEO, geography; APM, applied mathematics to social sciences; SPL, spanish language; TED, technical drawing; ARH, art history; HF, history of philosophy; LAT, lating language; VAL, valencian language; ECO, economy; HSP, history of Spain; ENG, english language; MAT, mathematics.

is not clear, as it does not serve as the basis for the specific purpose of each examination (He et al., 2018). For this reason, it is argued that, although all examinations require specific abilities, they also demand global cognitive processes related to the construct measurement.

With respect to the second objective, an optimal fit was observed in all the examinations, which led to considering the invariance properties assumed by the Rasch model in terms of person and item (or examination) comparison along the same latent construct (Bond and Fox, 2007). Therefore, the consequences of this type of estimation lie in the possibility of making inferences beyond the students' sample employed. At the same time, the examination fit allowed a comparison among them in terms of the difficulty parameters obtained together with the ability levels required to attain each possible score. From these results, a key concept was formed in the context of PAU – the selection of examinations, a topic widely discussed in international literature (Lamprianou, 2009). Bell et al. (2007) indicated that the perceived difficulty of a student in one or various examinations could be an obstacle to university entrance; as a consequence, other subjects might be favored with a higher enrollment fee. Taking into account the results of the present study, this may be happening with the subject of History of Spain to the detriment of History of Philosophy, as the students have to choose one and the number of candidates in the former is three times higher than the latter.

The analysis of the third objective highlights the need to consider the qualification scale employed in PAU as typical. Disorder rating category is observed between grades 3 and 4, which means that the 10-point category does not discriminate in some points of the latent trait. However, fit values were good for all categories, and the observed average of the measures increased with category values. It must be mentioned that the majority of countries that use comparative analysis employ a minor number of qualification categories. In this case, a smaller sample size may interfere with Andrich Threshold estimations. For this reason, in order to make similar estimations, future studies should analyze the general category structure in all Spanish communities that conclude general psychometric strategies.

Finally, the person separation index is low, showing that these examinations do not accurately differentiate students with high and low achievement. However, the Wright Map indicated that the difficulty levels of all examinations are within the students' ability range; therefore, there are adequate probability levels of obtaining positive results. The location of the examinations on the scale corresponds to a similar distribution of the categories on the latent construct. Again, these results showed the need to recodify the category system to improve the differentiation of individual levels, as a higher number of students might be included for each of the high and low categories.

CONCLUSION

In conclusion, the intention of this study was to initiate an effective analysis of standard scores comparison in Spain

under the construct comparability approach, a theoretical and methodological framework used in other countries (Newton, 2012). Limitations and future directions should be addressed. First, the samples utilized in other countries are considerably larger, as data were collected throughout the country, which provides better estimations. This study is implemented within a single Spanish community, and it confirms the potential need for future studies similar to those conducted in England. In the Spanish context, it would be essential to draw a comparison between autonomous communities in order to find the appropriate equity measurement. This possibility has not been explored in the scientific literature in this field. However, considering that the majority of Spanish examinations have a written format, the differences between examiners in the interpretation of tasks and the evaluation categories by different raters, together with other possible effects (halo effect, gender, and cultural bias), may contribute to error measurement, validity and justice in evaluation (Frederiksen, 1984; Eckes, 2015). In this context, a multi-faceted Rasch model may adequately address these issues in the future.

DATA AVAILABILITY STATEMENT

The datasets for this manuscript are not publicly available because they belong to the Valencian Government who are restrictive to their data policy. Requests to access the datasets should be directed to University Regulation Service (sru@gva.es).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Alicante. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

AV: literature review, data analysis, statistical procedures, and manuscript writing. LN: literature review, manuscript writing, and manuscript revision. TP-R: manuscript writing and manuscript revision. PM: manuscript revision.

FUNDING

This work was supported by the Vice Chancellor for Research of the University of Alicante (Award number: GRE17-16).

REFERENCES

Bell, J. F., Malacova, E., Vidal-Rodeiro, L., and Shannon, M. (2007). *A-Level Uptake: Crunchier Subjects and the Cracker Effect*. Cambridge, MA: A Cambridge Assessment publication.

Blanco, A., González, C., and Ordóñez, X. G. (2009). Patrones de correlación entre medidas de rendimiento escolar en evaluaciones longitudinales: un estudio de simulación desde un enfoque multinivel (correlation patterns of academic achievement measures in longitudinal assessment: a simulation study from a multilevel perspective). *Rev. Educ.* 348, 195–215.

Bond, T. G., and Fox, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. New York, NY: Psychology Press.

Coe, R. (2008). Comparability of GCSE examinations in different subjects: an application of the Rasch model. *Oxf. Rev. Educ.* 34, 609–636. doi: 10.1080/03054980801970312

Coe, R. (2010). Understanding comparability of examination standards. *Res. Pap. Educ.* 25, 271–284. doi: 10.1080/02671522.2010.49813

Coe, R., Searle, J., Barmby, P., Jones, K., and Higgins, S. (2008). *Relative Difficulty of Examinations in Different Subjects (Report from SCORE-Science Community Supporting Education)*. Durham: Durham University.

Dicke, T., Marsh, H. W., Parker, P. D., Pekrun, R., Guo, J., and Televantou, I. (2018). Effects of school-average achievement on individual self-concept and achievement: unmasking phantom effects masquerading as true compositional effects. *J. Educ. Psychol.* 110, 1112–1126. doi: 10.1037/edu0000259

Eckes, T. (2015). *Introduction to Many-Facet Rasch Measurement*, 2nd Edn. Frankfurt: Peter Lang Edition.

Frederiksen, N. (1984). The real test bias: influences of testing on teaching and learning. *Am. Psychol.* 39, 193–202. doi: 10.1037/0003-066X.39.3.193

Gaviria, J. L. (2005). Equalising baccalaureate records in the university entrance selection process. *Revista de Educación* 337, 351–387.

Guskey, T. R. (2006). Making high school grades meaningful. *Phi Delta Kappan* 87, 670–675. doi: 10.1177/003172170608700910

He, Q., Stockford, I., and Meadows, M. (2018). Inter-subject comparability of examination standards in GCSE and GCE in England. *Oxf. Rev. Educ.* 44, 494–513. doi: 10.1080/03054985.2018.1430562

Hübner, N., Wagner, W., Hochweber, J., Neumann, M., and Nagengast, B. (2019). Comparing apples and oranges: curricular intensification reforms can change the meaning of students' grades! *J. Educ. Psychol.* 112, 204–220. doi: 10.1037/edu0000351

Korobko, O. B., Glas, C. A. W., Bosker, R. J., and Luyten, J. W. (2008). Comparing the difficulty of examination subjects with item response theory. *J. Educ. Meas.* 45, 139–157. doi: 10.1111/j.1745-3984.2007.00057.x

Kuehn, S.-M. (2012). Exploring the use of statewide exit exams to spread innovation- The example of context in science tasks from an international comparative perspective. *Stud. Educ. Eval.* 37, 189–195. doi: 10.1016/j.stueduc.2012.01.003

Lamprianou, I. (2009). Comparability of examination standards between subjects: an international perspective. *Oxf. Rev. Educ.* 35, 200–226.

Linacre, J. M. (1998). Structure in Rasch residuals: why principal component analysis? *Rasch Meas. Trans.* 12:636.

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Meas. Trans.* 16:878.

Linacre, J. M. (2019). *Winsteps Rasch Measurement Computer Program [Version 4.4.0]*. Chicago, IL: Winsteps.

López-Martín, E., Kousmanen, T., and Gaviria, J. L. (2014). Linear and nonlinear growth models for value-added assessment. An application to Spanish primary and secondary schools' progress in reading comprehension. *Educ. Assess. Eval. Accountabil.* 26, 361–391. doi: 10.1007/s11092-014-9194-1

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psichometrika* 47, 149–174. doi: 10.1007/BF02296272

Neumann, M., Trautwein, U., and Nagy, G. (2011). Do central examinations lead to greater grading comparability? A study of frame-of-reference effects on the University entrance qualification in Germany. *Stud. Educ. Eval.* 37, 206–217. doi: 10.1016/j.stueduc.2012.02.002

Newton, P. E. (2005). Examination standards and the limits of linking. *Assess. Educ.* 12, 105–123. doi: 10.1080/09695940500143795

Newton, P. E. (2010). Conceptualizing comparability. *Measurement* 8, 172–179. doi: 10.1080/15366361003780956

Newton, P. E. (2012). Making sense of decades of debate on inter-subject comparability in England. *Assess. Educ.* 19, 251–273. doi: 10.1080/0969594X.2011.563357

Newton, P. E., Baird, J., Goldstein, H., Patrick, H., and Tymms, P., (eds). (2007). *Techniques for Monitoring the Comparability of Examination Standards*. London: Qualifications and Curriculum Authority.

Preece, P. (2002). Equal-interval measurement: the foundation of quantitative educational research. *Res. Pap. Educ.* 17, 363–372. doi: 10.1080/0267152022000031397

Rasch, G. (1980). *Probabilistic Models for Intelligence and Attainment Tests (Copenhagen, Danish Institute for Educational Research)*. Chicago, IL: The University of Chicago Press.

Raykov, T., and Pohl, S. (2013). Essential unidimensionality examination for multicomponent scale: an interrelationship decomposition approach. *Educ. Psychol. Meas.* 73, 581–600. doi: 10.1177/0013164412470451

Ruiz, J., Dàvila, P., Etxeberria, J., and Sarasúa, J. M. (2013). UPV-EHU University entrance testing in mathematics. Teachers' opinions and results. *Revista de Educación* 217–246. doi: 10.4438/1988-592X-RE-2011-362-159

Sahlberg, P. (2006). Education reform for raising economic competitiveness. *J. Educ. Change* 7, 259–287. doi: 10.1007/s10833-005-4884-6

Tasmanian Qualification Authority [TQA] (2006). *An Introduction to Rasch Modelling and the TCE*. Available at: https://www.tqa.tas.gov.au/4DCGI/_WWW_doc/003675/RND01/Rasch_intro.pdf (accessed May 20, 2019).

Tasmanian Qualification Authority [TQA] (2007). *How the Scaled Awards are Calculated and Used to Determine the Tertiary Entrance Score*. Available at: http://www.tqa.tas.gov.au/0477 (accessed May 20, 2014).

Tognolini, J., and Andrich, D. (1996). Analysis of profiles of students applying for entrance to universities. *Appl. Meas. Educ.* 9, 323–353. doi: 10.1207/s1532481ame0904_3

Valencian Community Government (2019). *Exámenes y Criterios de Corrección de 2018 [Examinations and correction criteria in 2018]*. Available at: http://www.ceice.gva.es/es/web/universidad/examenes-y-criterios-de-correccion-convocatoria-extraordinaria (accessed March 4, 2019).

Valle, A., Núñez, J. C., Cabanach, R. G., González-Pienda, J. A., Rodríguez, S., Rosário, P., et al. (2008). Self-regulated profiles and academic achievement. *Psicothema* 20, 724–731.

Westphal, A., Becker, M., Vock, M., Maaz, K., Neumann, M., and McElvany, N. (2016). The link between teacher-assigned grades and classroom socioeconomic composition: the role of classroom behavior, motivation, and teacher characteristics. *Contemp. Educ. Psychol.* 46, 218–227. doi: 10.1016/j.cedpsych.2016.06.004

Wright, B. D., and Masters, G. N. (1982). *Rating Scale Analysis. Rasch Measurement*. Chicago, IL: MESA Press.

Zimmermann, F., Schütte, K., Taskinen, P., and Köller, O. (2013). Reciprocal effects between adolescent externalizing problems and measures of achievement. *J. Educ. Psychol.* 105, 747–761. doi: 10.1037/a0032793

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.

Copyright © 2020 Veas, Navas, Pozo-Rico and Miñano. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 07 February 2020
doi: 10.3389/fpsyg.2019.02920

Schools That Promote the Improvement of Academic Performance and the Success of All Students

Pilar Arnaiz-Sánchez, Remedios de Haro, Salvador Alcaraz and Ana Belén Mirete Ruiz*

Department of Didactics and School Management, University of Murcia, Murcia, Spain

The movement for effective schools and school improvement has enjoyed a long history, at both the theoretical and practical level. The contextual variables focused on the educational process of the classrooms have been identified in numerous investigations, concluding that the improvement of academic performance is a key element of the movement. The main objective of this research is focused on verifying the effectiveness of the treatment based on collaborative/cooperative learning methodologies and projects to improve the linguistic and mathematical competence as an enhancing element of academic performance. The sample consists of 228 students belonging to two public schools located in the city of Murcia (Spain), selected through judgmental or discretionary sampling. A quasi-experimental design with pretest and post-test and control group was employed, verifying the effectiveness of the treatment, and how it influences the improvement of the academic performance of the students in the experimental group. It concludes by pointing out the importance of learning strategies and applied teaching methodologies, understanding both within the conglomerate of process factors in the improvement of academic performance.

Keywords: academic performance, effective schools, educational improvement, intervention program, quasi-experimental design, inclusive education

OPEN ACCESS

Edited by:
Raquel Gilar,
University of Alicante, Spain

Reviewed by:
Anabel Moriña,
University of Seville, Spain
Ángeles Parrilla Latas,
University of Vigo, Spain

****Correspondence:***
Pilar Arnaiz-Sánchez
parnaiz@um.es

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

Received: *19 October 2019*
Accepted: *10 December 2019*
Published: *07 February 2020*

Citation:
Arnaiz-Sánchez P, de Haro R, Alcaraz S and Mirete Ruiz AB (2020) Schools That Promote the Improvement of Academic Performance and the Success of All Students. Front. Psychol. 10:2920. doi: 10.3389/fpsyg.2019.02920

INTRODUCTION

An objective present in every government agenda is ensuring the development of an inclusive, equitable, and quality education that promotes lifelong learning opportunities. UNESCO (2015), in its enactment of the 2030 Agenda for Sustainable Development, proposes as one of its fundamental objectives the promotion of this type of education in order to promote success for all, as well as good levels of academic performance. To do this, schools must develop the necessary conditions and processes that realize this goal so that schools become "inclusive and effective learning environments within the framework of a school for all" (Arnaiz, 2012, p. 31).

Inclusive education is connected to the movement for effective schools and school improvement, making it clear that what happens in classrooms, in terms of organization, interventions, activities proposed by teachers, among other factors, has a critical value for the improvement of academic performance and success for all students, in short, for the attainment of expected educational achievement (Rutter et al., 1979; Brophy and Good, 1986; Mortimore et al., 1989; Davis and Thomas, 1992; Reynolds and Cuttance, 1992; Scheerens, 1992; Ramasut and Reynolds, 1993; Teddlie and Stringfield, 1993; Creemers, 1994; Reynolds et al., 1994; Castejón, 1996; Wyatt, 1996; Murillo, 2003a,b; Arnaiz et al., 2018).

An effective school is one that "achieves the comprehensive development of each and every one of its students, greater than would have been expected when considering their previous performance and the social, economic and cultural situation of their families" (Murillo, 2005, p. 25). The fundamental role of the school in improving academic performance, as considered by the movement for Effective Schools and Inclusive Educational, calls into question the decisive characteristic given to some variables in academic performance, such as the socioeconomic origin of families. An emblematic example of this was the Coleman (1968), which rejected any possibility of improvement and overcoming inequalities. Fifty years after its publication, research has shown that a school, through its interventions, procedures and actions, can promote the overcoming of inequalities and, therefore, facilitate educational success. In this line of argument, Flecha and Buslon (2016); Madigan et al. (2016), Marqués (2016); Muro et al. (2018), among others, indicate that successful educational interventions produce significant improvements in academic results, overcoming the determinative and deficient views present in said report.

Traditionally, the improvement of student academic performance has been focused on three main components: personal characteristics or individual factors (such as intelligence), contextual factors or improvement of the educational environment (such as school improvement) and factors related to one's self-beliefs, understanding of self, and the environment (such as responsibilities, mentality, and personal experiences) (Van Mieghem et al., 2018). In other works, procedural factors such as school environment or leadership have been taken into account. Along the same lines, the works of Murillo (2007) and Jornet et al. (2012) classify these factors as entry (gender, socio-economic level, mother tongue, school resources, etc.), process (study habits, academic expectations, family support, school environment, teaching methodology, etc.) and product (academic performance). In turn, these factors can be divided into two areas or levels: students and educational centers.

Other authors have identified multiple individual components, such as cognitive ability (Lu et al., 2011), self-perspective (Miñano et al., 2012; Dedrick et al., 2015), gender (Miralles et al., 2012), expectations (Zimmerman et al., 1992; Miller et al., 1993), socioeconomic status (Miralles et al., 2012), opportunities for physical activity (Takehara et al., 2019), and motivation (Castejón et al., 2016; Muro et al., 2018), as predictive factors in academic performance. Likewise, Escarbajal et al. (2019) and Pulido and Herrera (2019) identified family status, cultural origin, and age as influential variables on the issue at hand. According to these authors, students of immigrant origin obtain lower academic results in relation to students of native origin.

Ruiz-Esteban et al. (2018) relate academic performance to academic goals and motivational patterns. Similarly, Rodríguez and Guzmán (2019) highlight the influence of academic goals and the variables of environment, support, and socio-labor status of families on academic performance.

Contextual variables have also been highlighted as important predictive factors in academic performance (Jeynes, 2010; Zuffianò et al., 2013). Notable among them is the orientation to learning goals (Hsieh et al., 2007), learning strategies (Preckel and Brunner, 2015; Veas et al., 2017), popularity (Schwartz et al., 2006), adaptation to the school context (McCoach and Siegle, 2003), early attention (Franco et al., 2017), the use of methodologies such as peer mentoring (Durán and Vidal, 2004; Dunn et al., 2017), the participation of families in the school (Wilder, 2014), or the relationships established between students and teachers (Ramberg et al., 2019).

Hincapie et al. (2018) relate academic performance to teaching methodology. In this case, good results are obtained when problem-based learning is used to improve academic performance and critical thinking. Likewise, Karrera et al. (2019) highlight the importance of using teaching strategies such as project work to improve academic performance in Primary Education. Similarly, the relationship between access to, and the use of, information and communication technologies and academic performance has been studied in depth. Various studies indicate an increase in academic performance in those schools where the use of these technologies is utilized (García-Martín and Cantón-Mayo, 2019; Hinojo et al., 2019). Several authors (Molina-López et al., 2018) have even pointed out that competition between schools shows a positive effect on the academic performance of students.

Other studies show empirical evidence of the influence of some variables on academic performance such as the mean socio-economic and cultural level of the school (Perry and McConney, 2010a,b), school size or teacher–student ratio (Nath, 2012), as well as school process factors such as grouping students according to their academic ability (Meunier, 2011; Kunz, 2014), teaching methodology (Nath, 2012; Payandeh-Najafabadi et al., 2013), or the learning environment (Payandeh-Najafabadi et al., 2013; Santos et al., 2013), among others.

Given the great diversity of factors influencing the academic performance shown in the scientific literature, in this article, we focus attention on some variables – teaching methodologies – that influence the academic performance of the students participating in a particular school. To that end, the recommendation made by Miralles et al. (2012) has been followed, which urges experience studies to be carried out that allow us to understand which factors contribute to the effective functioning of schools from the point of view of student academic performance.

This work is immersed in one of those schools that work daily to be effective in responding to the goals set in order to offer inclusive and quality education and thus improve the academic performance of all students by overcoming potential inequalities. This is due to the commitment of teachers to improve the quality of teaching-learning processes in the center and guarantee the principle of equity for all students. Being a center located in a marginal neighborhood, socially and economically deprived, with families experiencing unemployment and other vulnerable circumstances, the academic level of the students was very low. This situation leads to those families who have the possibility of not sending their children to this center, but to other schools in the area, will do so. This school thereby becomes marginal. The teachers set out to change this situation through innovation

and improvement actions (included in the Center Educational Project) and are essentially focused on the areas of Language and Mathematics.

Explicitly, the purpose of this contribution is to analyze the improvement in student academic performance after introducing a series of changes in relation to the curriculum – linguistic and mathematical competence – and teaching–learning strategies used by teachers and students. In short, we are assessing the ability of variables focused on the educational process in the classrooms – didactic strategies – for the improvement of students' academic performance.

As has been shown, there has been in-depth analysis of academic performance in research with personal and social variables, but there are fewer studies focused upon school variables in vulnerable contexts. In this way, the questions that guide this research are as follows: Can didactic strategies improve the academic performance of students in these contexts? Can an educational intervention treatment based on collaborative/cooperative learning and project work improve the performance of students in vulnerable contexts in the areas of language and mathematics?

MATERIALS AND METHODS

Participants

The population under study is composed of Primary Education students (6–12 years old) belonging to two public schools located in an urban area 7 km away from the city of Murcia (Spain). As the management team informed, their families' socio-economic and cultural level is medium–low and they mostly work in the service sector. The expectation of families regarding the education of their children is centered on their children receiving a standard education, and only a small proportion of the parents aspire for their children to engage in higher studies.

Both schools have similar characteristics, are state run, and are accountable to the Ministry of Education. The school where the experimental group is based has 20 units, 6 of Infant Education and 14 of Primary Education, which represents a total of 476 students, distributed in the stages of Infants (148 students) and Primary (328 students). The control group school has 15 units, 5 of Infant Education and 10 of Primary Education, which represents a total of 367 students, 135 in Infants and 232 in Primary Education.

Discretionary non-probabilistic or judgment sampling was used (McMillan and Schumacher, 2005; Hernández-Pina and Maquilón, 2010), dependent on knowledge of the characteristics of both schools as well as their population. Finally, the actual sample was composed of 228 students, which, based on a total population of 560 students, is considered a 95% confidence level with 5% sample error (calculation made via surveymonkey.com).

Of this sample, 130 (57.01%) were part of the control group and 98 (42.99%) were part of the experimental group. Regarding its distribution by gender, 137 (60.1%) participants were boys, and the remaining 91 (39.9%) were girls, without knowledge of the sexual identity of any of the participants. The description of the participants is included in **Table 1**.

Measure

The measure used was the average score obtained in the assessment of language and mathematics in the first trimester, as well as those in the final evaluation in both centers (**Table 2**). These evaluations were not specifically designed for this project in respect of the performance and criteria of each of the centers, since both use the learning standards established in educational regulations at the national and regional level.

Procedure

The informed consent of the parents for the participation of the students in the proposed program was requested by the management team of the center by means of a signed letter. The program was carried out in the experimental group as a study/work activity in the center for the development of the curriculum in Language and Mathematics in which all students participated.

Both groups (experimental and control) used the same series of textbook for the development of linguistic and mathematical competencies (Ibarrola, 2017) covering first to sixth grade of Primary education. In addition, in the experimental group, the use of the textbook was combined with the program specifically designed to improve academic performance in language and mathematics.

Once the existing levels in Language and Mathematics of both groups, control and experimental, were established, the intervention procedure for experimental group was designed. This was focused on the process factors because they are the ones that have a significant impact on the academic performance of the students. In this sense, the objectives to be achieved were established and activities were designed according to the needs and characteristics of the students in order to improve their skills. In the same way, the teaching methodology and the didactic strategies, which most encouraged the students' motivation and confidence toward learning, were established. To do this, we turned to cooperative learning, peer tutoring, and discovery learning methods that connect the curricular content with real-life experiences and the interests of the students, carrying out small research projects and problem-solving tasks (problem-based learning). In addition, the intervention prioritized that the teacher–student interaction be characterized by a safe and close environment.

Therefore, the treatment carried out for the improvement of academic performance consisted, in general, in the realization of a series of activities differentiated by courses. In the linguistic area, students presented suggestions for poems, short texts, newspaper articles, stories, comics, puns, oral expression

TABLE 1 | Distribution of the participants.

Center	Experimental	Control	Total
Gender			
Male	60 (61.2%)	77 (59.23%)	137
Female	38 (38.2%)	53 (40.77%)	91
Total	98 (100%)	130 (100%)	228

TABLE 2 | Description of the measurement process.

Experimental group	1st-trimester average score (linguistics + mathematics)	Treatment	Final average score (linguistics + mathematics)
Control group	1st-trimester average score (linguistics + mathematics)		Final average score (linguistics + mathematics)

exercises, etc., turning the classroom into a creative literary workshop with the aim of integrating the development of skills and multiple intelligences. In the mathematical area, the work undertook highlighted the development of logical reasoning, using puzzles and specific mathematical challenges for each cycle.

The time allocated to the intervention was three weekly sessions each of 45 min, from January to June. The coordination and development of each session as well as the final evaluation of the students fell to each of the classroom tutors.

In the case of the control group, the dynamics of the learning process undertaken up to that point in the areas of language and mathematics were not changed, consisting of reading through the lessons, teacher explanation, and carrying out the exercises (Ibarrola, 2017).

Design and Data Analysis

A quasi-experimental design with pretest and post-test and control group was adopted (Hernández-Pina and Maquilón, 2010; Campbell and Stanley, 2015). This type of design allows comparison between a group that has received intervention, called an experimental group, and another one called the control group to which no modification in the educational process has been applied. It is one of the most commonly used designs in socio-educational research because it does not require a random assignment of participants, but allows for the attainment of balanced groups.

Given the characteristics of the design used, the independent factor or variable belonged to one or the other group (experimental or control), while the criterion or dependent variable was the mean of the academic performance of the subjects (Williams, 1952; Box, 1971; Berenblut and Webb, 1974), obtained in the evaluations in the areas of mathematics and language.

For data analysis, the general linear model (GLM) of repeated measures has been used in order to assess the effectiveness of the treatment, through which, groups of related dependent variables that represent different measures of the same attribute are analyzed (Freeman, 1973; Bryant and Paulson, 1976; Wood et al., 1978; Vuchkov and Solakov, 1980; Defeo and Myers, 1992). To carry out the data analysis, the statistical package SPSS, version 21.0 was used.

The researcher must be conscious of the ethical responsibility involved in the conduct of an investigation, especially when it deals with human beings (McMillan and Schumacher, 2005). In this way, following the principles and norms published by the American Psychological Association [APA] (2010), the rights and dignity of the participants were guaranteed at all times in this investigation; this was endorsed by the favorable report issued for the realization of this research by the Research Ethics Commission of the University of Murcia.

RESULTS

The objective of this work is focused on verifying the effectiveness of the treatment of improving linguistic and mathematical competences as a favorable element of academic performance at a general level. In order to respond to this, our first step was to ensure the similarity of the experimental and control groups before the intervention; therefore, a comparative analysis of means was performed in the pretest phase, verifying that there were no statistically significant differences between these groups (**Table 3**).

The analysis of the sample demonstrated that the populations are distributed normally. The Box M test was applied, obtaining non-homogeneous variance–covariance matrices (F = 13,561; gl = 5749165.693 and $p < 0.001$), but, since the groups are approximately the same size [according to Hair et al. (1999), the size of the largest group divided by the size of the smallest group should be less than 1.5] and the highest variance ratio between the groups does not exceed the 10:1 ratio considered as the maximum limit in the analysis of profiles for Tabachnick and Fidel (2007), the violation of this assumption has a minimal impact.

To evaluate the effectiveness of the treatment in the sample, a GLM of repeated measures was also used. The dependent variable was the mean academic performance. Gender was included as covariate, obtaining a non-significant interaction. This implies that there is a significant effect of the intervention program on the academic performance of the experimental group, while the covariate was not significant.

Intra-subject factors are represented in the evaluation times (pre- and post-test) for the dependent variable. Inter-subject factors depended on the presence or absence of the treatment (i.e., the experimental group or the control group).

As shown in **Table 4**, the effects of the intra-subject and inter-subject test show that the effect of the interaction between the time of the evaluation (pretest and post-test) and the implementation of the program of activities is significant ($p < 0.001$). In addition, the observed power (the correct rejection of the null hypothesis of equality of means) is optimal.

The effect size (η^2), that is to say, the proportion of the total variability attributable to a factor (or the magnitude of the difference between one time and another, as a result of the interaction between the moment of evaluation and the application of the program), obtains the best results when the

TABLE 3 | Test *t* of mean difference. Pretest.

	M Experimental	dt	M Control	dt	*p*	*t*	gl
Academic performance	6.00	1.86	5.66	1.94	0.18	1.35	229

TABLE 4 | Summary of intra-/inter-subject analysis.

Source	Type III	gl	*F*	*p*	η^2 partial	Obs. power
Academic achievement						
Intra	0.863	1	0.902	0.343	0.004	0.157
Intra × gender	0.367	1	0.384	0.536	0.002	0.095
Intra × inter	70.195	1	73.384	0.000	0.246	1.000
Error intra	215.224	225				
Inter	143.635	1	23.824	0.000	0.096	0.998
Error inter	1356.512	225				

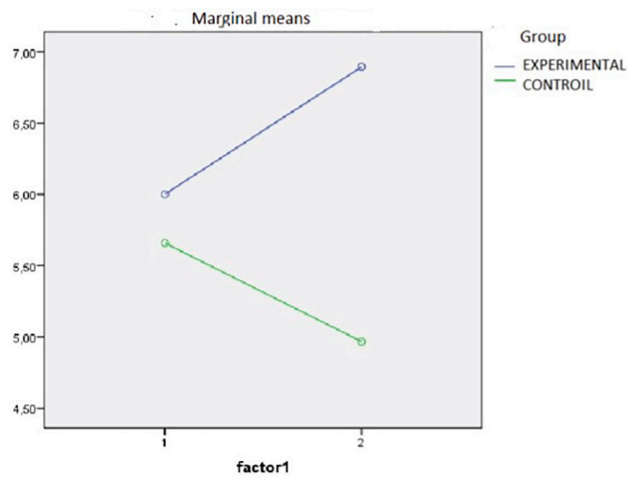


FIGURE 1 | Graphical representation of the means of academic performance of the two groups in pre- and post-test time.

interaction is analyzed according to the experimental and control group, reaching values of 0.246.

Figure 1 presents the interaction graph, which illustrates the directions of the differences. The total score of the academic performance of the experimental group was significantly higher once the intervention program was completed, the mean score being 6.90 (SD = 2.03), while it was 4.97 (SD = 1.65) for the control group.

Finally, it should be noted, in the analysis of the means obtained by both groups, that in each case statistically significant differences were found after the Student's *t* test for related samples, although in the case of the experimental group, as observed in **Figure 1**, they were increased ($p < 0.001$), while for the control group, the mean values decreased ($p < 0.001$).

DISCUSSION AND CONCLUSION

The results presented above demonstrate the value and effectiveness of the program used given the good results obtained and the improvement in the academic performance of the participating students compared to those who did not benefit from the advantages of the said program. Since the main objective of this work was to verify the effectiveness of the implementation of an intervention program based on cooperative learning, peer tutoring, and problem-based learning, it has been possible to verify that the final scores have risen in the experimental group by almost 1 point, based on a global average score from 6.00 points before treatment, to a final overall average score of 6.90 points after the application of the program.

In this sense, we agree with other studies that highlight the importance of learning strategies (Preckel and Brunner, 2015; Dunn et al., 2017; Hincapie et al., 2018; Karrera et al., 2019), or the teaching methodology used (Nath, 2012; Payandeh-Najafabadi et al., 2013) in the improvement of academic performance. Consequently, we are able to demonstrate how important it is for the school to improve performance and overcome inequalities as indicated by Ainscow (2005), Arnaiz et al. (2018), Flecha and Buslon (2016), Madigan et al. (2016), and Murillo (2005).

The school where the intervention program was developed had the goal of developing the necessary processes to promote quality education so that the students' language and math skills improved; desired to be more and more effective in achieving this end, as indicated by different experts, this is a basic requirement in achieving success for everyone and, consequently, good levels of academic performance (Murillo, 2008; Arnaiz, 2012; Van Mieghem et al., 2018).

This work has empirically proven the effectiveness of the implementation of a program for the improvement of skills in terms of academic performance. This program could be the start of a cycle of constant improvement, where the performance factor is eclipsed by other types of elements, such as the collective atmosphere of pursuing the enrichment of the teaching and learning processes, or the wider social recognition of the school's efforts (Murillo, 2007; Jornet et al., 2012).

Ramberg et al. (2019) established the relationship between student and teacher as a performance-enhancing element. We believe that teacher satisfaction with innovations aimed at the improvement of educational quality, as well as the acceptance of common goals with the school, could be a differentiating element to take into account in future work. In this sense, we do not want to lose the opportunity to point out some of the contributions that have been given to us by professionals involved in the implementation of the program, and collected in our field notebooks during some of the conversations held with the teaching staff, which appear to verify, subjectively, those data that we presented previously.

"My personal assessment is that students enjoy playing with language, having fun and making us participants in their compositions, which in many cases are extraordinarily original, allowing the transmission of this poetic flow to their future school studies" (5th grade teacher of Education Primary).

"The overall assessment of the work done with the students of the 5th and 6th grade of primary school has been highly satisfactory" (6th grade Primary Education teacher).

"Throughout the course we have been able to constantly enjoy inventing, creating and recreating to find appropriate learning processes that will take us away from routine and preconceived ideas" (5th grade Primary Education teacher).

In spite of the good results obtained, it is possible to deepen the research carried out by incorporating new study variables, which would allow advancement in the identification of the factors that promote the improvement of academic performance and, consequently, the development of effective schools. In the same way, it would be interesting to use other measurement tests in order to access the improvement of student performance – different to teacher grading – such as standardized tests in order to eliminate elements of subjectivity or equivalence between teachers, classrooms, and educational centers (McCoach and Siegle, 2003). Together with this, new variables could be incorporated to measure the effectiveness of the programs used, an element such as students in disadvantaged contexts, to show if successful actions improve academic results and consequently contribute to overcoming inequalities, such as indicated by Flecha and Buslon (2016).

Ultimately, the study carried out indicates the importance of learning strategies and applied teaching methodologies, understanding both within the conglomerate of process factors in the improvement of academic performance (Brophy and Good, 1986; Davis and Thomas, 1992; Reynolds and Cuttance, 1992; Scheerens, 1992; Ramasut and Reynolds, 1993; Creemers, 1994; Reynolds et al., 1994; Castejón, 1996; Wyatt, 1996; Murillo, 2003a,b). Hence, obviously, we want to point out the importance of what we found and the implementation of it in other educational centers in order to promote effective schools capable of offering a quality, equitable, and inclusive education for all, as UNESCO (2017) reminds us.

Finally, it is possible to express the limitations present in this study, which reside in the fact that only two curricular subjects or areas (language and mathematics) have been taken into account, ignoring others that are equally relevant in the measurement of academic performance. Another aspect to consider could be the application of a retest to verify the effectiveness of the implemented program, although this would take on greater importance if the experience were isolated as an anecdotal implementation of the program. Similarly, the participation of only two centers limits the extrapolation of results obtained to other contexts with similar characteristics. All this being said, it invites us to carry out new studies that overcome these limitations and promote educational improvement in this and other educational centers.

DATA AVAILABILITY STATEMENT

The data may be consulted by making an express request to the authors through the following email: parnaiz@um.es.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Commission of the University of Murcia. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

The study was carried out collaboratively, both in the development of the project and distribution tasks, although we found specific tasks had to be defined at different moments of the process. PA-S coordinated and led the group, in addition to revisions and corrections of text, analysis of data, discussion, and presentation of conclusions. SA and RH were directly responsible for the theoretical section, as well as for subsequent revisions and corrections. AM was responsible for data analysis and presentation of results.

REFERENCES

Ainscow, A. (2005). "El próximo gran reto: la mejora de la escuela inclusiva [The next big challenge: the improvement of inclusive school]," in *Proceedings of the Opening communication of Congress of School Effectiveness and Improvement*, Barcelona.

American Psychological Association [APA] (2010). Principios Éticos de los Psicólogos y Código de Conducta (APA) [Ethical principles of psychologists and code of conduct (APA)]. Available at: http://www.uhu.es/susana_paino/EP/CcAPA.pdf (accessed January 24, 2019).

Arnaiz, P. (2012). Escuelas eficaces e inclusivas: cómo favorecer su desarrollo [Effective and inclusive schools: how to promote their development]. *Educ. Siglo XXI.* 30, 25–44.

Arnaiz, P., De Haro, R., and Mirete, A. B. (2018). "Procesos de mejora e inclusión educativa en centros educativos de la Región de Murcia [Educational improvement and inclusion processes in educational centers in the Region of Murcia]," in *Inclusión y Mejora Educativa [Inclusion and Educational Improvement]*, eds J. C. Torrego, L. Rayón, Y. Muñoz, and P. Gómez, (Alcalá de Henares: Servicio Publicaciones Universidad), 271–281.

Berenblut, I. I., and Webb, G. I. (1974). Experimental design in presence of autocorrelated errors. *Biometrika* 61, 427–437. doi: 10.1093/biomet/61.3.427

Box, M. J. (1971). Experimental design criterion for precise estimation of a subset of parameters in a nonlinear model. *Biometrika* 58, 149–153. doi: 10.1093/biomet/58.1.149

Brophy, J., and Good, T. (1986). "Teacher behaviour and student achievement," in *Handbook of Research on Teaching*, ed. M. C. Wittrock, (New York: Macmillan), 328–375.

Bryant, J. L., and Paulson, A. S. (1976). Extension of Tukeys method of multiple comparisons to experimental-designs with random concomitant variables. *Biometrika* 63, 631–638. doi: 10.1093/biomet/63.3.631

Campbell, D., and Stanley, J. (2015). *Experimental and Quasi-Experimental Designs for Research*. Boston: Ravenio Books.

Castejón, J. L. (1996). *Determinantes del Rendimiento Académico de los Estudiantes y de los Centros Educativos: Modelos y Factores [Determinants of the Academic*

Performance of Students and Schools: Models and Factors. Alicante: Editorial Club Universitario.

Castejón, J. L., Gilar, R., Miñano, P., and Veas, A. (2016). Identificación y establecimiento de las características motivacionales y actitudinales de los estudiantes con rendimiento académico menor de lo esperado según su capacidad (underachievement) [Identifying and establishing the attitudes and motivations of low academic performance in underachieving students]. *Eur. J. Educ. Psychol.* 9, 63–71. doi: 10.1016/j.ejeps.2016.04.001

Coleman, J. (1968). The concept of equality of educational opportunity harvard. *Educ. Rev.* 38, 7–22. doi: 10.17763/haer.38.1.m3770776577415m2

Creemers, B. (1994). *The Effective Classroom*. London: Castell.

Davis, G., and Thomas, M. (1992). *Escuelas Eficaces y Profesores Eficientes [Efficent Schools and Efficient Teachers]*. Madrid: La Muralla.

Dedrick, R. F., Shaunessy-Dedrick, E., Suldo, S. M., and Ferron, J. M. (2015). Psychometric properties of the school attitude assessment survey-revised with international baccalaureate high school students. *Gift. Child Q.* 59, 38–54. doi: 10.1177/001698621459596

Defeo, P., and Myers, R. H. (1992). A new look at experimental-design robustness. *Biometrika* 79, 375–380. doi: 10.1093/biomet/79.2.375

Dunn, M. E., Shelnut, J., Ryan, J. B., and Katsiyannis, A. (2017). A systematic review of peer-mediated interventions on the academic achievement of students with emotional/behavioral disorders. *Educ. Treat. Child.* 40, 497–524. doi: 10.1353/etc.2017.0022

Durán, D., and Vidal, V. (2004). *Tutoría Entre Iguales: de la Teoría a la Práctica [Peer Tutoring: From Theory to Practice]*. Barcelona: Graó.

Escarbajal, A., Navarro, J., and Arnaiz, P. (2019). El rendimiento académico del alumnado autóctono y de origen inmigrante en la Región de Murcia [The academic performance of non-inmigrant students and those of inmigrant origin within the Region of Murcia]. *Tendencias Pedagógicas* 33, 5–17. doi: 10.15366/tp2019.33.001

Flecha, R., and Buslon, N. (2016). 50 años después del Informe Coleman. Las actuaciones educativas de éxito sí mejoran los resultados académicos [50 Years after the Coleman Report. SEAs Improve Academic Results]. *Int. J. Sociol. Educ.* 5, 127–143. doi: 10.17583/rise.2016.2087

Franco, V., Melo, M., Santos, G., Apolónio, A., and Amaral, L. (2017). A national early intervention system as a strategy to promote inclusion and academic achievement in Portugal. *Front. Psychol.* 8:1137. doi: 10.3389/fpsyg.2017.01137

Freeman, G. H. (1973). Statistical methods for the analysis of genotype environment interactions. *Heredity* 31, 339–354. doi: 10.1038/hdy.1973.90

García-Martín, S., and Cantón-Mayo, I. (2019). Uso de tecnologías y rendimiento académico en estudiantes adolecentes [Use of technologies and academic performance in adolescent students]. *Comunicar* 59, 73–81. doi: 10.3916/C59-2019-07

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1999). *Multivariate Data Analysis*. Upper Saddle River, NJ: Prentice Hall Iberia.

Hernández-Pina, F., and Maquilón, J. J. (2010). "Introducción a los diseños de investigación educativa [Introduction to educational research designs]," in *Principios, Métodos y Técnicas Esenciales Para la Investigación Educativa [Essential Principles, Methods and Techniques for Educational Research]*, ed. S. Nieto, (Madrid: Dykinson), 109–126.

Hincapie, D., Ramos, A., and Chirino, V. (2018). Aprendizaje Basado en Problemas como estrategia de Aprendizaje Activo y su incidencia en el rendimiento académico y Pensamiento Crítico de estudiantes de Medicina de Segundo Semestre [Problem based learning as an active learning strategy and its impact on academic performance and critical thinking of medical students]. *Rev. Complut. Educ.* 29, 35–52.

Hinojo, F. J., Aznar, I., Romero, J. M., and Marín, J. A. (2019). Influencia del aula invertida en el rendimiento académico. Una revisión sistemática [Influence of the flipped classroom on academic performance. A systematic review]. *Campus Virtuales* 8, 9–18.

Hsieh, P., Sullivan, J. R., and Guerra, N. S. (2007). A closer look at college students: self-efficacy and goal orientation. *J. Adv. Acad.* 18, 454–476. doi: 10.4219/jaa-2007-500

Ibarrola, B. (2017). *Sentir y Pensar. Cuadernos Prácticos [Feel and Think. Practical Notebooks]*. Madrid: GRUPO SM.

Jeynes, W. (2010). The salience of the subtle aspects of parental involvement and encouraging that involvement: implications for school-based programs. *Teach. Coll. Rec.* 112, 747–774.

Jornet, J. M., González-Such, J., and Perales, M. J. (2012). Diseño de cuestionarios de contexto para la evaluación de sistemas educativos: optimización de la medida de constructos complejos [Context questionnaire design for the evaluation of educational systems: optimization of complex constructs measurement]. *Bordón* 64, 89–110.

Karrera, I., Garmendia, M., and Arguiñano, A. (2019). Sinergias entre inclusión, competencias y rendimiento escolar: el caso de la escuela pública de Antzuola [Inclusion, competences and school performance. The case of Antzuola primary education public school]. Profesorado. *Rev. Curríc. Form. Prof.* 23, 83–102. doi: 10.30827/profesorado.v23i1.9145

Kunz, J. S. (2014). Analyzing Educational achievement differences between second generation immigrants: comparing Germany and German Speaking Switzerland. *German Econ. Rev.* 17, 61–91. doi: 10.1111/geer.12062

Lu, L., Weber, H. S., Spinath, F. M., and Shi, J. (2011). Predicting school achievement from cognitive and non-cognitive variables in a Chinese sample of elementary school children. *Intelligence* 39, 130–140. doi: 10.1016/j.intell.2011.02.002

Madigan, K., Cross, R. W., Smolkowski, K., and Strycker, L. A. (2016). Association between school wide positive behavioural interventions and supports and academic achievement: a 9-year evaluation. *Educ. Res. Eval.* 22, 402–421. doi: 10.1080/13803611.2016.1256783

Marqués, I. (2016). Apuntes sobre el informe coleman. Sobre la díficil convivencia de los principios igualitarios en un mundo desigual [Notes on coleman report. about the difficult coexistence of equality principles in a non equal world]. *Int. J. Sociol. Educ.* 5, 107–126. doi: 10.17583/rise.2016.2092

McCoach, D. B., and Siegle, D. (2003). Factors that differentiate underachieving gifted students from high-achieving gifted students. *Gift. Child Q.* 47, 144–154. doi: 10.1177/001698620304700205

McMillan, J. H., and Schumacher, S. (2005). *Research in Education: Evidence-Based Inquiry*. Madrid: Pearson.

Meunier, M. (2011). Immigration and student achievement: evidence from Switzerland. *Econ. Educ. Rev.* 30, 16–38. doi: 10.1016/j.econedurev.2010.06.017

Miller, B. R., Behrens, J. T., Greene, B. A., and Newman, D. (1993). Goals and perceived ability: impact on student valuing, self-regulation, and persistence. *Contemp. Educ. Psychol.* 18, 2–14. doi: 10.1006/ceps.1993.1002

Miñano, P., Gilar, R., and Castejón, J. L. (2012). A structural model of cognitive-motivational variables as explanatory factors of academic achievement in Spanish language and mathematics. *Anal. Psicol.* 28, 45–54.

Miralles, M. J., Castejón, J. L., Pérez, A. M., and Gilar, R. (2012). El análisis de los efectos de la escuela sobre el rendimiento académico en matemáticas: un análisis multinivel con datos de PISA 2003 [Analysing the influence of schools on academic performance in mathematics: a multilevel analysis using PISA 2003 data]. *J. Psychol. Educ.* 7, 83–110.

Molina-López, M. M., Sanz-Magallón, G., and García, M. C. (2018). Estimación del nivel de competencia entre centros educativos en la región de Madrid y análisis de su efecto sobre el rendimiento académico [Estimation of the school competition level in the region of Madrid and its effects on academic performance]. *Rev. Educ.* 382, 199–224. doi: 10.4438/1988-592X-RE-2018-382-397

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., and Russel, E. (1989). Un estudio de escuelas junior efectivas [A study of effective junior schools]. *Rev. Int. Investig. Educ* 13, 753–768. doi: 10.1016/0883-0355 (89)90026-8

Murillo, F. J. (2003a). "El movimiento de investigación de Eficacia Escolar [The school efficiency research movement]," in *La Investigación Sobre Eficacia Escolar en Iberoamérica. Revisión Internacional Sobre el Estado del Arte [Research on school Effectiveness in Latin America. International Review on the State of the Art]*, ed. F. J. Murillo, (Bogotá: Convenio Andrés Bello), 53–92.

Murillo, F. J. (2003b). Una panorámica de la investigación iberoamericana sobre eficacia escolar [Ibero-american School Effectiveness Research at a Glance]. *Rev. Electrón. Iberoamer. Sobre Calidad Eficacia Cambio Educ.* 1, 1–14.

Murillo, F. J. (2005). *La Investigación Sobre Eficacia Escolar [School Efficiency Research]*. Barcelona: Octaedro.

Murillo, F. J. (2007). *Investigación Iberoamericana Sobre Eficacia Escolar [Ibero-American Research on School Efficiency]*. Bogotá: Convenio Andrés Bello.

Murillo, F. J. (2008). Hacia un modelo de eficacia escolar. Estudio multinivel sobre los factores de eficacia en las escuelas españolas [Moving towards an school effectiveness model. multilevel study regarding effectiveness factors at

Spanish Schools]. *Rev. Iberoam. Sobre Calidad Eficacia Cambio Educ.* 6, 4–28. doi: 10.1016/j.aprim.2015.04.012

Muro, A., Soler, J., Cebolla, A., and Cladellas, R. (2018). A positive psychological intervention for failing students: does it improve academic achievement and motivation? A pilot study. *Learn. Motiv.* 63, 126–132. doi: 10.1016/j.lmot.2018.04.002

Nath, S. R. (2012). Factors influencing primary students 'learning achievement in Bangladesh. *Res. Educ.* 88, 50–63. doi: 10.7227/RIE.88.1.5

Payandeh-Najafabadi, A. T., Omidi-Najafabadi, M., and Farid-Rohani, M. R. (2013). Factors contributing to academic achievement: a Bayesian structure equation modelling study. *Int. J. Math. Educ. Sci. Technol.* 44, 490–500. doi: 10.1080/0020739x.2012.742149

Perry, L., and McConney, A. (2010a). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teach. Coll. Rec.* 112, 1137–1162.

Perry, L., and McConney, A. (2010b). School socio-economic composition and student outcomes in Australia: implications for educational policy. *Aust. J. Educ.* 54, 72–85. doi: 10.1177/000494411005400106

Preckel, F., and Brunner, M. (2015). Academic self-concept, achievement goals, and achievement: is their relation the same for academic achievers and underachievers? *Gift. Talent. Int.* 30, 68–84. doi: 10.1080/15332276.2015.1137458

Pulido, F., and Herrera, F. (2019). Estados emocionales contrapuestos como predictores del rendimiento académico en secundaria [Opposite emotional states as predictor of academic achievement in high school]. *Rev. Investig. Educ.* 37, 93–109. doi: 10.6018/rie.37.1.289821

Ramasut, A., and Reynolds, D. (1993). "Developing effective whole school. approaches to special educational needs: from school effectiveness theory to school development practice," in *Is There a Desk With my Name On IT*?, ed. R. Slee, (London: The Falmer Press), 219–240.

Ramberg, J., Brolin, S., Almquist, Y., and Modin, B. (2019). School effectiveness and students' perceptions of teacher caring: a multilevel study. *Improv. Sch.* 22, 55–71. doi: 10.1177/1365480218764693

Reynolds, D., Creemers, B. P. M., Nesselrodt, P. S., Schaffer, E. C., Stringfield, S., and Teddlie, S. (1994). *Advances in School Effectiveness Research and Practice*. Oxford: Elsevier Science.

Reynolds, D., and Cuttance, P. (1992). *School Effectiveness Research, Policy and Practice*. London: Cassell.

Rodríguez, D., and Guzmán, R. (2019). Rendimiento académico de adolescentes declarados en situación de riesgo [Academic achievement of adolescents declared at risk]. *Rev. Investig. Educ.* 37, 147–162. doi: 10.6018/rie.37.1.303391

Ruiz-Esteban, C., Méndez, I., and Díaz-Herrero, A. (2018). Evolución de las metas académicas en función del sexo y la edad y su influencia en el rendimiento académico en adolescentes murcianos [Evolution of academic goals base don gender and age and their influence on academic achievement in adolescents in Murcia]. *Educ. Siglo XXI.* 36, 319–332. doi: 10.6018/j/350021

Rutter, M., Maughan, B., Mortimore, P., and Ouston, J. (1979). *Fifteen Thousand Hours: Secondary Schools and Their Effects on Children*. Cambridge, MA: Harvard University Press.

Santos, M. A., Godás, A., and Lorenzo, M. (2013). Rendimiento académico y diversidad cultural: El eje lingüístico [School Performance and cultural diversity: the linguistic axis]. *Rev. Españ. Pedag.* 256, 461–478.

Scheerens, J. (1992). *Effective Schooling: Research, Theory and Practice*. London: Casell.

Schwartz, D., Gorman, A. H., Nakamoto, J., and McKay, T. (2006). Popularity, social acceptance, and aggression in adolescent peer groups: links with academic performance and school attendance. *Dev. Psychol.* 42:1116. doi: 10.1037/0012-1649.42.6.1116

Tabachnick, B., and Fidel, L. (2007). *Using Mulrivariate Statistics*, 5th Edn. Boston: Allyn & Bacon.

Takehara, K., Ganchimeg, T., Kikuchi, A., Gundegmaa, L., Altantsetseg, L., Aoki, A., et al. (2019). The effectiveness of exercise intervention for academic achievement, cognitive function, and physical health among children in Mongolia: a cluster RCT study protocol. *BMC Public Health* 19:697. doi: 10.1186/s12889-019-6986-8

Teddlie, C., and Stringfield, S. (1993). *Schools Make a Difference: Lessons Learned From 10- Year Study of School Effects*. New York, NY: Teachers College Press.

UNESCO, (2015). *Declaración de Incheon y Marco de Acción Para la Realización del Objetivo de Desarrollo Sostenible 4 [Incheon Declaration and Framework for Action to achieve the Sustainable Development Goal 4]*. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000245656_spa (accessed July 21, 2019).

UNESCO, (2017). *Guía Para Asegurar la Inclusión y la Equidad en la Educación [Guide to Ensure Inclusion and Equity in Education]*. París: UNESCO.

Van Mieghem, A., Verschueren, K., Petry, K., and Struyf, E. (2018). An analysis of research on inclusive education: a systematic search and meta review. *Int. J. Inclus. Educ.* doi: 10.1080/13603116.2018.1482012

Veas, A., López-López, J. A., Gilar, R., Miñano, P., and Castejón, J. L. (2017). Differences in cognitive, motivational and contextual variables between under-achieving, normally-achieving, and over-achieving students: a mixed-effects analysis. *Psicothema* 29, 533–538. doi: 10.7334/psicothema2016.283

Vuchkov, I. N., and Solakov, E. B. (1980). The influence of experimental design on robustness to non-normality of the F test in regression analysis. *Biometrika* 67, 489–492.

Wilder, S. (2014). Effects of parental involvement on academic achievement: a meta-synthesis. *Educ. Rev.* 66, 377–397. doi: 10.1080/00131911.2013.780009

Williams, E. J. (1952). Some exact tests in multivariate analysis. *Biometrika* 39, 17–31. doi: 10.2307/2332460

Wood, D., Wood, H., and Middleton, D. (1978). An experimental evaluation of four face-to-face teaching strategies. *Int. J. Behav. Dev.* 1, 131–147. doi: 10.1097/CEJ.0000000000000201

Wyatt, T. (1996). School effectiveness research: dead end, damp squib or smouldering fuse? *Issues Educ. Res.* 6, 79–112.

Zimmerman, B. J., Bandura, A., and Martínez-Pons, M. (1992). Self-motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *Am. Educ. Res. J.* 29, 663–676.

Zuffianò, A., Alessandri, G., Gerbino, M., and Luengo, B. P. (2013). Academic achievement: the unique contribution of self-efficacy beliefs in self-regulated learning be. *Learn. Individ. Differ.* 23, 158–162. doi: 10.1016/j.lindif.2012.07.010

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.

Copyright © 2020 Arnaiz-Sánchez, de Haro, Alcaraz and Mirete Ruiz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 March 2020
doi: 10.3389/fpsyg.2020.00233

Students' Evaluation of Teaching and Their Academic Achievement in a Higher Education Institution of Ecuador

Tarquino Sánchez[1], Raquel Gilar-Corbi[2], Juan-Luis Castejón[2], Jack Vidal[1] and Jaime León[3]*

[1] National Polytechnic School, Quito, Ecuador, [2] Developmental and Educational Psychology Department, University of Alicante, Alicante, Spain, [3] Department of Education, University of Las Palmas de Gran Canaria, Las Palmas, Spain

This paper addresses the relationship between student evaluation of teaching (SET) and academic achievement in higher education. Meta-analytic studies on teaching effectiveness show a wide range of results, ranging from small to medium correlations between SET and student achievement, based on diverse methodological approaches, sample size studies, and contexts. This work aimed to relate SET, prior academic achievement, and academic achievement in a large sample of higher education students and teachers, using different methodological procedures, which consider as distinct units of analysis the group class and the individuals, the variability between students within classes, and the variability between group-class means, simultaneously. The data analysis included the calculation of group-class means and its relationship with the group-class mean academic achievement, through correlation and hierarchical regression techniques; additionally, a multilevel path analysis was applied to the relationship between prior academic achievement, SET, and their academic achievement, considering the variability among group classes. A multisection analysis was also carried out in those course disciplines in which there was more than one class group (section). The results of individual and group-class analysis revealed that SET was moderately low but related to academic achievement in a significant way once the effect of previous academic achievement was controlled. In addition, multilevel path analysis revealed the effect of SET on achievement, both within and between group-class levels. The results of the analysis carried out in the course disciplines with different sections, according to a multisection design, yielded similar results to the individual and aggregated data analyses. Taken together, the results revealed that SET was low related to academic achievement, once the effect of previous academic achievement was controlled. From these results, it follows that the use of SET as a measure of teachers' effectiveness for making administrative decisions remains controversial.

Keywords: student evaluation of teaching ratings, academic achievement, teaching effectiveness, multisection study, multilevel analysis

OPEN ACCESS

Edited by:
José Jesús Gázquez, University of Almería, Spain

Reviewed by:
Olivia López Martínez, University of Murcia, Spain
Efstathios Stefos, The National University of Education (Ecuador), Ecuador

****Correspondence:***
Raquel Gilar-Corbi
raquel.gilar@ua.es

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

Received: *02 November 2019*
Accepted: *31 January 2020*
Published: *06 March 2020*

Citation:
Sánchez T, Gilar-Corbi R, Castejón J-L, Vidal J and León J (2020) Students' Evaluation of Teaching and Their Academic Achievement in a Higher Education Institution of Ecuador. Front. Psychol. 11:233. doi: 10.3389/fpsyg.2020.00233

INTRODUCTION

Student evaluation of teaching (SET) is a generalized practice in almost every institution of higher education around the world (Richardson, 2005; Zabaleta, 2007; Huybers, 2014) – from European countries (Husbands and Fosh, 1993) to Australian and North American universities (Richardson, 2005) and South American higher education institutions (Pareja, 1986).

However, this issue is contemporary and is a topic still open to question in higher education. Researchers working on SET have not yet provided a clear answer about some critical questions on the validity and utility of evaluations (Marsh, 2007; Spooren et al., 2013). Although the use of student evaluation as feedback for teachers is not so controversial, the utilization of student evaluation for measuring teaching effectiveness, based on the assumption that students learn better with highly rated teachers, is very controversial.

One of the central controversial points is the relation of SET ratings to their learning outcomes, such as academic achievement (Uttl et al., 2017). The evidence in support of SET as a measure of teachers' instruction effectiveness comes from the studies showing a correlation between measures of student evaluation and student achievement.

Methodological Concerns/Questions

Initially, the validity of students' judgments might be proven by the correlation between SET and academic achievement. However, the evaluation criteria for distinct course units may differ, and students' grades cannot be considered a simple measure of teaching effectiveness (Richardson, 2005).

The key evidence provided in favor of SET as a measure of the effectiveness of teachers' instruction is multisection studies (Uttl et al., 2017). Leventhal (1975) and Cohen (1981) defend that the stronger SET validation design implicates the designation of students to different sections of a multisection course. If the designation is random, between-section differences in student performance can be caused by differences in teachers. When students self-select into sections, it can be difficult to infer rating/achievement relationship. If this is the case, Marsh and Overall (1980) consider that, in these studies, they should provide adequate controls/measures for initial ability or prior achievement.

Some researchers (Cohen, 1981; Clayson, 2009; Uttl et al., 2017) point out that student achievement is highly dependent on factors such as intelligence or prior achievement and that to fully control these factors, it is necessary to randomly assign students to classes and teachers or, alternatively, use other control procedures of initial student ability or achievement, such as analysis of covariance using measures of prior academic achievement or capacity as covariates; using the change in grades based on pretest and posttest moments; or regressing individual students' performance scores on measures of students' prior achievement and using residual gains in performance, averaged across students within sections, as measures of learning. It is advisable to use a statistical procedure in which both ratings and performance are adjusted for initial student ability or performance.

An ideal multisection study design entails a course discipline or subject matter with many comparable group class – sections – taking the same program and assessment guidelines, in which students are randomly assigned to sections, with a different teacher in each section; all teachers are assessed through ratings before a final exam; and student academic achievement is evaluated by employing the same or an equivalent final exam. If a student shows better academic achievement due to highly rated teachers, a correlation between sections' average SET and sections' average final exam should be observed (Uttl et al., 2017).

This leads us to consider the appropriate unit of analysis in these types of studies (Cohen, 1981). Some researchers utilize the student as the unit of analysis, relating the student's academic achievement with his/her teacher rating. Other researchers utilize the group class as the unit of analysis, correlating mean group-class achievement with mean class SET. Researchers using individual student data follow a design that allows them to establish whether students who perform better, regardless of the class they attend, score the teachers better. To analyze the association between SET and student academic achievement for respective teachers, the group class (or teacher) must be used as the unit of analysis in the validity design (Cohen, 1981; Abrami et al., 1990; Marsh and Roche, 2000; Clayson, 2005; Richardson, 2005).

Although this solution is widely accepted, criticism has recently emerged. It is argued that the variability between students, despite being averaged, could confuse the variation between group means. Consequently, it may be found that there are no relationships between SET and achievement for individual students, even as the between-class mean data show a significant relationship (Clayson, 2007; Weinberg et al., 2009). It is necessary to use statistical methods that consider both the individual variability within the group class and the variability between group-class means.

Another methodological issue that can affect the results on the relationship between SET and student academic achievement is the number of sections (Cohen, 1981; Uttl et al., 2017). Kulik and McKeachie (1975) indicated that big correlations often appear with small sample sizes, suggesting that to find a stable validity coefficient, at least 30 sections are needed in a multisection study. More recently, Uttl et al. (2017) presented specific results on this topic in their meta-analysis of faculty's teaching effectiveness.

Revision Studies

To answer the question on the relationship between SET and academic achievement, a series of revision and meta-analytical studies have been carried out.

As early as the seventies, many researchers analyzed the association between SET and student achievement. However, as Kulik and McKeachie (1975) pointed out, "the most impressive thing about studies relating class achievement to class ratings of instructors is the inconsistency of the results" (p. 235).

Cohen (1981) performed the first meta-analysis based on 68 multisection studies, in which various equivalent sections/classes follow the same outline and the same or equivalent assessments; each section is instructed by a different professor, and these professors are evaluated using students' evaluation of teaching

ratings. Cohen's (1981) results indicated that SET scores correlated moderately with academic achievement (r = 0.43), concluding that these results support the validity of SET as a measure of teaching effectiveness. However, recent studies have questioned some aspects of Cohen's (1981) meta-analysis, referring to the repeatable search strategy followed by Cohen or the sample size of sections on which Cohen's meta-analysis studies are based, with as few as five sections (Uttl et al., 2017).

The primary objective of Feldman's (1989) meta-analysis was to extend Cohen's analysis of the correlation between several specific dimensions of the evaluation of the teacher's instruction. The four dimensions most correlated with academic achievement were, in this order, preparation and organization, clarity and understandableness, perceived outcome, and teacher's stimulation of interest in the course and its subject matter. Feldman's (1989) results showed that the correlation between preparation and organization, the dimension most strongly correlated with academic achievement, ranged from 0.36 to 0.57. However, this meta-analysis did not account for the size of individual studies, so the moderate to high correlations may be an artifact of small-study effects.

The objectives of Clayson's (2009) meta-analysis were to address situational questions and methodological questions. Criteria for including studies were related to college instruction, data based on multiple sections of the same course discipline, a measure of learning common across sections, a learning measure based on actual testing results and not on student perception, and SET conducted before the students took their final exam. Overall, 17 articles were included, containing 42 studies and 1,115 sections. Considering the situational dependence of previous meta-analysis on educational and/or psychological disciplines, studies were coded according to the subject matter of study.

The raw averaged correlation coefficient between SET and academic achievement was 0.33, whereas the weighted average correlation was 0.13, using between-group-class data. When within-class individual student data were used, this correlation was found to be very close to zero (-0.03). Furthermore, their results also showed a negative relation between Z-transformed r and the size of the sample, indicating that as the number of sections increases, the value of the correlation decreases. A moderator variable was identified; the association was greater in education and liberal arts disciplines, but lower in business classes. The more control was used – for example, considering the effect of previous academic achievement – the less association was found. Clayson (2009) concluded that "a small average relationship exists between learning and the evaluations but that association is situational and not applicable to all teachers, academic disciplines, or levels of instruction" (p. 16).

One of the criticisms of Clayson's (2009) work is that the number of articles included in the previous meta-analysis by Cohen (1981) exceeded 40 articles, while Clayson used 17 articles with 42 multisection studies. In addition, Clayson's meta-analysis was based on different individual multisection studies, mixed in as if it were a multisection study (Uttl et al., 2017).

The most extensive revision work on the relationship between the results of SET and their academic achievement is the one recently carried out by Uttl et al. (2017). On the one hand, they reanalyzed the previous meta-analyses of Cohen (1981); Feldman (1989), and Clayson (2009); on the other hand, they updated the previous meta-analyses of SET/achievement correlations included in multisection studies to date.

Both in the reanalysis of the previous meta-analyses and in Uttl et al.'s (2017) meta-analysis, special attention is paid to the effects of small study size or small number of sections. Furthermore, in this study, correlations weighted by sample size were used, instead of averaged correlations. The third objective was to analyze the effects of prior achievement on the relation between SET and final achievement.

The results of the reanalysis carried out by Uttl et al. (2017) indicate that, in these studies, the moderate SET/achievement correlations are close to zero when the small-study-size effects are considered. As noted by Kulik and McKeachie (1975), large correlations usually appear with small sample sizes; more low correlations are found when larger samples are used.

In the reanalysis of Cohen's (1981) data, Uttl et al. (2017) found that the SET/achievement correlation estimated by using only studies with 30 or more sections was 0.27. The reanalysis of Cohen's (1981) data did not support Cohen's conclusion that SET explains 18–25% of academic achievement variability (mean r = 0.47); instead, Uttl et al. (2017) conclude that SET explains at best 10% of variance in academic performance.

According to Uttl et al. (2017), the reanalysis of Feldman's (1989) meta-analysis also showed that Feldman's results were dependent on small-study effects and that the specific student rating dimensions do no correlate with achievement. Similarly, the reanalysis of Clayson's (2009) work also points out that the correlations estimated were lower than reported, once the small-study effects were considered.

In the updated meta-analysis carried out by Uttl et al. (2017), the overall SET/achievement means correlation was 0.23. The values for correlations adjusted for prior achievement/ability were 0.16 and 0.25, eliminating two studies considered as outliers. In addition, when small sample bias is into account and after outliers are removed, the SET/achievement correlation was 0.08 for all correlations and −0.03 for correlations adjusted for prior ability. Thus, individual differences in knowledge, ability, and motivation influence the academic performance more than teaching ratings did.

In sum, the different analyses carried out by Uttl et al. (2017) – with the assumption of fixed and random effects, with and without prior achievement, with outliers eliminated, and considering or not considering the effect of size – found correlations that varied approximately between 0.08 and 0.30, which were significantly lower than the values found in previous studies.

The Present Study

The present study aimed to check the relationships between SET and academic achievement, starting from the knowledge offered by previous studies. This study is carried out in a different context to most previous works. It is based in the South American country Ecuador and analyzes SET in the National Polytechnic School—a higher education institution for the study of technical subjects, such as engineering, architecture,

and biotechnology. If the association between SET and academic performance is situational and not applicable to all academic disciplines, appearing stronger in studies in the field of education and the liberal arts and less in other areas such as business classes (Clayson, 2009), it seems necessary to carry out new studies, focusing on technical areas different to previous studies where there are fewer studies on the subject.

Although there are no records on the beginning of the evaluation of teachers in higher education in Ecuador, this has been a widespread practice in Ecuadorian higher education institutions since the early 1980s (Pareja, 1986).

The Council of Ecuadorian Higher Education obligates the evaluation of the teaching staff of higher education institutions, both for their entry and for their promotion, in the Career and Ladder Regulations of the Professor and Researcher of the Higher Education System. Teachers' professorships may even be removed if they obtain a negative SET twice consecutively or if they obtain four negative evaluations throughout their careers (Consejo de Educación Superior [CES], 2017).

The variable prior knowledge/ability is found to be a powerful moderator of the relation between SET and academic achievement (Cohen, 1981; Clayson, 2009; Uttl et al., 2017). When prior academic achievement/ability is considered, the correlations between SET and achievement correlation decrease, even coming close to zero. The present study includes a measure of previous academic achievement and statistical procedures that adjust both measures of SET and achievement for prior student achievement. Although prior achievement is one of the variables that most influence the final achievement, this study examines whether SET makes a significant contribution to the final achievement, after the effect of previous achievement is controlled for.

An open methodological question, which seeks to address this study, is the unit of analysis. Most of the researchers in this field use the group-class average as the unit of analysis, arguing that the individual differences within the group class are eliminated and the differences between the means of the group classes, sections, or teachers (Cohen, 1981; Abrami et al., 1990; Marsh and Roche, 2000; Clayson, 2005; Richardson, 2005; Uttl et al., 2017) are clearly reflected; other researchers defend the need to account for the individual variability within the group classes (Clayson, 2007; Weinberg et al., 2009). Some studies in this field have considered both aspects separately (Clayson, 2009), but to our knowledge, none have considered the variability within and between group classes or teachers jointly. In this study, we will use methods that consider both sources of variability, the students and the group class, for multilevel analysis.

In addition, since multisection designs are the ones that offer the most valuable estimate of the relationship between SET and academic achievement, an aggregated data analysis is carried out following the procedure of a multisection design, using the data from course disciplines with two or more sections.

From this theoretical context, the following objectives were established:

(1) Correlate the individual students' teacher ratings and their academic achievement.
(2) Correlate the average of SET in the class-group means with the academic achievement means of each group class.
(3) Examine the relationship of SET with the final academic achievement, once the effect of the prior academic achievement has been controlled for, establishing the specific contribution of SET to the final academic achievement, using the group averages as the unit of analysis.
(4) Evaluate the joint contribution of the individual student and the group class evaluations of teaching to the final academic achievement, considering the previous achievement.
(5) Analyze the relationships between SET, academic achievement, and prior academic achievement, following the procedure of a multisection design, considering those course disciplines or subjects matters in which there are different sections.

MATERIALS AND METHODS

Participants

The sample included 1,538 students of the National Polytechnic School from Ecuador, enrolled in eight different faculties and schools and studying 28 different degrees. Of these students, 68.6% were male and 31.4% were female. The higher percentage of male students is representative of the population of students of polytechnic studies. The average age was 22.3 years ($SD = 3.2$). This sample was chosen from a larger sample of 6,100 students who rated the teachers during the 2016/2017 academic year. These 1,538 students attended 343 different course disciplines and were distributed into 453 class groups. Most of these course disciplines had only one class group, while 48 course disciplines had more than one class group or section (with 776 students in total). The number of sections ranged from 2 to 10, with a total of 158 sections across different course disciplines. The total number of students in the different sections was 776. The teachers' sample consisted of 310 teachers, who represented a varied sample in terms of age, category, and teaching experience. More than half of these teachers were male (62.8%).

Measures

Student evaluation of teaching was obtained from the "Cuestionario de Evaluación de la Enseñanza del Profesor de la Escuela Politécnica Nacional del Ecuador" (Teacher Evaluation Questionnaire of the National Polytechnic School), approved by the teacher staff for the 2016/2017 academic year. The scale consisted of 33 items grouped theoretically into four factors: planning, mastery, and clarity in the explanation of the subject; methodology and resources; teacher – student relationship; and evaluation.

The results of the validation of this questionnaire in a large sample of 6,100 students (Sánchez et al., 2019) showed the permanence of these four theoretical factors in an exploratory factor analysis, with a high reliability of internal consistency – Cronbach's alpha ranged between 0.94 and 0.86 and was 0.96 for

the total scale. The results also show a high correlation between the four factors (0.78–0.88).

Two measures of student academic achievement were taken: previous academic achievement and academic achievement at the end of the semester. Previous accumulated achievements were a measure of the mean academic achievement reached by students on all previous subject matters, among those who were enrolled until the beginning of the current semester. This measure was obtained from computerized administrative records. Although strictly it cannot be considered a measure of prior performance in the particular subject matter, it can be seen as being indicative of the general knowledge or ability with which the student begins the study of the subject.

The measure of academic achievement at the end of the semester was operationalized by grades awarded by the teacher, based on a final exam, consisting of theoretical and practical written examinations. These final exams in some cases were the same across sections and in others were different for different sections. The different sections follow the same program and have the same assessment criteria. These criteria are specified in the study program of each course. There are also common general rules for all exams in the Polytechnic School. The measures of previous accumulated academic achievement and the final grades ranged from 0 to 40 for all courses.

Students' age and gender as well as teachers' age, gender, and experience were collected from administrative records.

Procedure

The data were collected from the existing computer records in the administration of the Polytechnic School and permission was granted for access to the records by the academic staff of the institution. The data provided by the institution were anonymous, with an identification code for each student.

The application of the SET scale was carried out at the end of the semester, before the students knew their final grades. All teachers were evaluated by the students in the same term. All students had to evaluate the teachers to be able to access their final grades. The SET was made through an electronic platform, in which the data were recorded.

The impact of faculty procedures of SET on response rates has been studied by several authors, especially focusing on electronic evaluations. A high response rate is important, which in the field of evaluation in higher education is estimated at 70% (Richardson, 2005). Young et al. (2019) found that the number of responses was significantly higher when students had time in class to complete the evaluation of teaching compared to the electronic form of administration. When the response rate in electronic administration was lower than that with paper-and-pencil questionnaires, this work followed the procedure of forcing all students to answer the evaluation survey in order to access their final grades. This procedure has proved useful and valid in some higher education institutions (Leung and Kember, 2005; Nair and Adams, 2009).

Data Analysis

The data analysis was performed according to the design and goals of this research.

On the one hand, average class group was employed as a unit of analysis; on the other, the individual data of the students were analyzed.

When the class-group average was employed as the unit of analysis, a correlation analysis and a hierarchical regression analysis were performed. Correlation analysis was calculated with Pearson's product–moment correlation technique. The linear hierarchical multiple regression analysis included, in the first step, prior academic achievements and, in the second step, SET. This methodological approach establishes the specific contribution of a variable, which enters last in the analysis, to the prediction of the dependent variable – in this case, the academic achievement at the end of the semester. In addition, the extra amount of variance accounted for in the final academic achievement by SET can be estimated (Cohen and Cohen, 1983).

A multilevel path analysis was performed on the individual data, grouped into sections. This analysis accounts jointly for the variability among individual students within the class groups (level 1) and the variability between groups, taught by different teachers (level 2). A path analysis is established in which the influence of previous academic achievement on the final academic achievement and on SET is examined and in which the relation of SET with the final academic achievement is also included. All variables were observed; no latent variable was defined.

The program used was the structural equation modeling (EQS) by Bentler (2005). Parameter estimation was conducted on the basis of maximum likelihood (ML); ML estimation is based on the characteristics of multivariate normality that are used to produce optimal estimates of the population parameters, and thus, it requires relatively large sample sizes. Implementation of a diversity of fit indices is recommended when evaluating the model fit, including chi-square, chi-square relative to the degree of freedom, standardized root mean square residual (SRMR), root mean square error of approximation (RMSEA), and the comparative fit index (CFI) (Hu and Bentler, 1999).

The analysis of grouped data, although it may be considered more appropriate than the analysis of individual data (Cohen, 1981), raises some important methodological questions. An analysis of class groups mixing different course disciplines or subject matter and sections of the same courses raises questions about the validity of correlation coefficients estimated from a pooling of heterogeneous microarray data (Hassler and Thadewald, 2003; Almeida-de-Macedo et al., 2013). The effect of heterogeneous variance–covariances across a pool of data causes less efficient estimates of Pearson correlation coefficients across groups than does the approach of combining correlation coefficients of individual groups.

To overcome this question, an aggregated data analysis is carried out following the procedure of a multisection design, using the data from course disciplines with two or more sections. To consider the small-sample bias effect, correlations weighted by simple size were used.

RESULTS

The results presented are divided into two sections – those related to the aggregated data and those related to individual data – that consider the hierarchical nature of the data for the multilevel path analysis.

Average Group as Unit of Analysis

The data of the 1,538 students were averaged across the 453 class groups, from the same or different course disciplines.

Table 1 shows correlations between the mean group prior academic achievement, the mean group SET, and the mean group final academic achievement.

As **Table 1** shows, statistically significant correlations between mean prior academic achievement and mean final academic achievement were identified, as well as between mean SET ratings and final academic achievement. Prior academic achievement was not statistically correlated with SET.

To determine the specific contribution of SET on final academic achievement, a hierarchical multiple regression analysis was performed, in which independence of residuals was estimated (Durbin–Watson = 2.02).

A hierarchical linear regression analysis (see **Table 2**) was conducted in which prior academic achievement was entered in step 1 and SET in step 2.

Model 1 was significant (R^2 = 0.27, F = 145.95), and prior academic achievement significantly predicted the final academic achievement (β = 0.52, $p < 0.001$). In the second step (model 2), SET significantly predicted final academic achievement (β = 0.26, $p < 0.001$), beyond the effect of prior academic achievement. This model explained 34% of the variance of final performance.

The change between model 1 and model 2 was statistically significant (ΔR^2 = 0.07, F = 100.42, $p < 0.001$), indicating that the specific proportion of variance in final academic achievement accounted for by SET was 7%, and it is statistically significant.

TABLE 1 | Correlations between variables with data grouped into class groups.

Variable	1	2	3	*M*	*SD*
1. Prior achievement	1			25.48	4.31
2. SET ratings	03	1		4.01	0.70
3. Final achievement	0.52**	0.28**	1	27.79	6.71

*N = 453. **p < 0.01.*

TABLE 2 | Hierarchical regression of prior academic achievement and student evaluation of teaching (SET) on academic achievement.

Variable	*B*	*SE B*	β
Step 1			
Constant	7.26	1.72	
Prior achievement	0.81	0.07	0.52**
R^2			0.27
F			145.95**
Step 2			
Constant	−2.42	1.72	
Prior achievement	0.79	0.06	0.51**
SET	2.51	0.40	0.26**
R^2			0.34
ΔR^2			0.07
ΔF			100.42**

*N = 453. **p < 0.001.*

Individual Student as Unit of Analysis

Correlations between student prior academic achievement, SET, and student final academic achievement are shown in **Table 3**.

The results of individual students were similar, although slightly lower, to those averaged by groups. Statistically significant correlations were found between individual students' prior achievements and individual students' final academic achievement, as well as between SET and final academic achievement. Prior academic achievement was not statistically correlated with SET.

As individual students were grouped into class groups, a multilevel structural equation analysis with observed variables was performed, with individual students within the section as level 1 and the difference between groups as level 2. The total student sample was 1,538, distributed into 453 class groups.

The model tested the influence of previous academic achievement on final academic achievement and SET, as well as the influence of SET on final academic achievement. **Figure 1** shows the model and results of the multilevel structural analysis.

The ML method was employed for parameter estimation. This method assumes multivariate normal distributions, although the method of ML is robust for departures from normality, especially if the sample is large and the skewness is <2 and kurtosis <7, in absolute terms (West et al., 1995) – values that are below those obtained in this work.

Once the model displayed in **Figure 1** includes relationships between all the variables, it is a saturated model in which the number of parameters to estimate is equal to the data; since it makes theoretical sense to consider the similarity of the individual (within) and section (between) parameters, the three path coefficients were constrained to be equals.

This model provided a very good fit to the data (Bentler CFI = 0.996, χ^2 = 4.89, df = 3, p = 0.18; McDonald's MFI = 0.999; SRMR = 0.020: RMSEA = 0.030) (see **Table 4**).

Furthermore, for the test of equivalence of path coefficients across levels, the EQS reported a cumulative multivariate Lagrange multiplier (LM) test (χ^2) and an incremental univariate χ^2 value, along with their probability values, for each constraint. To find non-invariant parameters across groups, the probability associated with the incremental univariate χ^2 values of <0.05

TABLE 3 | Correlations between student individual variables.

Variable	1	2	3	*M*	*SD*	Skewness	Kurtosis
1. Prior achievement	1			5.88	5.17	−0.98	1.95
2. SET ratings	0.03	1		3.99	0.70	−1.23	2.16
3. Final achievement	0.50**	0.23**	1	28.08	6.71	−1.60	0.96

*N = 1,538. **p < 0.01.*

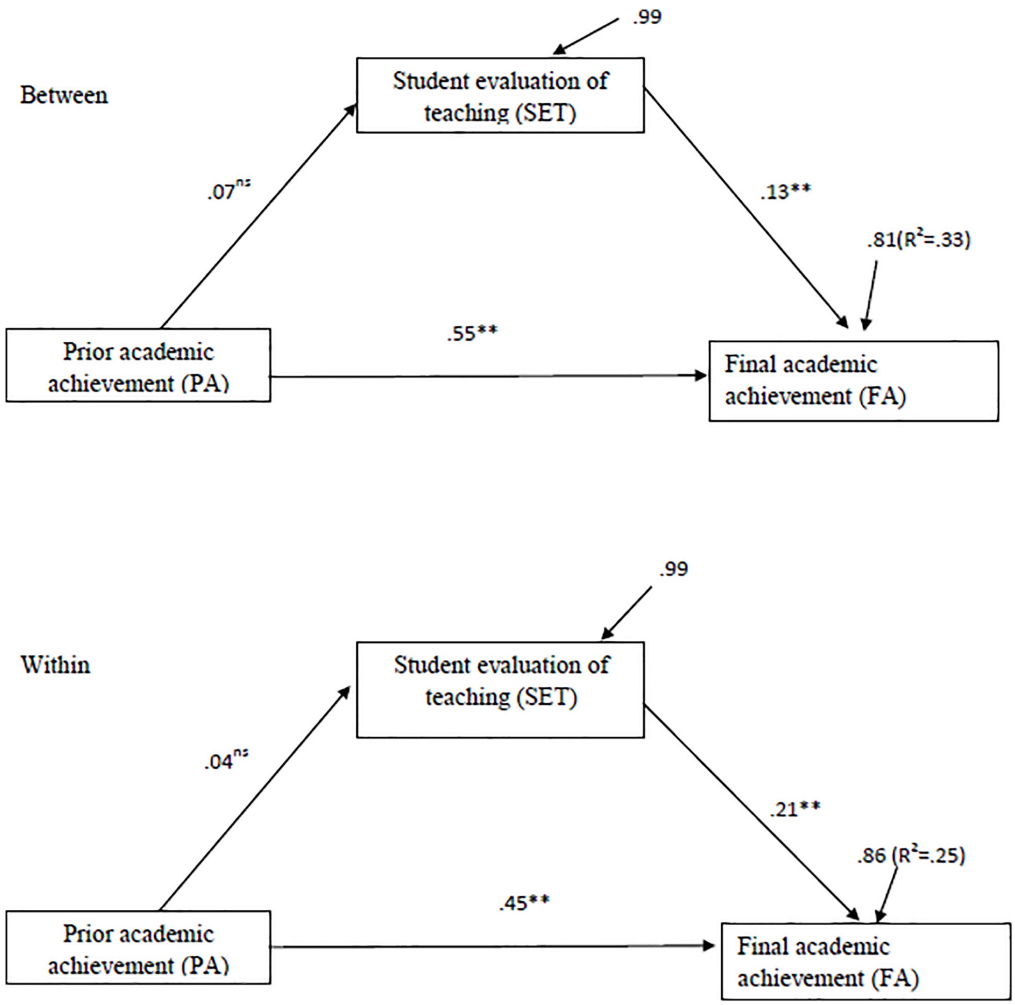


FIGURE 1 | Standardized coefficients of the multilevel model with between prior academic achievement (V1), SET (V3), and academic achievement (V2). Path coefficients constrained to be equal. $^{**}p < 0.01$; ns = not significant.

(Byrne, 2008) was checked; none of the equality constraints were significant (V3, V2, $p = 0.30$; V3, V4, $p = 0.36$; and V4, V2, $p = 0.46$), indicating the equivalence of the three coefficients across levels.

The relationships between the observed variables proposed in the model were significant ($p < 0.05$), except for the effects generated by prior academic achievement on SET. Both at the individual (within) and at the section levels (between), the highest regression coefficient was prior academic achievement on final academic achievement ($\beta = 0.45$, $p < 0.01$ for level 1; $\beta = 0.55$, $p < 0.01$ for level 2). SET also has an effect on final academic achievement, at both the student level ($\beta = 0.21$, $p < 0.01$) and group level ($\beta = 0.13$, $p < 0.01$). Conversely, prior academic achievement was not statistically related to SET, either at the individual level ($\beta = 0.04$, $p > 0.05$) or at the group level ($\beta = 0.07$, $p > 0.05$).

The total percentage of variance explained from the final academic achievement at the level of the students was 25%, while at the level of the sections, it was 33%.

TABLE 4 | Mean correlations between variables estimated with data grouped into sections.

Variable	1	2	3
1. Prior achievement	1		
2. SET	0.09	1	
3. Final achievement	0.16*	0.26**	1

*N = 150. *p < 0.05; **p < 0.01.*

Multisection Design Analysis

Sections' average SET, sections' average prior achievement, and sections' average academic achievement were correlated for each course discipline; then the mean of the correlations weighted by the sample size was estimated. Specifically, we transform *r*s to Fisher's *Z* scores, calculating average Fisher's *Z* scores across all course disciplines and weighing *Z*s by each sample size, and transformed average Fisher *Z* scores back to *r*. These results are shown in **Table 4**.

Again, there will be a moderate but significant correlation between the previous academic achievement and the final academic achievement, although in this case with a lower value; there was also an average positive correlation between the means of the SET and academic achievement, based on the means of each section weighted by the sample size. SET averages were not related to the prior performance.

To control for the effect of prior achievement on the relationship between SET and academic achievement, the partial correlation between the means of the sections within each course discipline of the SET and the final achievement means was estimated, considering the means of the previous achievement. Then the mean of the partial correlations weighted by the sample size was estimated. The average value of the partial correlation coefficient between SET and final achievement, considering the effect of the previous achievement, estimated in the average of the different course disciplines, was $r = 0.22$.

To examine the effect of small samples in multisection studies, the correlation between the number of sections and the absolute value of the correlation between SET and final achievement was calculated, obtaining $r = -0.18$, indicating that there is a tendency to obtain higher correlations when these correlations are based on a smaller number of sections.

DISCUSSION

This work aimed to clarify several of the issues raised about SET as a measure of teacher effectiveness. For this, a large number of individual students and group class were included; a multisection design was used when course disciplines had more than one class group; previous academic performance was considered, since the random allocation of students to the sections was not assured; and statistical methods were used which consider both the individual student variability within sections and the variability between sections. Furthermore, the study was carried out in a geographical and disciplinary context different from that of most previous studies.

The results obtained with aggregated data, taking the group class as the unit of analysis, showed a moderate but statistically significant correlation (0.28) between SET and final academic achievement. This value corresponds to the value obtained in the meta-analysis of Uttl et al. (2017) when the data of Cohen (1987) were reanalyzed considering small-sized studies and effects (i.e., only the studies with a number of 30 or more sections).

These results also showed a moderately high correlation between prior academic achievement and final academic achievement. This finding is in accordance with previous meta-analytic studies on the variables associated with achievement in higher education, in which prior knowledge/abilities appear as one of the main determinants of academic achievement (Schneider and Preckel, 2017).

However, the correlation between prior achievement and SET was not statistically significant, suggesting that SET is not affected by previous academic achievements.

Control for prior academic achievement with the hierarchical regression analysis procedure continued to show a significant effect of SET on academic achievement; this effect was around 7%, which corresponds to a correlation of 0.27, similar to that found in the reanalysis of Cohen's (1981) data, and is slightly higher than the value obtained in the meta-analysis of Uttl et al. (2017) based on nearly 100 multisection studies published to that date, which stood at 0.23.

The results obtained with the individual student data showed a statistically significant correlation (0.23) between SET and final academic achievement, which was a bit lower than that obtained with the data aggregated in sections. This result is consistent with previous studies about instructor's teaching effectiveness, in which it is considered that multisection studies that use the grouped data of the sections are more appropriate to apprehend the true relationship between SET and academic achievement (Cohen, 1981; Uttl et al., 2017).

The results of individual data showed again a moderately high correlation between prior academic achievement and final academic achievement, as well as a non-significant relation of SET with prior academic achievement.

Following the suggestion of several authors regarding these types of studies, both the individual variability within the sections and the variability between sections (Clayson, 2007; Weinberg et al., 2009) of the data of the present work included a multilevel structural equation analysis.

The results of the multilevel analysis showed that there was a significant effect of SET on the final academic achievement, at both the individual and the section levels, even after controlling the effect of prior academic achievement. In addition, the magnitude of the effect was similar in both levels. The total percentage of variance explained from the final academic achievement at the level of the sections was 33%, while at the level of the individual students, it was 25%, with 8% of the explained variance of final academic achievement attributable to the sections: that is, to the effect of the teacher.

The results obtained with aggregated data, taking the section as the unit of analysis, following the guidelines of a multisection design, show that a significant, although low, relationship remains between SET and academic achievement when the sample size effect is considered ($r = 0.26$), even when the effect of the prior academic achievement is controlled ($r = 0.22$). Therefore, the results of the individual and the group analyses do not differ substantially from the results obtained in the analysis of the sections, supporting partially the results of the individual analysis and aggregated group analysis, in which biased correlations could appear due to pooling of heterogeneous samples, when the analysis of the data is carried out following the guidelines of a multisection design.

These results were similar to those found in studies carried out in different geographical and disciplinary contexts. The study was conducted in a Higher Polytechnic School of Ecuador, which teaches scientific and technological disciplines, which are different from the humanistic and social disciplines rated in most of the studies on teaching effectiveness (Clayson, 2009).

On the basis of the large-scale datasets from Australia, Canada, and the United States ($N = 26{,}746$ students) in the Programme for International Student Assessment (PISA), 2012, Scherer et al. (2016) find support for significant relations to the educational outcomes. Students' achievement could be best predicted by perceived classroom management ($\beta = 0.20$ to 0.31).

Together, the results show the relation between SET and academic achievement, in a study where multiple sections are included, controlling previous academic achievement and considering both the student variability within sections and the variability between sections with different teachers, in subject matters of a scientific – technological nature.

However, the amount of influence of SET on academic achievement is lower than that found in some previous meta-analytic studies (Cohen, 1981; Feldman, 1989), but higher than that found in the meta-analysis of Uttl et al. (2017) carried out on the multisection studies published to that date; when small-study-size effects and prior academic achievement were considered, it was close to zero.

Although university student academic achievement depends mainly on various intellectual and non-intellectual factors (Richardson et al., 2012; Schneider and Preckel, 2017), the results of this work support the conclusion that SET has a modest, around 5%, but significant influence on academic achievement and is therefore related to teacher effectiveness.

However, taking into consideration our results and the results of previous meta-analyses, especially the comprehensive meta-analysis of Uttl et al. (2017), the influence of SET on academic achievement seems to be sufficiently limited to make relevant administrative decisions. Although use of SET as a feedback for teachers' use and as a measure of student satisfaction is not problematic (Spooren et al., 2013; Uttl et al., 2017), the use of SET as a measure of teachers' effectiveness for making administrative decisions about teachers' hiring, firing, promotions, and merit pay is controversial (Uttl et al., 2017, 2019; American Sociological Association, 2019).

Limitations

The analysis that takes into account individual student and average group as units of analyses, mixing different subject courses and sections of the same courses, raises questions about the validity of correlation coefficients estimated from pooling heterogeneous microarray data, given that it causes less efficient estimates of Pearson correlation coefficients than does the approach of combining correlation coefficients of individual groups, as is done in the analysis that follows a multisection design, although, on the other hand, and the results obtained from the multisection analysis are consistent with the individual and group analyses.

Final exams in some cases were the same across sections; however, in others, they were not identical for different sections; although different sections follow the same program and have the same assessment criteria, the exams should be identical or equivalent, as required for a multisection study.

This study uses a low number of sections, ranging from 2 to 10, which can lead to the small section size effect, given the tendency to obtain higher correlations when these correlations are based on a smaller number of sections.

Prior academic achievement in the subject matter was not measured; the measure was of the accumulated academic performance in all subject matters in which the student had been enrolled before the beginning of the semester. However, in scientific–technological disciplines, the academic achievement accumulated previously is a measure that is usually related to the final achievement, and it also seems to be an adequate measure to study the possible influence on SET.

Another question that arises in relation to this study is the procedure of obtaining the SET. Although research shows that, in general, electronic evaluation procedures are as valid as traditional procedures (Spooren et al., 2013), more research is necessary on this procedure of forcing all the students to answer the evaluations of teaching, in terms of social desirability, acquiescence, and stereotyped answers, etc. From a methodological perspective, in the path analysis, all the variables are observed variables and not latent; therefore, the measurement error could not be estimated.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Data collection was made from existing computer records of the Polytechnic School administration, and the academic staff of the institution granted access to them. The data provided by the institution were anonymous, with only one identification code for each student.

AUTHOR CONTRIBUTIONS

TS: theoretical review of the study. RG-C: quantitative methods and theoretical review of the study. J-LC and JL: quantitative methods. JV: data collection and review of the references.

FUNDING

This research was supported by National Secretariat of Higher Education, Science and Technology, SENESCYT (PIC-18-INE-EPN-002).

REFERENCES

Abrami, P. C., d'Appolonia, S., and Cohen, P. A. (1990). Validity of student ratings of instruction: what we know and what we do not. *J. Educat. Psychol.* 82, 219–231. doi: 10.1007/s10459-017-9783-0

Almeida-de-Macedo, M., Ransom, N., Feng, Y., Hurst, J., and Wurtele, E. S. (2013). Comprehensive analysis of correlation coefficients estimated from pooling heterogeneous microarray data. *BMC Bioinformatics* 14:214. doi: 10.1186/1471-2105-14-214

American Sociological Association, (2019). *Statement on Student Evaluations of teaching*. Avaliable at: https://www.asanet.org/sites/default/files/asa_statement_on_student_evaluations_of_teaching_sept52019.pdf (Septembre 9, 2019).

Bentler, P. M. (2005). *EQS 6 Structural Equations Program Manual*. Encino, CA: Multivariate Software.

Byrne, B. M. (2008). *Structural Equation Modeling with EQS: Basic Concepts, Applications and Programming*. London: Routledge.

Clayson, D. E. (2005). Within-class variability in student-teacher evaluations: example and problems. Decision. *Sci. J. Inno. Educ.* 3, 109–124. doi: 10.1111/j.1540-4609.2005.00055.x

Clayson, D. E. (2007). Conceptual and statistical problems of using between class data in educational research. *J. Mark. Educ.* 27, 122–129. doi: 10.1002/mono.12060

Clayson, D. E. (2009). Student evaluations of teaching: are they related to what students learn? A meta-analysis and review of the literature. *J. Mark. Educ.* 31, 16–30. doi: 10.1177/0273475308324086

Cohen, J., and Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cohen, P. A. (1981). Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies. *Rev. Educ. Res.* 51, 281–309. doi: 10.3102/00346543051003281

Cohen, P. A. (1987). "A Critical analysis and reanalysis of the multisection validity meta-analysis," in *Paper Presented at the Annual Meeting of the American Educational research Association*, Washington, DC). doi: 10.3102/00346543051003281

Consejo de Educación Superior [CES], (2017). *Reglamento de Carrera y Escalafón del Profesor e Investigador del Sistema de Educación Superior. [Career and Ladder Regulations of the Professor and Researcher of the Higher Education System]*. Avaliable at: https://procuraduria.utpl.edu.ec/sitios/documentos/NormativasPublicas/Reglamento%20de%20Carrera%20y%20Escalaf%C3%B3n%20del%20Profesor%20e%20Investigador%20del%20Sistema%20de%20Educaci%C3%B3n%20Superior%202018.pdf (20 April 2019).

Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: refining and extending the synthesis of data from multisection validity studies. *Res. Higher Educ.* 30, 583–645. doi: 10.1007/bf00992392

Hassler, U., and Thadewald, T. (2003). Nonsensical and biased correlation due to pooling heterogeneous samples. *Statistician* 52, 367–379. doi: 10.1111/1467-9884.00365

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Husbands, C. T., and Fosh, P. (1993). Students' evaluation of teaching in higher education: experiences from four european countries and some implications of the practice. *Assess. Eval. Higher Educ.* 18, 95–114. doi: 10.1080/0260293930180202

Huybers, T. (2014). Student evaluation of teaching: the use of best–worst scaling. *Assess. Eval. Higher Educ.* 39, 496–513. doi: 10.1080/02602938.2013.851782

Kulik, J. A., and McKeachie, W. J. (1975). "The evaluation of teachers in higher education," in *Review of Research in Education*, Vol. 3, ed. F. N. Kerlinger, (Itasca: Peacock), 201–240.

Leung, D. Y. P., and Kember, D. (2005). Comparability of data gathered from evaluation questionnaires on paper and through the Internet. *Res. Higher Educ.* 46, 571–591. doi: 10.1007/s11162-005-3365-3

Leventhal, L. (1975). Teacher rating forms: critique and reformulation of previous validation designs. *Can. Psychol. Rev.* 16, 269–276. doi: 10.1037/h0081814

Marsh, H. W. (2007). Do university teachers become more effective with experience? A multilevel growth model of students' evaluation of teaching over 13 years. *J. Educ. Psychol.* 99, 775–790. doi: 10.1037/0022-0663.99.4.775

Marsh, H. W., and Overall, J. U. (1980). Validity of students' evaluations of teaching effectiveness: cognitive and affective criteria. *J. Educ. Psychol.* 72, 468–475. doi: 10.1037/0022-0663.72.4.468

Marsh, H. W., and Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: popular myth, bias, validity, or innocent bystanders? *J. Educ. Psychol.* 92, 202–228. doi: 10.1037/0022-0663.92.1.202

Nair, C. S., and Adams, P. (2009). Survey PLATFORM: a factor influencing online survey delivery and response rate. *Q. Higher Educ.* 15, 291–296. doi: 10.1080/13538320903399091

Pareja, F. (1986). *La Educación Superior en el Ecuador [The higher education in Ecuador]*. Caracas: Regional Center For Higher Education in Latin America And the Caribbean (CRESALC)-UNESCO.

Richardson, J. T. E. (2005). Instruments for obtaining student feedback: a review of the literature. *Assess. Eval. Higher Educ.* 30, 387–415. doi: 10.1080/02602930500099193

Richardson, M., Abraham, C., and Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychol. Bull.* 138, 353–377.

Sánchez, T., Sandoval, I., Salazar, D., Gilar, R., and Castejón, J. L. (2019). "Validation of the teacher evaluation questionnaire of the National Polytechnic School, applying the method of factor analysis with extraction of principal components," in *17th LACCEI International Conference for Engineering, Education, and Technology, 24-26 July 2019*, Jamaica.

Scherer, R., Nilsen, T., and Jansen, M. (2016). Evaluating individual students' perceptions of instructional quality: an investigation of their factor structure, measurement invariance, and relations to educational outcomes. *Front. Psychol.* 7:110. doi: 10.3389/fpsyg.2016.00110

Schneider, M., and Preckel, F. (2017). Variables associated with achievement in higher education: a systematic review of meta-analyses. *Psychol. Bul.* 43, 565–600. doi: 10.1037/bul0000098

Spooren, P., Brockx, B., and Mortelmans, D. (2013). On the validity of student evaluation of teaching: the state of the art. *Rev. Educ. Res.* 83, 598–642. doi: 10.3102/0034654313496870

Uttl, B., Cnude, K., and White, C. A. (2019). Conflict of interest explain the size of student evaluation of teaching and learning correlations in multidrction studies: a meta-analysis. *PeerJ* 7, e7225. doi: 10.7717/peerj.7225

Uttl, B., White, C. A., and Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: student evaluation of teaching ratings and student learning are not related. *Stud. Educ. Eval.* 54, 22–42. doi: 10.1016/j.stueduc.2016.08.007

Weinberg, B. A., Hashimoto, M., and Fleisher, B. M. (2009). Evaluating teaching in higher education. *J. Econ. Educ.* 40227– 61,

West, S. G., Finch, J. F., and Curran, P. J. (1995). "Structural equation models with non-normal variables," in *Structural Equation Modeling: Concepts, Issues, and Applications*, ed. R. H. Hoyle, (Thousands, CA: Sage), 56–75.

Young, K., Joines, J., Standish, T., and Gallagher, V. (2019). Student evaluations of teaching: the impact of faculty procedures on response rates. *Assess. Eval. Higher Educ.* 44, 37–49. doi: 10.1186/s12909-015-0387-1

Zabaleta, F. (2007). The use and misuse of student evaluation of teaching. *Teach. Higher Educ.* 12, 55–76. doi: 10.1080/13562510601102131

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.

Copyright © 2020 Sánchez, Gilar-Corbi, Castejón, Vidal and León. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 25 March 2020
doi: 10.3389/fpsyg.2020.00431

How Schools Affect Student Well-Being: A Cross-Cultural Approach in 35 OECD Countries

Elena Govorova[1], Isabel Benítez[2] and José Muñiz[1]*

[1] Department of Psychology, University of Oviedo, Oviedo, Spain, [2] Department of Methodology of Behavioural Sciences, University of Granada, Granada, Spain

A common approach for measuring the effectiveness of an education system or a school is the estimation of the impact that school interventions have on students' academic performance. However, the latest trends aim to extend the focus beyond students' acquisition of knowledge and skills, and to consider aspects such as well-being in the academic context. For this reason, the 2015 edition of the international assessment system Programme for International Student Assessment (PISA) incorporated a new tool aimed at evaluating the socio-emotional variables related to the well-being of students. It is based on a definition focused on the five dimensions proposed in the PISA theoretical framework: cognitive, psychological, social, physical, and material. The main purpose of this study is to identify the well-being components that significantly affect student academic performance and to estimate the magnitude of school effects on the well-being of students in OECD countries, the school effect being understood as the ability of schools to increase subjective student well-being. To achieve this goal, we analyzed the responses of 248,620 students from 35 OECD countries to PISA 2015 questionnaires. Specifically, we considered non-cognitive variables in the questionnaires and student performance in science. The results indicated that the cognitive well-being dimension, composed of enjoyment of science, self-efficacy, and instrumental motivation, as well as test anxiety all had a consistent relationship with student performance across countries. In addition, the school effect, estimated through a two-level hierarchical linear model, in terms of student well-being was systematically low. While the school effect accounted for approximately 25% of the variance in the results for the cognitive dimension, only 5–9% of variance in well-being indicators was attributable to it. This suggests that the influence of school on student welfare is weak, and the effect is similar across countries. The present study contributes to the general discussion currently underway about the definition of well-being and the connection between well-being and achievement. The results highlighted two complementary concerns: there is a clear need to promote socio-emotional education in schools, and it is important to develop a rigorous framework for well-being assessment. The implications of the results and proposals for future studies are discussed.

Keywords: well-being, school effectiveness, Programme for International Student Assessment, science, hierarchical linear modeling

OPEN ACCESS

Edited by:
Juan Luis Castejon, University of Alicante, Spain

Reviewed by:
Pamela Woitschach, University of British Columbia, Canada
Antonio Valle, University of A Coruña, Spain

***Correspondence:**
Isabel Benítez
ibenitez@ugr.es

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

Received: *27 October 2019*
Accepted: *24 February 2020*
Published: *25 March 2020*

Citation:
Govorova E, Benítez I and Muñiz J (2020) How Schools Affect Student Well-Being: A Cross-Cultural Approach in 35 OECD Countries. Front. Psychol. 11:431. doi: 10.3389/fpsyg.2020.00431

INTRODUCTION

The effectiveness of an education system or a school is generally measured in terms of the impact that school interventions have on student performance, with the prevalent focus being on the cognitive elements, and mostly those associated with the requirements of the academic curriculum or competence areas. Following the definition of Murillo (2005), a school is considered effective when it achieves the maximum holistic development of every one of its students, and especially when this development is greater than might be expected considering the student's previous performance and/or the social, economic, and cultural situation of his/her family. Although a student's development is expected to be comprehensive, school effectiveness is traditionally estimated only through student attainment measures, such as the number or percentage of students who graduate (Grosskopf et al., 2014; Podinovski et al., 2014), standardized test scores in various subjects (Crespo-Cebada et al., 2014; Johnson and Ruggiero, 2014), scores on international and national assessments, or the percentage of students progressing to higher or further education (OECD, 2008c). However, it could be argued that the "results" of a school in terms of non-academic achievement should also be considered as educational objectives given that students with low levels of well-being are more likely to have a negative experience of school, as well as to suffer from depression and be involved in substance abuse or delinquency (Sun and Shek, 2010). As a result of the shared concerns of educational communities and families around the world, the latest trends aim to extend the focus of school effectiveness research beyond simple cognitive performance and also examine aspects such as well-being in the academic context. Some studies have focused on the effect of school on socio-emotional factors such as attitude to learning or academic self-concept (Opdenaker et al., 2002; Murillo and Hernández-Castilla, 2011; Belfi et al., 2012), although the results are not conclusive. Aware of the importance of socio-emotional development as an inseparable element of the integral learning process, the 2015 edition of the Programme for International Student Assessment PISA) incorporated a new instrument aimed at evaluating the socio-emotional variables related to the well-being of students, making it possible to assess school effectiveness in terms of improvements in student well-being at the international level.

Well-Being

In recent years, the importance of well-being and the quality of life concept has grown and has extended into many areas. There are numerous definitions of these, and other terms such as satisfaction and happiness, that, as Veenhoven (2000) points out, have traditionally been used interchangeably. There is, however, nowadays consensus that quality of life refers to both objective and subjective elements and reflects both the living conditions and the perceptions of individuals (Casas, 2004). Moyano-Díaz and Ramos-Alvarado (2007) also assume an integrative perspective based on a model where the quality of life measure is divided into an objective component that refers to a person's ability to access goods and services and a subjective one that incorporates the concept of subjective well-being, which, in turn, is divided into a cognitive and an affective component. In this case, the cognitive focuses on satisfaction (both global and in terms of specific domains), while the affective includes both positive and negative affects.

Assessing the impact of well-being on academic performance has also been the objective of several studies, the results of which have been equally diverse. For example, Novello et al. (1992) proposed a possible relationship between health and performance in which well-being seemed to play an important role, and, in the same vein, Berger et al. (2011) found, through a multilevel analysis, a relationship between socio-emotional well-being, well-being, self-esteem, social integration, positive perception of a school's ambience, and performance. Gutman and Vorhaus (2012) also found, in a longitudinal study, significant correlations between four dimensions of well-being (emotional, behavioral, social, and school) and performance. However, El Ansari and Stock (2010) found that the relationship between health, educational performance, and well-being, the latter operationalized in terms of motivation and satisfaction with the educational experience, was reciprocal.

However, the concept of well-being in childhood and adolescence in itself has been studied extensively (Casas, 2010). For instance, Pollard and Lee (2003) carried out a systematic review where they found that, although well-being has not been defined consistently and there is no agreement on the best way to measure it, five key dimensions are usually addressed (albeit not usually all at the same time), namely, physical, psychological, cognitive, social, and economic well-being. The physical dimension refers to health and physical habits; the psychological to emotions and mental health (often operationalized by the "absence" of negative indicators); the cognitive to intellectual and school-related elements; the social to relationships with others, support, and interpersonal or communicative skills; and the economic to economic resources of the family.

An international survey, PISA, in its addition of 2015, adopted a comprehensive model in the evaluation of well-being, which incorporates, in addition to the habitual evaluation of performance, items and scales aimed at measuring well-being. The PISA 2015 assessment formulates a model including indicators of five dimensions of well-being: psychological, social, physical, material, and cognitive (Borgonovi and Pál, 2016). The model differs from the proposals described above by incorporating in the material dimension aspects related to educational and cultural resources.

In the present study, we use the definition of well-being from the PISA theoretical framework, which describes it as "*a dynamic state characterised by students experiencing the ability and opportunity to fulfil their personal and social goals. It encompasses multiple dimensions of students' lives, including: cognitive, psychological, physical, social and material. It can be measured through subjective and objective indicators of competencies, perceptions, expectations and life conditions*" (Borgonovi and Pál, 2016).

Furthermore, the OECD has published recently the unified framework for the assessment of social and emotional skills (Kankaraš and Suárez-Álvarez, 2019), one of the fundamental

pillars of well-being, which reiterates the importance of socio-emotional development of individuals, crucial for students' maturity. The OECD defines these skills as "...individual capacities that can be (a) manifested as consistent patterns of thoughts, feelings and behaviors, (b) developed through formal and informal learning experiences, and (c) important drivers of socioeconomic outcomes throughout the individual's life" (OECD, 2015).

School Effectiveness

School effectiveness has been examined in hundreds of studies since the publication of the Coleman Report in 1966 (Coleman et al., 1966). The conclusions of this report highlighted the low impact of school factors on student performance in comparison with the strong effect exerted by family socioeconomic context, which educational institutions were ill-equipped to counter. That said, Coleman did also offer the first estimations of school effects, finding that the educational institution explains from 5 to 9% of the variance in mathematics results. Since then, a significant amount of work has been carried out that aims to identify the various factors related to performance and to quantify the magnitude of school effects on students' results (MacBeath and Mortimore, 2001; Hanushek and Luque, 2003; Scheerens and Demeuse, 2005). Teddlie et al. (2000), in the *International Handbook of School Effectiveness Research*, summarized the most important evidence in this field, concluding that there is great variation in estimates of school effectiveness between countries and depending on the methodological approach taken. In general, 5–35% of the variance in academic achievement results between schools is explained by educational policies and practices, a school's atmosphere, and learning climate, depending on the study involved (Martínez-Arias, 2009).

Studies that focus on school effectiveness in terms of the promotion of non-cognitive variables are much less common, although there are some notable exceptions. Murillo and Hernández-Castilla, 2011 performed a cross-country study in Latin America and Spain to estimate the magnitude of school, classroom, and country effects for non-cognitive variables such as self-concept, classroom behavior, social coexistence, and students' satisfaction with their school. Belfi et al. (2012) conducted a literature review of the influence of class composition (ability and gender) in secondary education on students' school well-being and academic self-concept. Lazarides and Buchholz (2019) studied the relationship between student-perceived teaching quality in mathematics classrooms and enjoyment, anxiety, and boredom, at both student and classroom levels, and estimated that these parameters accounted for 4–10% of school effects depending on the variable. Other studies in this area include those by Grisay (1996); Opdenakker and Van Damme (2000), Opdenaker et al. (2002); Sammons (1999), and Vandenberghe et al. (1994), all of which report schools' minimal impact on non-cognitive educational results and attribute less than 5% of variation to the educational institution.

The present study has two main objectives. The first is to identify the well-being components that significantly affect student academic performance. The second consists in estimating the magnitude of school effects on the well-being of students in the OECD countries, where school effect is understood as the ability of schools to increase students' subjective evaluation of their well-being. In addition, the relationship between socio-emotional variables and student- and school-level factors is examined.

METHOD

Sample

The PISA database developed by the OECD is the main source of information used in this study. PISA aims to evaluate the knowledge and skills acquired by students at the end of compulsory education in OECD member countries (35 countries at the time of the 2015 PISA report) and in non-member countries that have joined the project. The test systematically evaluates three areas of knowledge, reading, mathematics, and science. PISA evaluations are organized in such a way that in each cycle (PISA evaluations are carried out every 3 years), one of the evaluation areas is examined in depth. PISA 2015, the sixth edition of the study, focused on science achievement. In the present study, the full data set from all the OECD countries has been used, which comprises data collected from 248,620 15-year-old students. The summed data of all OECD countries were used to obtain the total OECD results, and the individual country data sets were employed for cross-country analysis. **Table 1** reflects sample configuration by country (sample size and percentage of girls), along with the country abbreviation used throughout the study.

Instruments

The cognitive test in PISA 2015 aimed to evaluate the level of acquisition of competences in science, reading, and mathematics, and the student questionnaire collected information about the students themselves, their family background, and school and learning environment. Additionally, school principals completed a questionnaire about the school, its resources, and management practices, and in some countries, optional teacher and parent questionnaires were also used. In this study, only the data relating to the student and school questionnaires as well as the performance test results were analyzed since the teacher and parent data are not available for many OECD countries.

The cognitive performance scale in PISA has become a worldwide reference as it is based on internationally agreed-upon theoretical frameworks. PISA uses the concept of competences, which in this context refers to the ability of students to extrapolate what they have learned and apply their knowledge and skills in real-life situations, as well as their ability to analyze, reason, and effectively communicate their findings and interpret and solve problems in different situations. The full PISA cognitive performance test comprises 528 questions about science, mathematics, reading, problem solving in collaboration, and financial competence and in total constitutes 13 h of tests. However, the test is constructed using a matrix design such that each student only answers a specific and limited combination of questions, resulting in a test that lasts approximately 2 h. Since the PISA 2015 edition

focused on science, this field of study was evaluated in greater detail, and hence, the number of items evaluating this area was higher than for other areas, a total of 184 items, which equates to about 6 h in terms of test time, although each student only answers a (different) subset of these questions (for more details on the design, see the PISA 2015 Theoretical Framework: OECD, 2016a).

The student questionnaire collected demographic data of the students and their perceptions of their school environment, their learning experience, the processes and practices employed by the school, and students' behavior. Based on students' self-reports, a number of instruments were constructed: simple indexes (i.e., gender, age, or repetition of the same school grade) and complex indexes (economic, social, and cultural status (ESCS), an index of the disciplinary climate in the classroom, index of instrumental motivation, etc.).

TABLE 1 | Sample configuration.

Abbreviation	Country	Total	% of girls
AUS	Australia	14,530	49
AUT	Austria	7,007	49
BEL	Belgium	9,651	49
CAN	Canada	20,058	50
CHL	Chile	7,053	50
CZE	Czech Republic	6,894	50
DNK	Denmark	7,161	50
EST	Estonia	5,587	50
FIN	Finland	5,882	49
FRA	France	6,108	51
DEU	Germany	6,504	49
GRC	Greece	5,532	49
HUN	Hungary	5,658	50
ISL	Iceland	3,371	52
IRL	Ireland	5,741	49
ISR	Israel	6,598	56
ITA	Italy	11,583	50
JPN	Japan	6,647	50
KOR	Korea	5,581	48
LVA	Latvia	4,869	50
LUX	Luxembourg	5,299	51
MEX	Mexico	7,568	50
NLD	Netherlands	5,385	50
NZL	New Zealand	4,520	50
NOR	Norway	5,456	50
POL	Poland	4,478	49
PRT	Portugal	7,325	50
SVK	Slovak Republic	6,350	48
SVN	Slovenia	6,406	45
ESP	Spain	6,736	51
SWE	Sweden	5,458	50
CHE	Switzerland	5,860	48
TUR	Turkey	5,895	50
GBR	United Kingdom	14,157	49
USA	United States	5,712	50
OECD total		248,620	50

In terms of the new element added to the PISA study in 2015, that is, the assessment of both subjective and objective measures of student well-being, as mentioned earlier, five dimensions were examined in order to consider well-being as a multidimensional element.

The cognitive dimension comprises students' self-beliefs about their acquisition of subject-specific skills. As science was the main domain in PISA 2015, the questions regarding self-beliefs related to this area of knowledge. The constructs measured were: *science self-efficacy*, *broad interest in science*, *interest in broad science topics*, and *instrumental motivation to learn science*.

The psychological dimension encompassed psychological functioning in relation to educational aspects such as *students' career and educational expectations*, measured in terms of the *expected job* and *the highest level of education* each student aspired to, *achievement motivation*, and *test and learning anxiety*, along with the *overall satisfaction with life*.

The physical dimension in PISA 2015 measured two aspects of students' lifestyle: *the amount and frequency of physical activity* and *eating habits*. Specifically, students were asked if they exercised or did any sport before or after going to school, how many days per week they had physical education classes in school, and how often they were engaged in moderate or vigorous physical activities outside school. Students also reported whether they ate breakfast before going to school and dinner in the evening after school.

The assessment of the social well-being dimension was particularly important, as the quality of 15-years old relationships with teachers and peers is strongly linked to subjective well-being perception. PISA 2015 assessed five aspects of social well-being: *students' sense of belonging at school*; *social learning experiences*, assessed through the value given to and enjoyment of cooperative learning; *the relationship with their teachers*, assessed through the perception of teachers' unfair treatment of students; *the relationship with their peers*, as measured by the constructs *engagement with peers* and *bullying*; and *the relationship with their parents*, assessed through the scales *parental support* and *engagement with parents*.

Lastly, the material dimension investigated both the material resources available in the students' households and the infrastructure of their school. The material conditions at home focused on *parental occupation status* and *physical resources at home*, data that also contributed to the computation of ESCS. Moreover, the students were asked if they *worked for pay* or *worked in households* before or after school. Information about the quality of the material environment of the school was collected through the questionnaire directed at school principals, which sought to quantify *human resources* in terms of the professional profile of the teachers employed by the school and any staff shortages, *material resources*, measured as the availability of *physical educational resources* and *computer availability*, and lastly, the *extracurricular activities* offered by the school. **Table 2** describes the well-being model based on the OECD well-being framework.

The original version of the student questionnaire can be found in Annex A of the PISA 2015 Theoretical Framework (OECD, 2016a), while the items of the specific well-being

TABLE 2 | Well-being model dimensions.

Dimension	Constructs
Cognitive dimension	Enjoyment of science
	Instrumental motivation in science
	Science self-efficacy
	Interest in broad science topics
Material dimension	Parental occupation
	Physical resources at home
	Shortage of educational material
	Shortage of educational staff
	Index proportion of all teachers fully certified
	Total number of all teachers at school
Physical dimension	Eating breakfast/dinner
	Exercise or practice sport outside of school
Psychological dimension	Overall life satisfaction
	Achievement motivation
	Students' career and educational expectations
	Test and learning anxiety
Social dimension	Belongingness at school
	Relationship with teachers: teacher fairness
	Collaboration and teamwork dispositions: enjoy cooperation
	Collaboration and teamwork dispositions: value cooperation
	Bullying

scales and constructs are collated in "A Framework for the Analysis of Student Well-Being in the PISA 2015 Study" (Borgonovi and Pál, 2016).

Procedure

The students participating in PISA 2015 took a computer-based test, with assessments lasting a total of 2 h for each student. They also answered a background questionnaire, which took around 35 min to complete. The data collected were processed and published by OECD.

To achieve the objectives of our study, we used OECD data to perform a two-step analysis. Firstly, the well-being model configured through the dimensions or components that significantly impact students' performance in an international context was identified. As a preliminary step, each dimension of the proposed model was analyzed individually, discarding variables until the model adequately fitted the data. Then, the well-being-performance model was constructed by introducing science performance (the major domain of the 2015 edition of PISA) as the dependent variable. Science performance was estimated as the mean of the 10 plausible values, the estimators of student proficiency used in PISA. The proposed well-being model was configured for the whole sample of the OECD students.

Secondly, the magnitude of school effects in terms of the various measures of well-being were estimated at the international and country level. With this purpose, the gross variance of the well-being indicators accounted for by clustering as well as the variance adjusted by students' characteristics were assessed.

In addition, the relationships between student/school-level factors and the well-being indicators at the international and country level were analyzed. With this purpose, the previous model was enriched with the predictor variables related to school characteristics.

Data Analyses

During the first step, the well-being model was evaluated using confirmatory factor analysis (CFA), where the latent variables were those represented by student responses to the student questionnaire. The estimation method employed was maximum likelihood with robust standard errors. The fit of the model was analyzed according to different criteria: the comparative fit index (CFI), the Tucker–Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR), taking into account the usual criteria as set out in Hu and Bentler (1999): CFI and TLI should be greater than 0.95, RMSEA should be below 0.06, and SRMR below 0.08. Then, the multiple regressions for the OECD countries as a whole and for individual countries were used to compute the standardized beta weights and the percentage of variance in academic achievement as a function of the studied variables. CFA was carried out using the *lavaan* package of R software (Rosseel, 2012), and multiple regressions using the *rms* package (Harrell, 2019).

The second step aimed to measure, at the OECD level and the individual country level, school effectiveness in the promotion of the well-being dimension, as well as those variables identified in step 1 as being important in relation to performance. At the country level, the PISA data have a hierarchical structure, where the individuals at level 1 (students) are nested in clusters at level 2 (schools). It is generally accepted that school effectiveness studies require multilevel techniques, such as those developed by Aitkin and Longford (1986), to be employed both in order to estimate the magnitude of school effects and to analyze the impact of student- and school-related factors (Aitkin and Longford, 1986; Hill and Rowe, 1996; Kennedy and Mandeville, 2000; Goldstein, 2003; Murillo, 2008; Gamazo et al., 2018). In this work, therefore, hierarchical linear modeling was used to estimate school effects on well-being indicators (Snijders and Bosker, 2012) whereby the two-level technique was applied in the cross-country analysis, the first level corresponding to students and the second to schools.

The estimation of the variance components of the model allows the calculation of the intraclass correlation coefficient (ICC), which represents the proportion of variation in dependent variables that is accounted for by clustering (Snijders and Bosker, 2012), i.e., ICC, is the ratio of the between-school variance to the sum of the between-school and within-school variance. ICC was calculated in two phases.

Phase 1. Null Model Estimation

In the first phase, gross school effects were estimated through the null model, which contained only the dependent variables and the constant. In this configuration, the model has random effects at both levels without taking into account any control variables. The null model is usually established as the starting point of multilevel

analysis. It makes it possible to obtain the "gross" school effects, assessed through the ICC, i.e., those effects that are not adjusted for contextual variables (Lee, 2000; Hayes, 2006).

Phase 2. Estimation of the Model Incorporating Adjustment Variables

There is a consensus that school effects cannot be measured in terms of "gross" results but should be adjusted by relevant factors related to student progress (Goldstein et al., 1993; Mortimore et al., 1994; Goldstein and Thomas, 1996; Gray et al., 1996). To this end, in the second phase, the model was enriched with the control variables (**Table 3**), and the adjusted school effects measured in terms of ICC were estimated.

The ESCS index at the student level and the mean ESCS at the school level were incorporated in the model. These indicators have continuously been demonstrated to be strong predictors of school outcomes in all OECD countries (Perry and McConney, 2010a,b; Cordero et al., 2014; Suárez-Álvarez et al., 2014; OECD, 2016b; Gamazo et al., 2018). In PISA, the ESCS index is constructed from three components: *the occupational status of the parents*, *the educational level of the parents* (selecting in both cases the data for the parent with the higher level), and *home possessions*.

The impact of student-level background information, like gender and immigration status, has also been widely studied, the results underlining the importance of gender as a predictor of achievement (Stoet and Geary, 2013; Karakolidis et al., 2016; Özdemir, 2016). The model used in this work also included information about repetition of the same grade. Although its benefits are not compared between OECD countries here (Jacob and Lefgren, 2004, 2009; Manacorda, 2012), this strategy is widely used in some countries, like Spain and Portugal. For categorical variables like gender and immigrant status, dummy variables were generated (as many as the number of categories of the original variable minus one).

At the last stage of the study, with the purpose of assessing the relationship between student and school factors related to well-being, the complete model was configured whereby the predictor variables from the previous step were widened to include school factors such as school type, class size, or teaching methodology. There is evidence that supports the notion that these factors influence educational outcomes. For instance, the meta-analysis by Hattie (2009) suggests that reduced class size is a determining factor for improving student achievement, along with a reduced teacher–student ratio (Nath, 2012).

The model was also enriched with the variables that evaluated teaching strategies and teacher support, concepts that have recently gained interest in the academic field (Hattie, 2009; Nath, 2012; Gil et al., 2018) with respect to measuring their effects on student well-being. The OECD classifications distinguish between teacher-directed and student-centered instruction methodologies. Teacher-directed instruction, assessed through the scale *teacher-directed science instruction*, is focused on the role of teacher leading and managing the activities taking place in the classroom. Student-centered instruction, referred to as *inquiry-based science teaching and learning practices*, is associated with the teacher facilitating students' own learning by allowing them time to find solutions to problems on their own before the teacher confirms or demonstrates the solution (Hoad et al., 2007; Rowe, 2007; OECD, 2009).

Teacher support is also gaining importance (OECD, 2016b; Ricard and Pelletier, 2016). Following the PISA measurement construct, teacher support consists in the teacher showing an interest in every student's learning separately, giving extra help when needed, helping students with their learning, continuing to work on a teaching point until all students understand the material, and giving students an opportunity to express their opinions.

The school-level variables, i.e., teaching methodology and teacher support, were calculated as the across-school average of these student-level indexes, constructed on the basis of student responses to the context questionnaires following PISA methodology (OECD, 2017b). The predictor variables of the complete model are shown in the **Table 4**. The package lme4 of R software was used for multilevel modeling (Bates et al., 2015).

TABLE 3 | Control variables.

Level	Variable
Student level	Economic, social, and cultural status
	Gender
	Immigration status
	Grade
School level	School-level economic, social, and cultural status

RESULTS

Relationship Between Student Performance and Well-Being

As a preliminary step to data analysis, for each dimension of the proposed well-being model, a CFA was performed on the summed data for all OECD countries. The cognitive dimension, represented by the four constructs explained above, was the only one that achieved appropriate model fit according to the criteria

TABLE 4 | Predictor variables.

Level	Variable
Student level	Economic, social, and cultural status
	Gender
	Immigration status
	Grade
School level	School type
	Student–teacher ratio
	School size
	Class size
	Teacher-directed science instruction (school level)
	Inquiry-based science teaching and learning practices (school level)
	Teacher support of students' choices in a science classes (school level)

of Hu and Bentler (1999). The material dimension, defined as the economic resources of a student's household, represented through the *home possessions* index and the index of parental occupational, was also confirmed. Although psychological, social, and physical dimensions, assessed through the respective OECD scales, did not exhibit construct solidity, individual scales aimed at assessing these dimensions were introduced into the well-being-performance model in order to capture whether, and how, subjective non-cognitive well-being indicators predict performance in science.

Figure 1 presents the final well-being-performance model, which is the one that achieved the highest fit values. **Table 5** indicates the values obtained for the CFA. In the model representing the relationship between well-being and performance, the cognitive dimension was finally reduced to three scales: *enjoyment of science*, *instrumental motivation in science*, and *science self-efficacy*. In the psychological dimension, the variables *achievement motivation* and *test anxiety,* considered individually, acted as good predictors of science performance. The material dimension, measured through the level of *parents' occupation* and *home possessions* of students' families, was strongly related to performance. Finally, although four social dimension variables (*belongingness at school, teacher fairness, enjoy cooperation*, and *value cooperation)* were demonstrated to have a significant impact on science achievement, the model that included all four of them did not fit the data well. However, the variable *enjoy cooperation* contributed positively to the final model. The variables of the physical dimension did not provide reliable information about the well-being-performance model, probably due to their dichotomous nature.

Multiple regression (**Table 6**) indicated that well-being variables explained around 22% of the variance related to students' achievement in science.

It can be observed that in the regression model performed for the overall OECD sample, the greatest weight corresponded to the material well-being dimension. Nevertheless, the impact of cognitive well-being is also both high and constant across countries: on average, an increase of 1 point in terms of cognitive well-being would result in an increment of 18 points on the PISA science achievement scale. In six countries (Australia, Canada, Finland, Iceland, Ireland, and Korea), the cognitive variables are able to predict achievement as much as, or in some cases better than, economic background does. Test anxiety was found to reduce science performance by up to 12 points, with the strongest negative relationship observed in Finland. These results are also constant across countries, excepting Korea, where higher test anxiety corresponds to higher performance in science. *Achievement motivation* and *enjoy cooperation* both also relate positively to the cognitive results in most of the countries, although their impact is weaker.

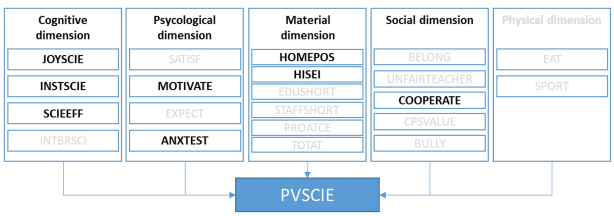


FIGURE 1 | Confirmed well-being/performance model. Source, Prepared by the authors, based on Programme for International Student Assessment (PISA) Well-being Framework.

TABLE 5 | Model fit test statistics of well-being-performance model.

	χ^2	CFI	TLI	RMSEA	SRMR
Total OECD	144701. 688***	0.957	0.953	0.038	0.036

****Significant at $p < 0.01$; CFI, comparative fit index (CFI); TLI, Tucker–Lewis index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.*

School Effectiveness in Well-Being Promotion

Once the well-being components that were strongly related to performance were identified, we studied the school effects on the well-being components that can be modulated by the school. These effects on science cognitive scores are also presented.

Table 7 summarizes the school effects for the null model and for the model adjusted for student background and ESCS information. Consistent with previous research, the results indicate that the school seems to have only a weak influence on student well-being, although there is some variation depending on the country analyzed and on the predictor variable considered. In the null model, the total OECD school-level variation in science performance was around 39%, while it barely reached 9% for the well-being components, indicating that the school's role turns out to be much less important in promoting students' well-being. School effects accounted for 9% of variation in the cognitive well-being dimension, 8% of test anxiety, and 5% of enjoyment of cooperation. The model adjusted to incorporate the control variables did not result in any significant differences in terms of school effects, explaining only 1% of variation for enjoyment of science and enjoyment of cooperation variables, while the school-level variation in science performance was reduced up to 25%.

Figure 2 reflects cross-country school effects for the cognitive dimension and for the psychological and social variables (for the country-level results and for the variables that compose the cognitive well-being dimension, please refer to **Supplementary Material**). In comparison with the rest of the variables, the role of the school in cognitive well-being promotion is systematically higher than other dimensions in OECD countries. Adjusted school effects in Italy and Japan were around 10% in terms of cognitive well-being. Italy, along with Belgium, also showed higher variability in students' perception of achievement motivation at the school level. Enjoyment of cooperation is the component of social well-being where schools had less impact, a result that is consistent across countries, with Switzerland being the only country where it exceeded 5%. Schools also do not seem to play any great part in test anxiety reduction. Italy was the only country where any great amount (10%) of

TABLE 6 | Regression coefficients.

Abbreviations	Country	R^2	Beta				
			COGWB	MATWB	MOTIVATE	ANXTEST	COOPERATE
Total OECD		0.224	17.88***	31.44***	0.41***	−12.24***	4.35***
AUS	Australia	0.238	28.73***	25.69***	5.25***	−8.85***	2.00***
AUT	Austria	0.247	16.41***	34.56***	2.28***	−16.02***	1.74***
BEL	Belgium	0.256	20.07***	38.00***	−5.73***	−10.12***	7.98***
CAN	Canada	0.197	21.77***	21.22***	5.15***	−11.67***	1.39***
CHL	Chile	0.241	4.93***	29.52***	5.95***	−18.01***	8.50***
CZE	Czech Republic	0.317	14.69***	42.68***	6.47***	−15.94***	8.46***
DNK	Denmark	0.280	21.76***	26.05***	12.41***	−14.40***	1.83***
EST	Estonia	0.273	16.64***	22.74***	9.65***	−17.89***	5.65***
FIN	Finland	0.262	24.39***	23.68***	11.72***	−20.38***	NS
FRA	France	0.279	22.00***	36.99***	NS	−10.44***	9.44***
DEU	Germany	0.184	18.03***	39.22***	1.14***	−12.93***	6.59***
GRC	Greece	0.247	18.66***	27.13***	9.35***	−11.37***	5.10***
HUN	Hungary	0.295	3.26***	43.61***	7.54***	−12.87***	10.52***
ISL	Iceland	0.229	17.16***	14.48***	12.56***	−16.23***	2.30*
IRL	Ireland	0.220	27.45***	23.42***	7.55***	−13.31***	−1.19***
ISR	Israel	0.258	13.11***	29.54***	2.64***	−6.88***	−0.50*
ITA	Italy	0.277	15.98***	30.97***	−1.58***	−9.20***	9.08***
JPN	Japan	0.214	23.89***	24.37***	5.71***	−1.80***	−0.88***
KOR	Korea	0.130	25.08***	25.06***	11.60***	1.12***	−1.35***
LVA	Latvia	0.180	10.80***	23.41***	11.15***	−18.23***	12.67***
LUX	Luxembourg	0.168	17.35***	39.38***	NS	−16.33***	6.02***
MEX	Mexico	0.209	4.68***	17.17***	8.64***	−12.43***	5.16***
NLD	Netherlands	0.341	23.29***	36.80***	6.69***	−1.71***	4.89***
NZL	New Zealand	0.192	27.67***	29.33***	3.61***	−15.71***	1.71***
NOR	Norway	0.173	24.57***	24.94***	9.03***	−11.76***	3.42***
POL	Poland	0.198	12.96***	28.40***	9.91***	−15.64***	7.03***
PRT	Portugal	0.228	18.02***	30.10***	13.50***	−15.43***	−1.87***
SVK	Slovak Republic	0.247	11.03***	34.18***	11.21***	−9.54***	14.63***
SVN	Slovenia	0.197	17.10***	33.63***	6.97***	−12.89***	10.96***
ESP	Spain	0.287	20.62***	24.50***	10.44***	−16.14***	5.24***
SWE	Sweden	0.224	22.96***	29.19***	6.33***	−9.23***	2.95***
CHE	Switzerland	0.201	18.54***	39.44***	2.06***	−13.14***	4.49***
TUR	Turkey	0.226	10.87***	26.58***	4.36***	−5.14***	6.19***
GBR	United Kingdom	0.188	27.71***	28.21***	−0.37***	−6.30***	1.99***
USA	United States	0.182	18.35***	26.93***	0.26***	−10.71***	2.21***

****Significant at $p < 0.01$; *significant at $p < 0.1$; NS, not significant. COGWB, cognitive well-being; MATWB, material well-being; MOTIVATE, achievement motivation; ANXTEST, test and learning anxiety; COOPERATE, enjoy cooperation.*

variation in this dependent variable was accounted for by school nesting. In Iceland and Germany, no school variation in test anxiety was observed.

Student and School Factors Related to Performance

Finally, the impact of student and school factors on well-being variables was assessed. The results for science performance are presented in order to reflect the differences in the influence of these factors on achievement results and well-being. **Table 8** shows the estimates of multilevel modeling on each dependent variable for the overall OECD sample. **Table 9** identifies the number of countries where the factors are significantly positively related to the dependent variables.

ESCS has traditionally been positively related to performance, a tendency that, in this work, persists when well-being variables are taken into account. Students with higher ESCS exhibited higher cognitive well-being, with the strongest impact being on their perception of self-efficiency in science. More advantaged students also had higher achievement motivation and were more resistant to stress as a result of exams. This relationship was reproduced at the individual country level.

At the OECD level, girls demonstrated lower levels of cognitive well-being along with higher levels of test anxiety, although they enjoyed cooperation more than their male classmates and had higher levels of achievement motivation. At the individual country level, these results were repeated, except for achievement motivation, where, in 9 countries, no clear relationship with gender was observed, while in 14 countries, boys were more highly motivated to achieve academically.

Once ESCS was controlled for, students with immigrant backgrounds reported higher motivation to achieve than non-immigrant students. They also demonstrated higher levels of cognitive well-being, especially for enjoyment of science and

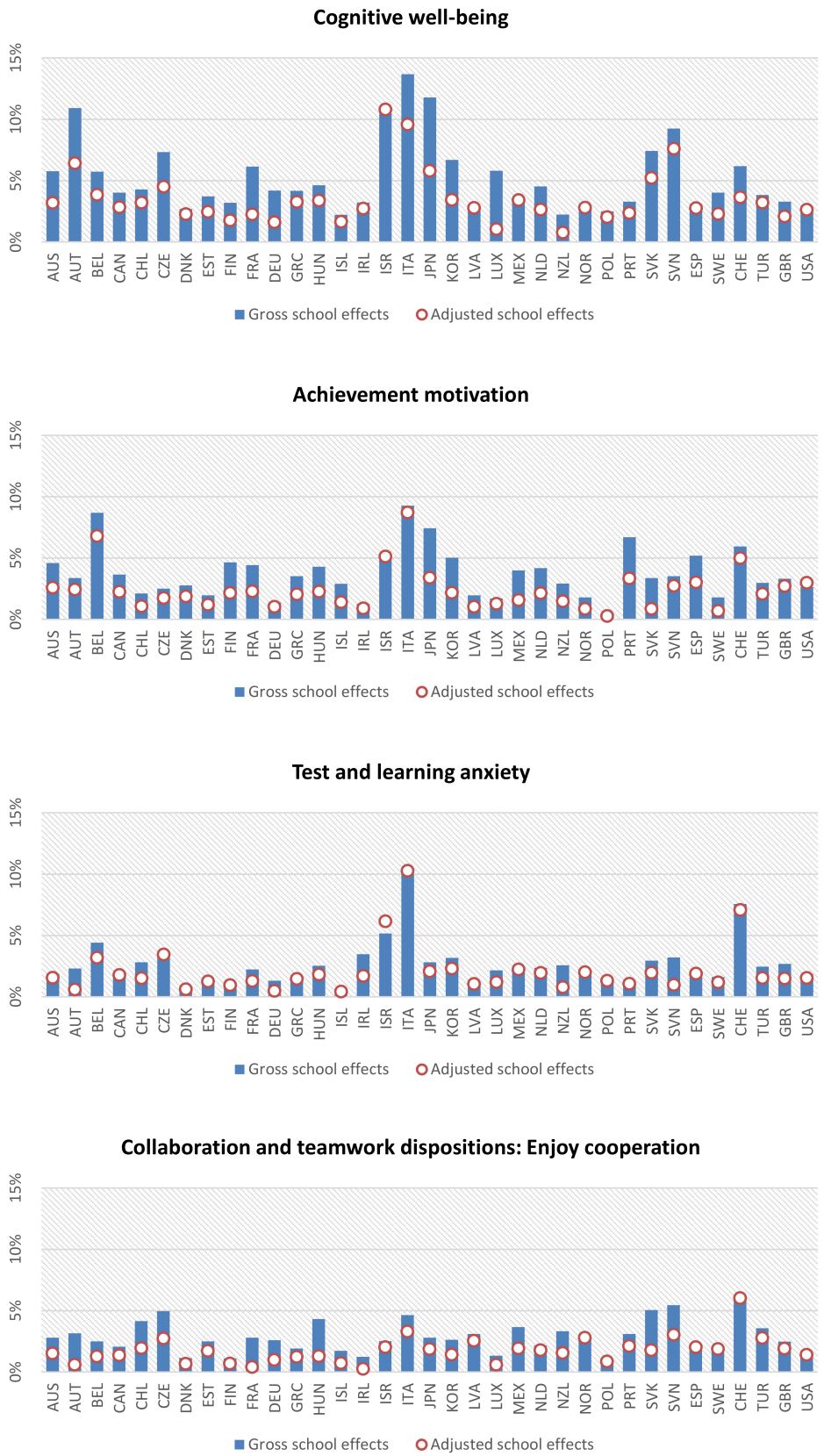


FIGURE 2 | Country-level school effects in terms of intraclass correlation coefficient (ICC).

TABLE 7 | School effects in terms of ICC.

	SCIE	COGWB	JOYSCIE	SCIEEFF	INSTSCIE	MOTIVATE	ANXTEST	COOPERATE
Gross school effect								
Total OECD school effect	39.0%	9.0%	9.0%	5.0%	7.0%	3.0%	8.0%	5.0%
Min school effect	5.0%	2.2%	2.1%	1.7%	0.5%	0.5%	0.6%	1.0%
Max school effect	62.1%	13.7%	10.7%	7.2%	12.6%	9.3%	10.0%	6.3%
Net school effect								
Total OECD school effect	25.0%	9.0%	8.0%	5.0%	7.0%	3.0%	8.0%	4.0%
Min school effect	3.6%	0.8%	0.9%	0.4%	0.3%	0.3%	0.4%	0.3%
Max school effect	41.5%	10.8%	8.6%	3.9%	11.8%	8.7%	10.3%	6.0%

ICC, intraclass correlation coefficient; SCIE, science performance; COGWB, cognitive well-being; JOYSCIE, enjoyment of science; SCIEEFF, science self-efficacy; INSTSCIE, instrumental motivation in science; MOTIVATE, achievement motivation; ANXTEST, test and learning anxiety; COOPERATE, enjoy cooperation.

instrumental motivation. On the other hand, being an immigrant was associated with higher test anxiety.

At the school level, the influence of school characteristics, along with teaching methods and teacher support, on the students' subjective well-being was assessed. Although public schools consistently performed worse than private schools even after controlling for ESCS, this tendency was reversed in terms of students' perception of their cognitive well-being. In public schools, students tended to demonstrate higher levels of self-efficiency and science enjoyment. Nevertheless, they were less motivated to achieve and more prone to feeling anxious about exams. The school and class size seemed to have a very low impact on students' perception of well-being both at the OECD and at the individual country level.

Teaching methodology, measured as the use of teacher-directed or inquiry-based instruction, and teacher support are strongly and positively related to the well-being indicators, while they have an opposite effect in relation to science performance: the more frequent use of inquiry-based teaching and higher teacher support are associated with a decrease in science performance of around 25 points on the PISA scale. However, more inquiry-based instruction, when students are given opportunities to explain their ideas, spend time in the laboratory doing practical experiments, or are required to discuss science questions, increases students' perception of self-efficiency and promotes intrinsic motivation by increasing science enjoyment. Furthermore, it reduces exam anxiety and raises achievement motivation. The positive relationship between inquiry-based teaching and the cognitive well-being dimension is confirmed individually in 17 OECD countries.

Enjoyment for science is higher when the teacher regularly explains scientific ideas, a whole class discussion takes place with the teacher, and the teacher addresses students' questions and practically explains an idea. Teacher-directed instruction also increases students' positive predisposition toward cooperation. The positive impact of teacher-directed methodologies on cognitive well-being and cooperation is observed in 19 separate OECD countries.

Teacher support was the school-level variable that demonstrated the strongest relationship with student well-being in the model proposed. Showing an interest in every student's learning, giving extra help when students need it, and continuing with explanations until all students understand the material turn out to be extremely important for the promotion of achievement motivation and for positive predisposition toward cooperation. In addition, these practices reduce test anxiety in 7 of the 10 OECD countries where teacher support is significant.

DISCUSSION

The aim of the study was twofold. On the one hand, the present study sought to reach a global definition of well-being across the countries assessed for the PISA 2015 report and to assess its relationship with performance. On the other, we focused on ascertaining the impact of school effects on student welfare and identifying those factors positively related to well-being in the educational context.

The results evidenced the complexity of the well-being concept and the need for further research on its definition. Of the four dimensions described in the original model, only the cognitive dimension was confirmed as having an impact across all countries in PISA 2015. In the evaluation of the material dimension, only student-level variables contributed positively to the model, while school environment and resources did not demonstrate any significant effect once the students' economic background was taken into account. Psychological and social dimensions were found to be multifaceted concepts represented by a variety of individual variables but not confirmed as solid constructs. Finally, the physical dimension did not provide reliable information with respect to the construct definition.

Consequently, in the well-being-performance model, well-being was finally defined by the cognitive and material dimensions, along with the individual psychological and social variables *achievement motivation*, *test anxiety*, and *enjoyment of cooperation*, i.e., the variables that were found to be good predictors of performance in science. The results showed that student well-being significantly impacts student performance. Higher cognitive well-being is associated with better achievement results, increasing science performance by up to 22 points on the PISA scale. In six countries, the promotion of cognitive well-being was even demonstrated to counteract the effect of socioeconomic background. Lower *test anxiety* is also linked to better results, along with *enjoyment of cooperation*.

Nevertheless, currently, school interventions do not appear strong enough to make an impact on subjective well-being.

TABLE 8 | Estimation of fixed effects and random effects of the complete model for the overall OECD sample.

	PVSCIE	COGWB	JOYSCIE	SCIEEFF	INSTSCIE	MOTIVATE	ANXTEST	COOPERATE
Student level								
ESCS	19.65***	0.17***	0.15***	0.22***	0.08***	0.14***	−0.06***	0.09***
GENDER_girl	−8.08***	−0.13***	−0.13***	−0.20***	−0.02***	0.02***	0.45***	0.21***
IMMIG_yes	−18.29***	0.10***	0.13***	0.03**	0.09***	0.19***	0.08***	0.06***
School level								
SCHLTYPE_pub	−11.63***	0.04***	0.03***	0.06***	0.01	−0.08***	0.03***	−0.05***
STRATIO	0.00	0.00***	0.00**	0.00**	0.00***	0.00***	0.00	0.00***
SCHSIZE	0.00	0.00	0.00***	0.00***	0.00	0.00***	0.00	0.00***
CLSIZE	0.46***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***	0.01***
TDTEACH_S	52.59***	0.18***	0.29***	0.06***	0.10***	0.05**	0.04***	0.17***
IBTEACH_S	−27.20***	0.17***	0.13***	0.24***	0.07***	0.09***	−0.07***	0.04***
TEACHSUP_S	−24.97***	0.23***	0.21***	0.09***	0.28***	0.33***	0.21***	0.05***
Random effects								
σ2	5248.41(71%)	0.88(95%)	1.15(96%)	1.44(97%)	0.94(96%)	0.81(87%)	0.87(94%)	0.94(97%)
τ00 (CNTSCHID)	2116.7(29%)	0.05(5%)	0.05(4%)	0.05(3%)	0.04(4%)	0.12(13%)	0.06(6%)	0.03(3%)

****Significant at $p < 0.01$; **significant at $p < 0.05$; NS, not significant. CNTSCHID, country school ID; PVSCIE, science performance; COGWB, cognitive well-being; JOYSCIE, enjoyment of science; SCIEEFF, science self-efficacy; INSTSCIE, instrumental motivation in science; MOTIVATE, achievement motivation; ANXTEST, test and learning anxiety; COOPERATE, enjoy cooperation; ESCS, economic, social, and cultural status; GENDER_girl, gender (the student is a girl); IMMIG_yes, (the student is an immigrant); SCHLTYPE_pub, school type (the school is public); SCHSIZE, school size; STRATIO, student–teacher ratio; CLSIZE, class size; TDTEACH_S, teacher-directed science instruction (school level); IBTEACH_S, inquiry-based science teaching and learning practices (school level); TEACHSUP_S, teacher support of students' choices in science classes (school level).*

School effects explain barely 5% of the variation in well-being perception within schools, and school-level variation is highest for the cognitive well-being dimension, accounting for up to 9% of school effects on average across all the OECD countries. These results are consistent with previous studies (Murillo and Hernández-Castilla, 2011; Lazarides and Buchholz, 2019) and provide further evidence in support of these effects both in the across-OECD context as well as for each member country. Our results highlight that some countries, like Italy and Switzerland, are more successful with school-level interventions, while others, like Poland and Iceland, have a very limited school-level influence on well-being. There may be multiple reasons for this low school-level impact on well-being, the most likely being a lack of socio-emotional education within schools, the low availability of tools and policies for well-being improvement, or the limited time dedicated to achievement in non-academic aspects of learning (Murillo and Hernández-Castilla, 2011), although it is becoming more common to introduce school practices aimed at the promotion of cognitive, social, and emotional well-being and stress reduction (Jennings et al., 2013; Schonert-Reichl et al., 2015). Research findings provide evidence to support the notion that the implementation of such methods improves attention deficits, reduces stress, and promotes self-regulation among adolescents (Albrecht et al., 2012; Carboni et al., 2013).

The student and school factors associated with higher levels of cognitive well-being, motivation, and cooperation were also assessed. At the student level, the socioeconomic background was again a good predictor of student well-being, which clearly makes it difficult for schools to combat its substantial influence. However, on the positive side, the analysis provides evidence that teachers employing a methodology that combines the traditional teacher-led approach with more innovative practices based on inquiry and teamwork seems to be a powerful tool for improving non-cognitive educational achievement. Science teaching and learning practices that include experimentation and critical thinking increase students' self-efficacy in science and reduce test anxiety. These insights are especially important given that student-oriented teaching methods seem to be negatively linked to academic achievement (Gil et al., 2018). A classic teacher-directed approach, where the teacher leads class discussions and explains ideas, is associated with higher levels of science enjoyment and better predisposition toward cooperative working. These results support the idea of the importance of an adaptive pedagogy that brings together innovation and teacher-directed instruction, rather than teachers opting exclusively for either one of these approaches (Rowe, 2007; OECD, 2008b).

Teacher support of pupils at the school level was initially negatively related to science performance in the multilevel model proposed here. This was probably due to the fact that teachers at schools in disadvantaged areas report supporting students in their learning more frequently than teachers in schools in more advantaged areas, as is also the case for teachers in rural as opposed to urban schools (OECD, 2017a). Schools in disadvantaged and rural areas tend to perform worse in the PISA assessments, and therefore, their students are in greater need of teacher support. Nevertheless, in this study, teacher support turned out to be the strongest predictor of student well-being, i.e., when the teacher works to ensure the students' complete understanding of the problem, provides extra help when it is required, and aims to integrate learning, students report higher subjective well-being. Previous research (Ahmed et al., 2014) has also shown that student-perceived teacher support is negatively related to student anxiety and boredom at the student level, and positively related to

TABLE 9 | Cross-country summary of fixed effects (number of countries with significant and positive impact).

Variable	SIG	ESCS	GENDER_girl	IMMIG_yes	SCHLTYPE_pub	SCHSIZE	STRATIO	CLSIZE	TDTEACH_S	IBTEACH_S	TEACHSUP_S
COGWB	SIG (N°)	35	34	18	5	0	7	4	19	17	17
	POSITIVE (N°)	35	3	13	3	0	3	4	19	17	15
JOYSCIE	SIG (N°)	33	25	16	6	6	9	7	29	12	17
	POSITIVE (N°)	33	2	15	2	0	6	6	28	10	15
SCIEEFF	SIG (N°)	35	30	13	9	7	6	5	13	18	8
	POSITIVE (N°)	35	1	12	3	0	4	3	11	18	4
INSTSCIE	SIG (N°)	31	19	14	5	3	6	3	4	14	17
	POSITIVE (N°)	31	7	14	3	0	4	2	3	14	16
MOTIVATE	SIG (N°)	34	26	22	12	7	3	6	12	6	10
	POSITIVE (N°)	34	12	21	4	0	2	4	9	3	8
ANXTEST	SIG (N°)	28	35	15	6	5	4	7	4	7	10
	POSITIVE (N°)	0	35	13	4	0	1	5	3	2	3
COOPERATE	SIG (N°)	33	34	12	9	10	4	11	20	4	13
	POSITIVE (N°)	33	34	9	3	0	1	7	19	1	12

SCIE, science performance; COGWB, cognitive well-being; JOYSCIE, enjoyment of science; SCIEEFF, science self-efficacy; INSTSCIE, instrumental motivation in science; MOTIVATE, achievement motivation; ANXTEST, test and learning anxiety; COOPERATE, enjoy cooperation; ESCS, economic, social, and cultural status; GENDER_girl, gender (the student is a girl); IMMIG_yes, (the student is an immigrant); SCHLTYPE_pub, school type (the school is public); SCHSIZE, school size; STRATIO, student–teacher ratio; CLSIZE, class size; TDTEACH_S, teacher-directed science instruction (school level); IBTEACH_S, inquiry-based science teaching and learning practices (school level); TEACHSUP_S, teacher support of students' choices in science classes (school level).

enjoyment and negatively related to anxiety at the classroom level (Lazarides and Buchholz, 2019).

In the 21st century, the era of knowledge and innovation, the school has gained great importance in the development and learning of individuals, as well as it having become an extraordinarily complex and multidisciplinary facility. On the one hand, the purpose of schools is to promote knowledge acquisition, but on the other, they must help children build confidence and develop a variety of learning strategies for the future (OECD, 2008a). This research aims to contribute to the growing concern about students' quality of life and happiness and to emphasize the importance of a comprehensive approach to education where socio-emotional development is integrated in a schools' day-to-day functioning.

The principal limitation of the study lies in the need for improvement in the instruments available for assessing well-being in an educational context. Although the OECD provides a solid framework for the measurement of well-being, some dimensions, like physical well-being, still need to include reliable and unidimensional scales. Moreover, it should be taken into account that instruments based on self-reporting will never achieve the same level of sensitivity in measuring latent constructs such as those involved in well-being as do academic achievement tests (Murillo and Hernández-Castilla, 2011).

The results of this research should be considered with cautions, as there is no evidence of causality for the relationships observed. The reciprocal relationship between well-being and performance should be taken into account. For instance, previous research has shown that higher levels of achievement are positively related to enjoyment (Ma, 1997) and reduce exam anxiety (Ma and Xu, 2004). In addition, the impact of student and family characteristics should not be forgotten, as they are connected to the achievement and behavior of students at school, as García-Crespo et al. (2019) indicate. Nevertheless, the conclusions regarding teaching methodology are more consistent, although it would be interesting to study the persistence of the positive impact of teachers' interventions with respect to students with different academic profiles (low/average/high academic performance, etc.). Future research within our research team will focus on expanding on the results obtained in this work by extending the analysis to primary education data, where it is expected that school involvement in socio-emotional variables is more common and efficient.

DATA AVAILABILITY STATEMENT

This study was based on the public data available on the OECD web page: http://www.oecd.org/pisa/data/2015database/.

ETHICS STATEMENT

The study did not require ethics approval. The data collection was performed on behalf of the OECD following the standards

for the protection of privacy and the processing of personal data (http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm).

AUTHOR CONTRIBUTIONS

EG prepared the data set for analysis, conducted the analysis, and was involved in writing the manuscript. IB and JM supervised the analysis and participated in writing the manuscript.

FUNDING

This study was partially funded by the Spain Ministry of Science, Innovation and Universities (Project number PSI2017-85724-P) and by 2E Estudios, Evaluaciones e Investigación, S.L.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.00431/full#supplementary-material

REFERENCES

Ahmed, I., Khairuzzaman Wan Ismail, W., Mohamad Amin, S., and Islam, T. (2014). Role of perceived organizational support in teachers' responsiveness and students' outcomes: evidence from a public sector University of Pakistan. *Int. J. Educ. Manag.* 28, 246–256. doi: 10.1108/ijem-02-2013-0031

Aitkin, M., and Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *J. R. Stat. Soc. Ser. A* 149, 1–26.

Albrecht, N. J., Albrecht, P. M., and Cohen, M. (2012). Mindfully teaching in the classroom: a literature review. *Aust. J. Teach. Educ.* 37, 1–15.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Belfi, B., Goos, M., De Fraine, B., and Van Damme, J. (2012). The effect of class composition by gender and ability on secondary school students' school well-being and academic self-concept: a literature review. *Educ. Res. Rev.* 7, 62–74. doi: 10.1016/j.edurev.2011.09.002

Berger, C., Alcalay, L., Torretti, A., and Milicic, N. (2011). Socio-emotional well-being and academic achievement: evidence from a multilevel approach. *Psicol. Reflex. Crít.* 24, 344–351. doi: 10.1590/s0102-79722011000200016

Borgonovi, F., and Pál, J. (2016). *A Framework for the Analysis of Student Well-Being in the PISA 2015 Study: Being 15 In 2015,* OECD Education Working Papers, No. 140. Paris: OECD Publishing, doi: 10.1787/5jlpszwghvvb-en

Carboni, J. A., Roach, A. T., and Fredrick, L. D. (2013). Impact of mindfulness training on the behavior of elementary students with attention-deficit/hyperactive disorder. *Res. Hum. Dev.* 10, 234–251. doi: 10.1080/15427609.2013.818487

Casas, F. (2004). De afrontar problemas sociales a promover la calidad de vida. *Rev. Hist. Psicol.* 25, 305–322.

Casas, F. (2010). El bienestar personal: su investigación en la infancia y la adolescencia. *Encuentros Psicol. Soc.* 5, 85–101.

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., et al. (1966). *Equality of Educational Opportunity*. Washington, DC: United States Government Publishing Office.

Cordero, J. M., Manchon, C., and Simancas, R. (2014). Repetition and explanatory factors in Spain. *Rev. Educ.* 2014, 12–37.

Crespo-Cebada, E., Pedraja-Chaparro, F., and Santín, D. (2014). Does school ownership matter? An unbiased efficiency comparison for regions of Spain. *J. Productivity Anal.* 41, 153–172. doi: 10.1007/s11123-013-0338-y

El Ansari, W., and Stock, C. (2010). Is the health and wellbeing of university students associated with their academic performance? Cross sectional findings from the United Kingdom. *Int. J. Environ. Res. Public Health* 7, 509–527. doi: 10.3390/ijerph7020509

Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., and Rodríguez-Conde, M. J. (2018). Evaluación de factores relacionados con la eficacia escolar en PISA 2015. Un análisis multinivel. [Assessment of factors related to school effectiveness in PISA 2015. A multilevel analysis]. *Rev. Educ.* 379, 56–84.

García-Crespo, F. J., Fernández-Alonso, R., and Muñiz, J. (2019). Resilient and low performer students: personal and family determinants in European countries. *Psicothema* 31, 363–375. doi: 10.7334/psicothema2019.245

Gil, M., Cordero, J. M., and Cristobal, V. (2018). Las estrategias docentes y los resultados en PISA 2015. Teaching practices and results in PISA 2015. *Rev. Educ.* 379, 32–55.

Goldstein, H. (2003). *Multilevel Statistical Models*, 3rd Edn. London: Hodder Arnold.

Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttall, D., et al. (1993). A multilevel analysis of school examination results. *Oxf. Rev. Educ.* 19, 425–433. doi: 10.1080/0305498930190401

Goldstein, H., and Thomas, S. (1996). Using examination results as indicators of school and college performance. *J. R. Stat. Soc. Ser. A* 159, 149–163.

Gray, J., Goldstein, H., and Jesson, D. (1996). Changes and improvements in schools' effectiveness: trends over five years. *Res. Pap. Educ.* 11, 35–51. doi: 10.1080/0267152960110104

Grisay, A. (1996). *Evolution des Acquis Cognitifs et Socio-Affectifs des Élèves au Cours des Années de Collège*. Liège: Université de Liège.

Grosskopf, S., Hayes, K. J., and Taylor, L. L. (2014). Efficiency in education: research and implications. *Appl. Econ. Perspect. Policy* 36, 175–210. doi: 10.1093/aepp/ppu007

Gutman, L. M., and Vorhaus, J. (2012). *The Impact of Pupil Behaviour and Wellbeing on Educational Outcomes.* Research report No. DFE-RR253. London: Department for Education.

Hanushek, E. A., and Luque, J. A. (2003). Efficiency and equity in schools around the world. *Econ. Educ. Rev.* 22, 481–502. doi: 10.1016/s0272-7757(03)00038-4

Harrell, F. E. Jr. (2019). *rms: Regression Modeling Strategies.* R package version 5.1-3.1.

Hattie, J. (2009). *Visible Learning: A Synthesis of Meta-Analyses Relating to Achievement*. London: Routledge.

Hayes, A. F. (2006). A primer on multilevel modeling. *Hum. Commun. Res.* 32, 385–410. doi: 10.1111/j.1468-2958.2006.00281.x

Hill, P. W., and Rowe, K. J. (1996). Multilevel modelling in school effectiveness research. *Sch. Eff. Sch. Improv.* 7, 1–34. doi: 10.1080/0924345960070101

Hoad, K., Munro, J., Pearn, C., Rowe, K., and Rowe, K. (2007). *Working Out What Works (WOWW) Training and Resource Manual: A Teacher Professional Development Program Designed to Support Teachers to Improve Literacy and Numeracy Outcomes for Students with Learning Difficulties in Years 4, 5 and 6*. Canberra, ACT: Commonwealth Department of Education, Science and Training.

Hu, L. T., and Bentler, P. M. (1999). Cut off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6, 1–55. doi: 10.1080/10705519909540118

Jacob, B. A., and Lefgren, L. (2004). Remedial education and student achieve – ment: a regression–discontinuity analysis. *Rev. Econ. Stat.* 86, 226–244. doi: 10.1162/003465304323023778

Jacob, B. A., and Lefgren, L. (2009). The effect of grade retention on high school completion. *Am. Econ. J. Appl. Econ.* 1, 33–58. doi: 10.1037/edu0000243

Jennings, P. A., Snowberg, K. E., Coccia, M. A., and Greenberg, M. T. (2013). Improving classroom learning environments by cultivating awareness and resilience in education (CARE): results of two pilot studies. *J. Classroom Interact.* 46, 37–48.

Johnson, A. L., and Ruggiero, J. (2014). Nonparametric measurement of productivity and efficiency in education. *Ann. Oper. Res.* 221, 197–210. doi: 10.1007/s10479-011-0880-9

Kankaraš, M., and Suárez-Álvarez, J. (2019). *Assessment Framework of the OECD Study on Social and Emotional Skills*, OECD Education Working Papers, No. 207. Paris: OECD Publishing, doi: 10.1787/5007adef-en

Karakolidis, A., Pitsia, V., and Emvalotis, A. (2016). Examining students' achievement in mathematics: a multilevel analysis of the programme for international student assessment (PISA) 2012 data for Greece. *Int. J. Educ. Res.* 79, 106–115. doi: 10.1016/j.ijer.2016.05.013

Kennedy, E., and Mandeville, G. (2000). *Some Methodological Issues in School Effectiveness Research The International Handbook of School Effectiveness Research*. London: Routledge, 203–219.

Lazarides, R., and Buchholz, J. (2019). Student-perceived teaching quality: how is it related to different achievement emotions in mathematics classrooms? *Learn. Instr.* 61, 45–59. doi: 10.1016/j.learninstruc.2019.01.001

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: the case of school effects. *Educ. Psychol.* 35, 125–141. doi: 10.1207/s15326985ep3502_6

Ma, X. (1997). Reciprocal relationships between attitude toward mathematics and achievement in mathematics. *J. Educ. Res.* 90, 221–229. doi: 10.1080/00220671.1997.10544576

Ma, X., and Xu, J. (2004). The causal ordering of mathematics anxiety and mathematics achievement: a longitudinal panel analysis. *J. Adolesc.* 27, 165–179. doi: 10.1016/j.adolescence.2003.11.003

MacBeath, J., and Mortimore, P. (2001). *Improving School Effectiveness*. Buckingham: Open University.

Manacorda, M. (2012). The cost of grade retention. *Rev. Econ. Stat.* 94, 596–606. doi: 10.1162/rest_a_00165

Martínez-Arias, M. R. (2009). Usos, aplicaciones y problemas de los modelos de valor añadido en educación. *Rev. Educ.* 348, 217–250.

Mortimore, P., Sammons, P., and Thomas, S. (1994). School effectiveness and value added measures. *Assess. Educ.* 1, 315–332.

Moyano-Díaz, E., and Ramos-Alvarado, N. (2007). Bienestar subjetivo: midiendo satisfacción vital, felicidad y salud en población chilena de la Región Maule. *Universum (Talca)* 22, 177–193.

Murillo, F. J. (2005). *La Investigación Sobre Eficacia Escolar*. Barcelona: Octaedro.

Murillo, F. J. (2008). Los modelos multinivel como herramienta para la investigación educativa. *Magis. Rev. Int. Invest. Educ.* 1, 45–62.

Murillo, F. J., and Hernández-Castilla, R. (2011). Factores escolares asociados al desarrollo socio-afectivo en Iberoamerica. *RELIEVE Rev. Electron. Invest. Eval. Educ.* 17:2.

Nath, S. R. (2012). Factors influencing primary students' learning achievement in Bangladesh. *Res. Educ.* 88, 50–63. doi: 10.7227/rie.88.1.5

Novello, A. C., Degraw, C., and Kleinman, D. V. (1992). Healthy children ready to learn: an essential collaboration between health and education. *Public Health Rep.* 107, 3–15.

OECD (2008a). *Assessment for Learning: Formative Assessment, OECD/CERI International Conference Learning in the 21st Century: Research, Innovation and Policy, the Centre for Education Research and Innovation*. Paris: OECD Publishing.

OECD (2008b). *Innovating to Learn, Learning to Innovate*. Paris: OECD Publishing, doi: 10.1787/9789264047983-en

OECD (2008c). *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*. Paris: OECD Publishing.

OECD (2009). *Creating Effective Teaching and Learning Environments: First Results from Teaching and Learning International Survey*. Paris: OECD Publishing.

OECD (2015). *Skills for Social Progress: The Power of Social and Emotional Skills*. Paris: OECD Publishing, doi: 10.1787/9789264226159-en

OECD (2016a). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*. Paris: OECD Publishing, doi: 10.1787/9789264255425-en

OECD (2016b). *PISA 2015 Results: Excellence and Equity in Education*, Vol. I. Paris: OECD Publishing, doi: 10.1787/9789264266490-en

OECD (2017a). *PISA 2015 Results: Students' Well-Being*, Vol. III. Paris: OECD Publishing.

OECD (2017b). *PISA 2015 Technical Report*. Paris: OECD Publishing.

Opdenaker, M.-C., Van Damme, J., De Fraine, D. F., Van Landeghem, G., and Onghena, P. (2002). The effect of schools and classes on mathematics achievement. *Sch. Eff. Sch. Improv.* 13, 399–427. doi: 10.1076/sesi.13.4.399.10283

Opdenakker, M.-C., and Van Damme, J. (2000). Effects of schools, teaching staff and classes on achievement and well-being in secondary education: similarities and differences between school outcomes. *Sch. Eff. Sch. Improv.* 11, 165–196. doi: 10.1076/0924-3453(200006)11:2;1-q;ft165

Özdemir, C. (2016). Equity in the Turkish education system: a multilevel analysis of social background influences on the mathematics performance of 15-year-old students. *Eur. Educ. Res. J.* 15, 193–217. doi: 10.1177/1474904115627159

Perry, L., and McConney, A. (2010a). School socio-economic composition and student outcomes in Australia: implications for educational policy. *Aust. J. Educ.* 54, 72–85. doi: 10.1177/000494411005400106

Perry, L. B., and McConney, A. (2010b). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teach. Coll. Rec.* 112, 1137–1162.

Podinovski, V. V., Ismail, I., Bouzdine-Chameeva, T., and Zhang, W. (2014). Combining the assumptions of variable and constant returns to scale in the efficiency evaluation of secondary schools. *Eur. J. Oper. Res.* 239, 504–513. doi: 10.1016/j.ejor.2014.05.016

Pollard, E. L., and Lee, P. D. (2003). Child well-being: a systematic review of the literature. *Soc. Indic. Res.* 61, 59–78.

Ricard, N. C., and Pelletier, L. G. (2016). Dropping out of high school: the role of parent and teacher self-determination support, reciprocal friendships and academic motivation. *Contemp. Educ. Psychol.* 44, 32–40. doi: 10.1016/j.cedpsych.2015.12.003

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling and more. Version 0.5–12 (BETA). *J. Stat. Softw.* 48, 1–36.

Rowe, K. (2007). *The Imperative of Evidence-Based Instructional Leadership: Building Capacity within Professional Learning Communities via a Focus on Effective Teaching Practice*. Jolimont, VIC: Centre for Strategic Education.

Sammons, P. (1999). School effectiveness: coming of age in the 21st century. *Manag. Educ.* 13, 10–13. doi: 10.1111/add.12232

Scheerens, J., and Demeuse, M. (2005). The theoretical basis of the effective school improvement model (ESI). *Sch. Eff. Sch. Improv.* 16, 373–385. doi: 10.1080/09243450500234567

Schonert-Reichl, K. A., Oberle, E., Lawlor, M. S., Abbott, D., Thomson, K., Oberlander, T. F., et al. (2015). Enhancing cognitive and social–emotional development through a simple-to-administer mindfulness-based school program for elementary school children: a randomized controlled trial. *Dev. Psychol.* 51, 52–66. doi: 10.1037/a0038454

Snijders, T. A. B., and Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London: SAGE, 304–307.

Stoet, G., and Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: within-and across-nation assessment of 10 years of PISA data. *PLoS One* 8:e57988. doi: 10.1371/journal.pone.0057988

Suárez-Álvarez, J., Fernández-Alonso, R., and Muñiz, J. (2014). Self-concept, motivation, expectations, and socioeconomic level as predictors of academic performance in mathematics. *Learn. Individ. Differ.* 30, 118–123. doi: 10.1016/j.lindif.2013.10.019

Sun, R. C. F., and Shek, D. T. L. (2010). Life satisfaction, positive youth development, and problem behaviour among chinese adolescents in Hong Kong. *Soc. Indic. Res.* 95, 455–474. doi: 10.1007/s11205-009-9531-9

Teddlie, C., Reynolds, D., and Sammons, P. (2000). "The methodology and scientific properties of school effectiveness research," in *The International Handbook of School Effectiveness Research*, eds C. Teddlie and D. Reynolds (London: Falmer Press), 55–133.

Vandenberghe, R., Bohets, A., Claus, G., Vernelen, A., and Viaene, M. (1994). *De Determinanten van het Professionele Handelen van Leerkrachten Secundair Onderwijs en de Invloed op de Onderwijskwaliteit [The Determinants of Secondary Teacher's Professional Capacities and the Influence on the Quality of Education]*. Leuven: Katholieke Universiteit Leuven.

Veenhoven, R. (2000). The four qualities of life. *J. Happiness Stud.* 1, 1–39.

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.

Copyright © 2020 Govorova, Benítez and Muñiz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 26 May 2020
doi: 10.3389/fpsyg.2020.00876

Predicting Mathematics Achievement in Secondary Education: The Role of Cognitive, Motivational, and Emotional Variables

Amanda Abín[1], José Carlos Núñez[1,2], Celestino Rodríguez[1], Marisol Cueli[1], Trinidad García[1] and Pedro Rosário[3]*

[1] Department of Psychology, Universidad de Oviedo, Oviedo, Spain, [2] Facultad de Ciencias Sociales y Humanidades, Universidad Politécnica y Artística de Paraguay, Asunción, Paraguay, [3] Department of Applied Psychology, Escola de Psicologia, Universidade do Minho, Campus de Gualtar, Braga, Portugal

OPEN ACCESS

Edited by:
José Jesús Gázquez,
University of Almería, Spain

Reviewed by:
Miguel Angel Carbonero Martín,
University of Valladolid, Spain
Ana Miranda,
University of Valencia, Spain

***Correspondence:**
José Carlos Núñez
jcarlosn@uniovi.es

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

Received: 28 December 2019
Accepted: 08 April 2020
Published: 26 May 2020

Citation:
Abín A, Núñez JC, Rodríguez C, Cueli M, García T and Rosário P (2020) Predicting Mathematics Achievement in Secondary Education: The Role of Cognitive, Motivational, and Emotional Variables. Front. Psychol. 11:876. doi: 10.3389/fpsyg.2020.00876

Academic achievement in general, and in mathematics in particular, is positively associated not only with cognitive abilities, but also with emotional and motivational skills. The objective of this study was to analyze the prediction strength of cognitive, motivational, and emotional variables in mathematics achievement throughout high school, considering students' gender and age. A large sample of 2,365 Spanish students from the 4 years of high school (12–16 years old) participated in the study. Students provided information about their intellectual skills, perceived competence in mathematics, perceived utility of mathematics, intrinsic interest in learning, mathematics anxiety, and their causal attributions (for failure and for success), and of their achievement in mathematics. Data showed differences according to gender and the school grade level. The motivational and affective variables did not seem to play an important role in this relationship as predicted in the current study. The results of this study are discussed in light of previous research.

Keywords: intellectual abilities, perceived usefulness, perceived competence, anxiety, intrinsic motivation, achievement motivation, mathematics achievement

INTRODUCTION

Researchers' growing interest in studying mathematical achievement is driven by the importance of mathematics in both formal education and people's daily lives (Jansen et al., 2013; Namkung et al., 2019). Jain and Dowson (2009), for example, underlined the fact that mathematical comprehension is crucial for personal and professional success. Furthermore, Lipnevich et al. (2016) noted that success in mathematics is related to well-being, satisfaction with life, health, income, employability, and longevity.

Extant research has analyzed the influence of cognitive variables on mathematics achievement, but researchers have paid little attention to the role of emotional or motivational variables

(see Miñano and Castejón, 2011). Specifically, these authors found that intelligence did not explain a higher proportion of academic achievement than that provided by variables of an emotional or motivational nature. More recently, García et al. (2016a) concluded that motivation and enjoyment of mathematics were powerful predictors of mathematics achievement. Similarly, Lipnevich et al. (2016) stated that although intelligence was a significant predictor of mathematics achievement, attitudes toward this subject were key to explanation students' higher achievement. In short, and consistent with Zimmerman (2008), findings indicated that students' skills and abilities did not offer complete explanations about the magnitude or nature of mathematics achievement. In sum, perceived competence, perceived utility, motivation, and academic achievement may be considered related constructs. For this reason, the present work is aimed to examine the prediction strength of cognitive, motivational, and emotional variables in mathematics achievement, considering students' gender and age.

Perceived Competence, Perceived Utility, Motivation, and Emotions

Perceived competence in mathematics is defined as student perceptions about themselves as learners and of their capacity to successfully tackle mathematics tasks. This perception may fit reality to a greater or lesser extent, but in any case, it is a relevant source of students' motivation (García et al., 2016b). Literature reports a close association between students perceiving themselves as more capable in a particular subject and them being more willing to commit themselves to tasks related to that subject (for example, Pajares, 2008; Cabanach et al., 2009; Rosário et al., 2009). For example, Peixoto et al. (2017) have reported perceived competence to be strongly, significantly, and positively related to mathematics achievement. Similar results have been found with Portuguese students from fifth to ninth grade (Rosário et al., 2012), with British adolescents (Tosto et al., 2016) and with ninth and tenth-grade students from the US (Stevens et al., 2004).

However, feeling oneself to be capable may not be sufficient to explain personal commitment with academic tasks. Furthermore, task commitment needs to be perceived as useful. Perceived utility of mathematics refers to students' understanding about the applicability and benefits of learning that subject to their lives (Adelson and McCoach, 2011). Findings on the association between the perception of the value of the subject and their ability to learn new concepts and achieve higher in mathematics are mixed, while some researchers found positive relationships (Guy et al., 2005; Jordan et al., 2010). Other studies, did not find a positive association between perceived utility and the use of self-regulated learning strategies to improve the quality of learning (e.g., Cerezo et al., 2019). According to these authors, this could be because students often do not perceive a meaningful relationship between the use of cognitive strategies, high-quality learning, and their academic achievement.

As already noted, students' involvement in deep learning need both actual and perceived cognitive abilities, but also a strong motivation on the task (Suarez-Alvarez et al., 2014). Motivation for learning may be defined as the initiation of a learning process, the direction set, and the perseverance in path chosen. The relationship between academic motivation and mathematics achievement has been well studied both in elementary (e.g., Mercader et al., 2017) and in high school (e.g., Moenikia and Zahed-Babelan, 2010). Recently, Hammoudi (2019) reported that students more motivated were more willing to find learning opportunities and achieved higher in mathematics than their counterparts. However, and regardless of the theoretical model considered, it is essential to distinguish between motivation for success, or the tendency to succeed in the realization of a task – achievement goals (Wigfield et al., 2015), and intrinsic motivation, or the will to improve mastery on the task – mastery goals (Rodríguez et al., 2001; Murayama et al., 2013). In fact, although both types of motivation are positively related to perceived competence for mathematics (Hammoudi, 2019), intrinsic motivation is related to higher enjoyment of mathematics, higher effort displaced (García et al., 2016a), and higher success rates, even when the difficulty level is high. Recent literature defended the idea that students can pursue both intrinsic and achievement goals. The focus chosen is related to their personal characteristics, the nature of the task, and contextual variables (Wormington and Linnenbrink-Garcia, 2017). In addition, more than 30 years ago, Weiner (1986) found that the strength of students' motivation to learn was closely related to their reactions to academic successes and failures. Causal attributions may be defined as the explanations people ascribe to their successes and failures and play a determinant effect on students' motivation and academic achievement (González-Castro et al., 2014). In general, it has been shown that the more adaptive attributional patterns, the more the school achievement (i.e., success is attributed to internal and stable causes, and failure is attributed to changeable, but also internal causes) (Miñano and Castejón, 2011; Miranda et al., 2012).

Finally, learning in general and mathematics tasks in particular are experienced with certain amount of anxiety and a variety of emotional reactions (Rosário et al., 2008). Recently, Chang and Beilock (2016) related motivation with anxiety about mathematics. Math anxiety is the sensation of unease and worry felt when thinking about mathematics or while doing a mathematics task (Buckley et al., 2016). More specifically, math anxiety is characterized by negative feelings toward mathematics, which is likely to result in avoiding mathematics classes and show low math skills (Pizzie and Kraemer, 2017). In summary, many authors have emphasized the strong relationships between math anxiety, motivation, and mathematics achievement, noting that the lower the student's perceived competence in mathematics, the lower the motivation and the performance in mathematics (Lee and Stankov, 2013; Chang and Beilock, 2016; Passolunghi et al., 2016; Henschel and Roick, 2017).

Gender and Age

The relationship between cognitive, motivational, and affective variables and achievement in mathematics is significantly influenced by students' gender and age. Achievement in mathematics seems to vary depending on students' gender. However, although some researchers have indicated that for 30 years the gender gap in mathematics achievement has been in

favor of boys (García et al., 2007; Williams et al., 2016), others have reported that, mainly in countries with equal education for both sexes, boys and girls exhibit few or no differences in mathematical achievement (e.g., Spelke, 2005). As far as our knowledge, there are no data about gender differences regarding the predictive power of cognitive, motivational, or affective variables in relation to mathematics achievement.

When it comes to age, both transversal (e.g., Roskam and Nils, 2007) and longitudinal studies (e.g., Pintrich, 2000; Peetsma et al., 2005), have systematically reported that mathematics results diminish throughout schooling during adolescence. In addition, various studies have indicated that the motivational and emotional variables related to mathematics tend to change over time (Dowker et al., 2016), with perceived competence, perceived utility, intrinsic motivation, and even anxiety diminishing as students go through their schooling (Dowker, 2005; Mata et al., 2012). Nonetheless, as occurs with the gender variable, there is still little information about the interaction between student age and the predictive power of cognitive, motivational, or affective variables in mathematics achievement (Namkung et al., 2019).

The Current Study

Prior research has been analyzing the relationships between perceived competence, perceived utility, and math anxiety together with motivational variables and academic achievement (Miñano and Castejón, 2011; Lambic and Lipkovski, 2012; Chang and Beilock, 2016). As already noted, literature has reported recurrently positive relationships between mathematics achievement and cognitive competence, perceived competence, motivation (both intrinsic and for success), and adaptive attributional patterns. In addition, negative relationships with anxiety have been reported. However, there is little information about the predictive power of these variables in mathematics achievement when analyzed together. Furthermore, literature lacks information on the effects of the interaction of gender and age while estimating the effect size for each of these variables in mathematics achievement.

Consequently, in this study, we analyze the prediction strength of cognitive, motivational, and affective variables in mathematics achievement, considering students' gender and age. Grounded on data from previous research, the following hypotheses were set:

1. Cognitive variables (intellectual abilities), along with motivational variables (perceived competence, perceived utility, intrinsic and success motivation, and causal attributions for success and failure), and emotional variables (math anxiety) are good predictors of mathematics achievement.
2. The strength of the association between intellectual abilities and mathematics achievement is lower than that of the motivational or emotional variables.
3. The predictive power of the cognitive, motivational, and emotional variables in mathematics achievement varies depending on the students' age and gender.

MATERIALS AND METHODS

Participants

Participants in this study were 2,365 secondary school students from various schools in Asturias in the North of Spain. The total high school population (9th to 12th grade) in Asturias is ∼30,000. Data from the international PISA (2018) indicate that adolescents in Asturias scored slightly higher in mathematics than the Organization for Economic Cooperation and Development (OECD) average (Asturias = 491; OECD = 489), but slightly lower than the European Union average (494). The PISA report also indicates that despite boys scored higher in mathematics achievement than girls, the differences were not statistically significant. It is estimated that in the Asturian population as a whole, 7% are immigrants. Data indicate that immigrant students score far below non-immigrants (about 15 points, which is equivalent to a school grade level gap). Concerning socioeconomic status, the PISA report presents Asturias on the OECD average level. There is no evidence of differences between the schools in Asturias (which may be interpreted as an index of educational equality).

The current sample selection procedure was not random; although schools were initially chosen at random, not all agreed to participate. In addition, within the schools, a small number of students declined to participate for various reasons (e.g., being absent in one or more of the evaluation sessions, lack of permission from the family).

The sample subgroups by gender were similar sizes [girls: n = 1180 (49.9%); boys: n = 1185 (50.1%); $z = -0.145$, $p = 0.884$], although there were significant differences with respect to school grade level [1st year: n = 465 (19.7%), 2nd year: n = 487 (20.6%), 3rd year: n = 731 (30.9%), 4th year: n = 682 (28.8%); $\chi^2(3) = 92.462$, $p < 0.001$]. The gender distribution in each school grade was balanced, with no statistically significant differences: 1st year (50.5% girls; $z = 0.327$, $p = 0.743$), 2nd year (47.8% girls; $z = -1.345$, $p = 0.178$), 3rd year (48.4% girls; $z = -1.203$, $p = 0.229$), and 4th year (52.5% girls; $z = 1.843$, $p = 0.065$). The study did not include students with special educational needs or learning difficulties.

Instruments

Intellectual Abilities

To evaluate students' intellectual abilities, we used the Triarchic Intelligence Test (STAT). This is a test to measure intellectual abilities according to the Triarchic theory of intelligence (Sternberg, 1993). Its structure is the result of combining the three types of thinking (analytical, creative, and practical) with the content (verbal, numerical, and figurative). Although it is possible to get a score for each subscale, in this study, we only used the total test score. The test has adequate validity and reliability (Sternberg et al., 2001).

Motivational and Affective Variables

We measured perceived competence, perceived utility, intrinsic and success motivation, causal attributions, and anxiety from

the responses of the students to the Spanish adaptation of the Inventory of Attitudes Toward Mathematics from Fennema and Sherman (1978). In this adaptation, the dimensions used show satisfactory reliability (Cueli et al., 2014): perceived competence (six items, e.g., I believe I can do even the most difficult mathematics tasks; α = 0.85), perceived utility (eight items, e.g., Mathematics is a valuable and necessary subject; α = 0.85), intrinsic motivation (eight items, e.g., I find mathematics enjoyable and stimulating; α = 0.77), motivation for success (five items, e.g., I would like to be one of the best at mathematics; α = 0.86), math anxiety (six items, e.g., Normally, mathematics makes me nervous and uneasy; α = 0.78), attribution of success to internal causes (two items, e.g., I am convinced that to get good grades in mathematics you have to be intelligent; α = 0.71), and attribution of success to external causes (four items, e.g., To get good grades in mathematics, above all you have to be lucky; α = 0.78).

Mathematics Achievement

Data about the students' achievement in mathematics were gathered from the final grades in the subject. The secretaries of the participating schools with the permission of the parents provided data.

Procedure

The study was conducted in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki), which reflects the ethical principles for research involving humans (Williams et al., 2016). The study had the approval of the pertinent Ethical Committee of the Principality of Asturias (reference: CPMP/ICH/135/95, code: TDAH-Oviedo), and all procedures were performed in compliance with relevant laws and institutional guidelines. Data related to the predictor variables (cognitive, motivational, and affective) were collected 3 months before the mathematics tests were taken. Three qualified educational psychologists of the research project visited the schools and collected the data. Parents were informed about the study by the school authorities, and once they were assured of data confidentiality policy, they were asked to sign the informed consent document.

Data Analysis

Data were analyzed in two stages. Firstly, the descriptive data, correlation matrix, and distribution of means were examined, along with missing values (0.2%). Secondly, we performed various hierarchical regression analyses using SPSS 24. The strength of the associations and effect sizes were evaluated using R^2 (where R^2 <0.01 = null; R^2>0.01 and <0.09 = low/slight; R^2>0.09 and <0.25 = medium/moderate; R^2>0.25 = high/strong) and Cohen's d (1988), where $d < 0.20$ indicates a minimal effect size, $0.20 < d < 0.50$ indicates a small effect size, $0.50 < d < 0.80$ indicates a moderate effect size, and $d > 0.80$ indicates a large effect size.

RESULTS

Preliminary Analysis

Table 1 presents the descriptive statistics and the Pearson correlation matrix. The result of the KMO test indicated that the sampling was adequate (KMO = 0.733), and the Bartlett Sphericity test suggested that the matrix was suitable for multivariate analysis (χ^2 = 16556.93, p <0.001). According to the values for asymmetry and kurtosis, and according to commonly accepted criteria, the variables in the study complied with the criteria for univariate normality (Gravetter and Wallnau, 2014).

Prediction of Mathematical Achievement: Overall Sample

The hierarchical regression analysis was performed in three phases: (a) firstly, only intellectual ability was included as the sole predictor (model 1); (b) secondly, gender and school grade level were added as predictors (model 2); and (c) thirdly, perceived competence, perceived utility, intrinsic motivation, motivation for success, attribution of success and failure to internal or external causes, and math anxiety were added (model 3). The results are shown in **Table 2**.

Data show that intellectual abilities were strong, positive predictors of mathematics achievement (students with higher intellectual abilities tended to achieve higher results than students with lower intellectual abilities). Nonetheless, although the amount of explained variance was low (6.3%), the predictive capacity was similar in the two subsequent models, with a moderate effect size (d = 0.538). In fact, the predictive capability hardly suffered as a consequence of the inclusion of gender and school grade level (model 2) or the motivational and affective variables (model 3).

Gender and school grade level were also predictors of mathematics achievement, with a low percentage of explained variance (and both with small effect sizes: d = 0.307 and d = 0.257, respectively). This association was stable even after including the motivational and affective variables in the regression model (model 2 vs. model 3). Girls tended to show higher mathematics achievement than boys, $F_{(1,2365)}$ = 24.234; p < 0.001; η^2 = 0.010, although the effect size for these differences was small (d = 0.20). As students progressed through the school years, their mathematics achievement tended to fall, $F_{(3,2381)}$ = 30.261; p <0.001; η^2 = 0.037, again with a small effect size (d = 0.39).

Data indicated that including the motivational and affective variables in the model was statistically significant, $F_{(2,2354)}$ = 63.341; p <0.001, with a moderate strength for the association: R^2 = 0.14. From the seven variables included in the third model, the only predictors of mathematics achievement were perceived competence (albeit with a small effect size; d = 0.39), the perceived utility of mathematics (with a very small effect size; d = 0.17), intrinsic motivation (again with a small effect size; d = 0.29), and the attributions of successes and failures to external causes for which, although statistically significant at p < 0.05, the size of the coefficient of prediction was not significant (d = 0.08). Neither motivation for success, nor attribution to

TABLE 1 | Descriptive statistics and Pearson correlation.

	1	2	3	4	5	6	7	8	9	10	11
1. Mathematics achievement	–	0.253**	−0.101**	−0.169**	0.386**	0.339**	0.340**	0.171**	−0.225**	−0.071**	−0.218**
2. Intellectual abilities		–	0.043*	0.226**	0.108**	0.111**	0.073**	0.068**	−0.108**	0.005	−0.135**
3. Gender (0 = girls; 1 = boys)			–	−0.017	0.088**	−0.033	0.087**	−0.007	−0.142**	0.144**	0.165**
4. School year				–	−0.250**	−0.215**	−0.197**	−0.157**	0.146**	0.004	0.071**
5. Perceived competence					–	0.450**	0.517**	0.312**	−0.476**	−0.054**	−0.176**
6. Perceived utility of mathematics						–	0.467**	0.388**	−0.268**	−0.149**	−0.478**
7. Intrinsic motivation							–	0.305**	−0.378**	−0.041*	−0.188**
8. Motivation for success								–	−0.074**	0.110**	−0.177**
9. Math anxiety									–	−0.051*	−0.203**
10. Attribution to internal causes										–	0.318**
11. Attribution to external causes											–
M	2.464	11.377	0.50	2.69	3.496	3.582	3.071	3.639	2.843	2.845	2.326
SD	1.293	5.471	0.500	1.087	0.852	0.857	0.748	0.980	0.885	1.095	1.095
Asymmetry	0.514	−0.029	−0.004	−0.276	−0.355	−0.078	0.038	−0.437	0.040	0.014	0.511
Kurtosis	−0.875	0.028	−2.002	−1.214	−0.290	−0.514	0.203	−0.327	−0.201	−0.682	−0.686

*$p < 0.05$; **$p < 0.01$.

TABLE 2 | Results of the regression analysis for the overall simple (*N* = 2365).

	Model 1	Model 2	Model 3
Intellectual ability	0.251***	0.312***	236***
Gender (0 girl, 1 boy)		−0.118***	−0.135***
School year (1st to 4th)		−0.243***	−0.119***
Perceived competence			0.215***
Perceived utility			0.099***
Intrinsic motivation			0.153***
Motivation for success			−0.025
Anxiety			0.010
Internal causal attribution			−0.005
External causal attribution			−0.042*
R^2	0.063	0.131	0.269
ΔR^2		0.068	0.138

*$p < 0.05$, ***$p < 0.001$.

TABLE 3 | Results of hierarchical regression models for the variable gender.

	Model 1	Model 2	Model 3
Girls (*n* = 1180)			
Intellectual ability	0.230***	0.280***	0.204***
School year (1st to 4th)		−0.232***	−0.110***
Perceived competence			0.203***
Perceived utility			0.067
Intrinsic motivation			−0.008
Motivation for success			0.136***
Anxiety			0.051
Internal causal attribution			−0.034
External causal attribution			−.022
R^2	0.053	0.104	0.228
ΔR^2		0.051	0.124
Boys (*n* = 1185)			
Intellectual ability	0.283***	0.346***	0.270***
School year (1st to 4th)		−0.257***	−0.129***
Perceived competence			0.239***
Perceived utility			0.136***
Intrinsic motivation			−0.047
Motivation for success			0.167***
Anxiety			−0.036
Internal causal attribution			−0.048
External causal attribution			−0.035
R^2	0.080	0.143	0.306
ΔR^2		0.062	0.164

***$p < 0.001$.

internal causes, nor anxiety were found to be predictors of mathematics achievement.

Finally, model 3 predicted a significant and relevant amount of the variability of mathematics achievement (with a large effect size: $R^2 = 0.27$). Nonetheless, it is important to note, as the data in **Table 2** shows, that while intellectual abilities explained a small amount of the variance in mathematics achievement ($R^2 = 0.063$), the motivational and affective variables explained a moderate amount of the variability in achievement ($R^2 = 0.138$).

Prediction of Mathematics Achievement by Gender

Table 3 presents the results of the regression analysis for the girls' and boys' samples. Data were similar for both subsamples and did not differ significantly from what has already been reported for the overall sample. Specifically, we learned that intellectual abilities, despite the low amount of variance explained ($R^2 = 0.050$ girls; 0.080 boys), were good predictors of mathematics achievement in both samples, even after the inclusion of the other variables. Likewise, the perceived capability for mathematics explained academic achievement to the same extent for boys and girls, with similar results for intrinsic motivation. However, perceived utility has not significantly predicted mathematics achievement for the girls sample, which

was not true for boys. Irrespective of the samples, for the other variables (i.e., motivation for success, attributional processes, and mathematics anxiety), data were not found to be significantly associated with mathematics achievement.

Finally, analyzing the coefficients of determination, we found that the three models explained a higher quantity of the variance in the boys than in the girls sample, with the effect size being moderate for the girls (R^2 = 0.228) and large for the boys (R^2 = 0.306) sample. In both samples, the amount of variance explained by intellectual abilities was small, while the variance explained by the motivational and affective variables was moderate (R^2 = 0.124 girls; 0.164 boys).

Prediction of Mathematical Achievement by School Year

Table 4 presents the results of the predictions of mathematics achievement in the four school grade levels. The following are some of the most interesting results.

Firstly, as students' progress through the school years, up to the third year, there was a significant fall in the importance of intellectual abilities while explaining mathematics achievement. In first and second years, the amount of variance was moderate (R^2 = 0.162 in 1st year; 0.134 in 2nd year), but small in the third and fourth years (R^2 = 0.067 in 3rd year; 0.062 in 4th year). At the same time, perceived competence was a significant predictor of mathematics achievement in all four school years, and there was no decrease over time. Secondly, we found that intrinsic motivation was also a good predictor of achievement, except in the first year, in which this relationship was not statistically significant. The remaining motivational and affective variables were not clear, consistent predictors of mathematics achievement in the four school years. Taken together, and considered as a trend, the variance explained by motivational and affective variables decreased as students progressed from the 1st to 4th year high school grades. For the four school grade levels, the size of the association between the predictor variables and mathematics achievement was moderate or medium (17.5, 16.9, 14, and 13% of the variance explained, respectively). Thirdly, we also found that, in general, the explained variance for mathematics achievement was higher for the first two school years (34.4 and 32.3% of the variance explained for mathematics achievement) than for the last two (22.3 and 20.6% of variance explained, respectively). For the first two school grade levels, the size of the association was large and moderate for the final two. Finally, regarding gender, with the exception of the first year, in which the association was not statistically significant, for the other three school grade levels, girls tended to be more likely in showing higher mathematics achievement than boys (although the effect size was small in all cases).

TABLE 4 | Results of hierarchical regression models for the variable school year.

	Model 1	Model 2	Model 3
School year: 1st year (*n* = 465)			
Intellectual ability	0.403***	0.408***	0.305***
Gender (0 girl, 1 boy)		−0.085***	−0.058
Perceived competence			0.222***
Perceived utility			0.089
Intrinsic motivation			−0.029
Motivation for success			0.068
Anxiety			0.019
Internal causal attribution			−0.162***
External causal attribution			−0.031
R^2	0.162	0.169	0.344
ΔR^2		0.070	0.175
School year: 2nd year (*n* = 487)			
Intellectual ability	0.367***	0.372***	0.260***
Gender (0 girl, 1 boy)		−0.143***	−0.183***
Perceived competence			0.149***
Perceived utility			0.078
Intrinsic motivation			−0.013
Motivation for success			0.162***
Anxiety			0.130***
Internal causal attribution			−0.119***
External causal attribution			0.075
R^2	0.134	0.155	0.323
ΔR^2		0.021	0.169
School year: 3rd year (*N* = 731)			
Intellectual ability	0.259***	0.257***	0.204***
Gender (0 girl, 1 boy)		−0.126***	−0.154***
Perceived competence			0.271***
Perceived utility			0.106***
Intrinsic motivation			−0.013
Motivation for success			0.107***
Anxiety			−0.010
Internal causal attribution			0.052
External causal attribution			−0.024
R^2	0.067	0.083	0.223
ΔR^2		0.016	0.140
School year: 4th year (*n* = 682)			
Intellectual ability	0.250***	0.264***	0.207***
Gender (0 girl, 1 boy)		−0.117***	−0.139***
Perceived competence			0.184***
Perceived utility			0.073
Intrinsic motivation			−0.039
Motivation for success			0.213***
Anxiety			−0.056
Internal causal attribution			−0.041
External causal attribution			0.036
R^2	0.062	0.076	0.206
ΔR^2		0.014	0.130

***$p < 0.001$.

DISCUSSION

In this study, we aimed to assess the predictive capacity of a set of variables: cognitive variables (intellectual ability), motivational variables (perceived competence, perceived utility, intrinsic motivation, motivation for success, and attribution of causality for success and failure), and emotional variables (math anxiety) in determining students achievement in mathematics. Our goal was focused on determining not only their explanatory

power but also to further understand their interactions with the variables gender and school grade level. Although vast research has examined the predictive capacity of one or more of these variables, there are not much data analyzing them together, nor addressing whether the resulting predictive models would vary depending on variables such as gender and school grade level.

We formulated various hypotheses based on previous research. Firstly, we hypothesized that cognitive variables, motivational variables, and affective variables are good predictors of mathematics achievement. We also hypothesized that the size of the association between intellectual abilities and mathematics achievement is smaller than the size of the association between the motivational and emotional variables. Current data partially supported this general hypothesis.

In general terms, we found that students tended to be more likely to perform well in mathematics tasks when they had better intellectual abilities, higher perceived competence for mathematics, higher intrinsic motivation (i.e., interest in understanding mathematics and becoming more expert), and when they perceived mathematics to be useful. In line with previous research (e.g., Stevens et al., 2004; Miñano and Castejón, 2011; Lambic and Lipkovski, 2012; Miñano et al., 2012; Rosário et al., 2012; Tosto et al., 2016; Hammoudi, 2019), data showed the relationship of intellectual abilities and motivational variables (particularly perceived competence and intrinsic motivation, and perceived utility to a lesser extent). In addition, similarly to other studies (e.g., Miñano and Castejón, 2011; Miñano et al., 2012; García et al., 2016b; Lipnevich et al., 2016; Gilar-Corbi et al., 2019), we also concluded that the motivational variables were stronger predictors of mathematics achievement than the students' intellectual abilities.

In this regard, there are some aspects worth noting. Firstly, the fact that when it comes to explain student's achievement, their perceived capabilities are as important as their actual abilities (see also, Erturan and Jansen, 2015). This is interesting because perceived competence is a personal construction, and therefore can be modified according to student's experiences with mathematics. For this reason, teachers could consider helping students on their work, which offers the chance of successfully constructing confidence to tackle challenges and improve learning in mathematics. Secondly, it seems that at these ages, students still trust that what they learn in mathematics class will be useful; on the contrary, findings from Cerezo et al. (2019) indicate that college students fail to see the utility of what they are learning as a significant variable to organize their learning behaviors. For this reason, teachers and school administrators may wish to consider teaching learning strategies to help students link what they are learning with the near future (Cabanach et al., 2009; Rosário et al., 2015). Thirdly, as expected (e.g., Miñano and Castejón, 2011; García et al., 2016a; Lipnevich et al., 2016), the interest in learning a subject, such as mathematics, was associated with positive results. However, this did not happen, as our data showed, when learning mathematics was understood as an opportunity to outshine others or to gain some kind of reward. For this reason, the design of appropriate instructional strategies should include not only tasks focused on increasing students' self-confidence, and likely to be perceived as useful, but also tasks likely to increase students' interest and encourage them to deep their learning and compete with themselves rather than with their peers (Rosário et al., 2013).

Nonetheless, in contrast to some previous studies, the variables of a more emotional nature were not shown to be predictors of mathematics achievement (in either boys or girls samples), except in the second year of high school, in which anxiety and causal attribution processes significantly predicted mathematics achievement, thought. There may be various explanations for this.

When it comes to math anxiety, as mentioned in the beginning of this paper, prior data suggested a significant, strong, and negative relationship between anxiety and mathematics achievement (e.g., Rosário et al., 2008; Ashcraft and Moore, 2009; Miñano and Castejón, 2011; Maloney and Beilock, 2012; Miranda et al., 2012; Suárez-Pellicioni et al., 2015; Chang and Beilock, 2016; Passolunghi et al., 2016; Henschel and Roick, 2017). Firstly, and despite data from our study are not consistent with those results, they are in line with findings from Erturan and Jansen (2015), showing that when data are analyzed with regression equations, which include studying anxiety together with other variables (e.g., perceived competence for mathematics as predictors of mathematics achievement), anxiety is no longer predictive of mathematics achievement. Secondly, the magnitude of the relationship between anxiety and mathematics achievement could be affected by which dimension of anxiety is examined (Mammarella et al., 2018). Specifically, Dowker et al. (2016), and Henschel and Roick (2017) noted that the cognitive and affective dimensions of anxiety could be differently related to mathematics achievement. Similarly, Goetz et al. (2013) and Bieg et al. (2015) observed high levels of trait anxiety about mathematics, girls scoring higher, but this did not happen with state anxiety. Our findings could be related to the fact that the items of the questionnaire used, although not referencing very specific situations, could be understood as more related to state anxiety than to trait anxiety. Thirdly, another possible explanation may be related to the role of anxiety in the association with mathematics achievement and other variables such as perceived competence for mathematics (Erturan and Jansen, 2015). In a recent study, Pérez-Fuentes et al. (2020) attempted to learn whether mathematics anxiety, rather than directly predicting mathematics achievement, functioned as a mediating or moderating variable for other variables involved. In that study, they attempted to learn whether the relationship between perceived math ability and math achievement was mediated, at least partially, by anxiety, and whether it may even differ (in intensity or direction) depending on anxiety levels. Their results indicated that anxiety partially mediated, and moderated, the relationship between perceived competence and achievement. In terms of the moderating role, Pérez-Fuentes et al. (2020) found that when mathematics anxiety was high, the effect size of perceived competence for mathematics was large, whereas with low levels of anxiety, the effect was small. Authors suggested that when students experience high levels of math anxiety, the importance of their confidence in themselves grows significantly as a determinant of mathematics achievement. In contrast, when

anxiety is low, students' self-confidence is a much less strong determinant of achievement. However, more research is needed to confirm these findings.

The third hypothesis raised the possibility of the influence of gender and age on the predictors of mathematics achievement and of the magnitudes of these relationships. The direction of the impacts could not be further specified due to the limited available knowledge. In fact, to the best of our knowledge, available data only relates gender differences to some of the variables analyzed in the current study. For example, mathematics achievement (Spelke, 2005; García et al., 2007; Reilly et al., 2015; Fahle, 2016; Williams et al., 2016), achievement depending on the school grade level (Fahle, 2016), mathematics anxiety (Hill et al., 2016; Henschel and Roick, 2017) and perceived competence for mathematics (Henschel and Roick, 2017).

Regarding gender, data from our study add literature as follows. Firstly, there were no relevant gender differences regarding the predictor variables for mathematical achievement (i.e., intellectual abilities, perceived competence, or intrinsic motivation), although for boys, unlike the girls, the perceived utility of mathematics has shown to be a significant and positive predictor of mathematics achievement. Secondly, the variability in mathematics achievement that could be explained by the predictors was substantially higher in boys (large effect size) than in the girls (moderate effect size) sample. Thirdly, we found that in both samples, the predictive capacity of the non-cognitive variables (mainly perceived competence for mathematics and intrinsic motivation) was substantially higher than that shown by cognitive abilities (intellectual abilities). Whereas the non-cognitive variables exhibited a moderate predictive capacity, a small association was found for the cognitive abilities. Finally, it is worth noting that, consistent with recent studies (e.g., Erturan and Jansen, 2015), we did not find gender differences related to the magnitude of the association between anxiety and mathematics achievement, although there were differences in the direction of the relationship (positive for boys, negative for girls). As in the study by Erturan and Jansen (2015), in our research, perceived competence strongly and positively predicts performance in mathematics, for both boys and girls, but anxiety does not. So, we can conclude with Erturam and Jansen that "perceived math competence is more important in predicting performance than math anxiety" (p. 431).

With respect to the school grade level, this study adds literature by showing a decrease in the level of some of the variables taken as students' progress (e.g., a decrease in perceived competence, motivation, perceived utility of mathematics, and mathematics achievement; Peetsma et al., 2005; Roskam and Nils, 2007; Mata et al., 2012; Regueiro et al., 2015; Dowker et al., 2016). To be specific, we found that as students' progress throughout high school, the cognitive, motivational, and affective variables taken explain less of mathematics achievement. These findings indicate that mathematics achievement progressively depends less on the personal variables examined (e.g., intellectual abilities, perceived competence, motivation, anxiety, and attributional processes) and more on other variables: personal (e.g., personal engagement) and non-personal (e.g., school and family variables). In fact, despite the fact that it is reasonable to think that students' learning and achievement depend to a certain extent on family, school, and teaching variables, it is also expected that the main strong factors would be those personal to the students themselves (cognitive, motivational, and emotional). Thus, although the cognitive, motivational, and emotional variables considered in this study explain a significant proportion of the variability in mathematics achievement (with a large effect size), 70% of the variance remains to be explained. Although it may seem like a key strength of this study, it is clearly a shortcoming, since the remaining 70% have cleared educational and research implications. It does not seem feasible that 70% of adolescents' mathematics achievement can be explained by variables external to the student. It is reasonable to think that the different non-personal conditions (family, school, and teaching) may be important in students' learning and development, but through their influence on student variables (e.g., mainly those that can be changed, such as perceived competence, motivation, attitudes toward mathematics, attributional processes, anxiety) rather than separately from them. Future research, perhaps through causal relationship models, preferably with longitudinal, or repeated measure designs, should examine this idea more deeply.

In sum, considering the results of the present work, there are some educational implications that is necessary to highlight. First, if teachers focus in the cognitive skills of students in order to analyze or predict their academic results, they would be omitting important factors as their motivational situation. In this sense, beyond other variables of emotional nature, working on the perceived competence, intrinsic motivation, and perceived utility of students could have a positive impact on their mathematics achievement, especially in the first years of high school. Also, teachers have to consider other variables in their professional practice (e.g., family environment, instructional processes or math, or practice implicit theories).

Finally, it is important to note that despite the fact data in this study collected data from a wide sample of students and were representative in terms of gender and school grade level, it should be taken with caution when generalizing to different educational communities or to societies with substantially different educational systems. Nonetheless, the fact that mathematics is important for all of the OECD countries might reduce the likelihood of bias in generalizing these results. It is also essential to bear in mind that data about motivational and affective variables were collected by self-reports, which may bring bias. However, most of the research reviewed also used self-reported data, which should facilitate comparison.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Committee of the Principality of Asturias.

Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

JN, AA, and CR contributed conception and design of the study. MC and TG organized the database. JN and PR performed the statistical analysis. CR, AA, and MC wrote the first draft of the manuscript. PR, TG, and JN wrote sections of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

FUNDING

This research was partially funded by Principality of Asturias (FC-GRUPIN-IDI/2018/000199).

REFERENCES

Adelson, J. L., and McCoach, D. B. (2011). Development and psychometric properties of the math and me survey: measuring third through sixth graders' attitudes toward mathematics. *Measur. Eval. Counsel. Dev.* 44, 225–247. doi: 10.1177/0748175611418522

Ashcraft, M., and Moore, A. (2009). Mathematics anxiety and the affective drop in performance. *J. Psychoeduc. Assess.* 27, 197–205. doi: 10.1177/0734282908330580

Bieg, M., Goetz, T., Wolter, I., and Hall, N. C. (2015). Gender stereotype endorsement differentially predicts girls' and boys' trait-state discrepancy in math anxiety. *Front. Psychol.* 6:1404. doi: 10.3389/fpsyg.2015.01404

Buckley, S., Reid, K., Goos, M., Lipp, O., and Thomson, S. (2016). Understanding and addressing mathematics anxiety using perspectives from education, psychology and neuroscience. *Austr. J. Educ.* 60, 157–170. doi: 10.1177/0004944116653000

Cabanach, R., Valle, A., Gerpe, M., Rodríguez, S., Piñeiro, I., and Rosário, P. (2009). Design and validation of a motivational management questionnaire. *Rev. Psicodidact.* 14, 29–47.

Cerezo, R., Fernández, E., Amieiro, N., Valle, A., Rosário, P., and Núñez, J. C. (2019). The mediating role of self-efficacy and perceived usefulness between strategy knowledge and its use. *Rev. Psicodidact.* 24, 1–8. doi: 10.1016/j.psicoe.2018.09.001

Chang, H., and Beilock, S. L. (2016). The math anxiety-math performance link and its relation to individual and environmental factors: a review of current behavioral and psychopysiological research. *Curr. Opin. Behav. Sci.* 10, 33–38. doi: 10.1016/j.cobeha.2016.04.011

Cueli, M., González-Castro, P., Álvarez, L., García, T., and González-Pienda, J. A. (2014). Affective-motivational variables and performance in mathematics: a bidirectional analysis. *Rev. Mexican. Psicol.* 31, 153–163.

Dowker, A., Sarkar, A., and Looi, C. Y. (2016). Mathematics anxiety: what have we learned in 60 years? *Front. Psychol.* 7:508. doi: 10.3389/fpsyg.2016.00508

Dowker, A. D. (2005). *Individual Differences in Arithmetic: Implications for Psychology Neuroscience and Education*. Hove: Psychology Press.

Erturan, S., and Jansen, B. (2015). An investigation of boys' and girls' emotional experience of math, their math performance, and the relation between these variables. *Eur. J. Psychol. Educ.* 30, 421–435. doi: 10.1007/s10212-015-0248-7

Fahle, E. (2016). Patterns of change in U.S. gender achievement gaps during elementary and middle school. *Croatian Soc. Res. Educ. Effect.*

Fennema, E., and Sherman, J. A. (1978). Sex-related differences in mathematics achievement and related factors: a further study. *J. Res. Math. Educ.* 9, 189–203. doi: 10.3102/00028312014001051

García, M. I. B., Tello, F. P. H., Abad, E. V., and Moscoso, S. C. (2007). Actitudes, hábitos de estudio y rendimiento en Matemáticas: diferencias por género. *Psicothema* 19, 413–421.

García, T., Rodríguez, C., González-Castro, P., Torrance, M., and Gonzalez-Pienda, J. A. (2016a). Elementary students' metacognitive processes and post-performance calibration on mathematical problem-solving tasks. *Metacogn. Learn.* 11, 139–170. doi: 10.1007/s11409-015-9139-1

García, T., Rodríguez, C., Betts, L., Areces, D., and González-Castro, P. (2016b). How affective-motivational variables and approaches to learning predict mathematics achievement in upper elementary levels. *Learn. Individ. Differ.* 49, 25–31. doi: 10.1016/j.lindif.2016.05.021

Gilar-Corbi, R., Miñano, P., Veas, A., and Castejón, J. L. (2019). Testing for invariance in a structural model of academic achievement across underachieving and non-underachieving students. *Contemp. Educ. Psychol.* 59:101780. doi: 10.1016/j.cedpsych.2019.101780

Goetz, T., Bieg, M., Lüdke, O., Pekrun, R., and Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychol. Sci.* 24, 2079–2087. doi: 10.1177/0956797613486989

González-Castro, P., Rodríguez, C., Cueli, M., Cabeza, L., and Álvarez, L. (2014). Math competence and CPT variables in students with attention deficit/hiperactivity disorder and/or learning disabilities. *Rev. Psicodidact.* 19, 125–143. doi: 10.1387/RevPsicodidact.7510

Gravetter, F. J., and Wallnau, L. B. (2014). *Statistics for the Behavioral Sciences*. Belmont, CA: Wadsworth.

Guy, G. M., Cornick, J., and Beckford, I. (2005). More than math: on the affective domain in developmental mathematics. *Intern. J. Scholarsh. Teach. Learn.* 9, 1–7. doi: 10.20429/ijsotl.2015.090207

Hammoudi, M. H. (2019). Predictive factors of students' motivation to succeed in introductory mathematics courses: evidence from higher education in the UAE. *Intern. J. Math. Educ. Sci. Technol.* 50, 647–664. doi: 10.1080/0020739X.2018.1529339

Henschel, S., and Roick, T. (2017). Relationships of mathematics performances, control and value beliefs with cognitive and affective math anxiety. *Learn. Individ. Differ.* 55, 97–107. doi: 10.1016/j.lindif.2017.03.009

Hill, F., Mammarella, I. C., Devine, A., Caviola, S., Passolunghi, M. C., and Szucs, D. (2016). Maths anxiety in primary and secondary school student's: Gender differences, developmental changes and anxiety specifity. *Learn. Individ. Differ.* 48, 45–53. doi: 10.1016/j.lindif.2016.02.006

Jain, S., and Dowson, M. (2009). Mathematics anxiety as a function of multidimensional self-regulated and self-efficacy. *Contempo. Educ. Psychol.* 34, 240–249. doi: 10.1016/j.cedpsych.2009.05.004

Jansen, B. R. J., Louwerse, J., Straatemeier, M., Van der Ven, S. H. G., Klinkenberg, S., and Van der Maas, H. L. J. (2013). The influence of experiencing success in math on math anxiety, perceived math competence, and math performance. *Learn. Individ. Differ.* 24, 190–197. doi: 10.1016/j.lindif.2012.12.014

Jordan, N. C., Glutting, J., and Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learn. Individ. Differ.* 20, 82–88. doi: 10.1016/j.lindif.2009.07.004

Lambic, D., and Lipkovski, A. (2012). Measuring the influence of student's attitudes on the process of acquiring knowledge in mathematics. *Croatian J. Edu.* 14, 187–205. doi: 10.15516/cje.v14i1.90

Lee, J., and Stankov, L. (2013). Higher-order structure of noncognitive constructs and prediction of PISA 2003 mathematics achievement. *Learn. Individ. Differ.* 26, 119–130. doi: 10.1016/j.lindif.2013.05.004

Lipnevich, A. A., Preckel, F., and Krumm, S. (2016). Mathematics attitudes and their unique contribution to achievement: Going over and above cognitive ability and personality. *Learn. Individ. Differ.* 47, 70–79. doi: 10.1016/j.lindif.2015.12.027

Maloney, E. A., and Beilock, S. L. (2012). Math anxiety: who has it, why it develops, and how to guard against it. *Trends Cogn. Sci.* 16, 404–406. doi: 10.1515/9781400847990-016

Mammarella, I. C., Donolatoa, E., Caviolab, S., and Giofrèc, D. (2018). Anxiety profiles and protective factors: a latent profile analysis in children. *Pers. Individ. Differ.* 124, 201–218.

Mata, L., Monteiro, V., and Peixoto, F. (2012). Attitudes towards mathematics: effects of individual, motivational, and social support factors. *Child Dev. Res.* 2012, 1–10. doi: 10.1155/2012/876028

Mercader, J., Presentación, M. J., Siegenthaler, R., Molinero, V., and Miranda, A. (2017). Motivation and mathematics performance: a longitudinal study in early educational stages. *Rev. Psicodidáct.* 22, 157–163. doi: 10.1016/j.psicod.2017.05.007

Miñano, P., and Castejón, J. L. (2011). Cognitive and motivational variables in the academic achievement in language and mathematics subjects: a structural model. *Rev. Psicodidáct.* 16, 203–230.

Miñano, P., Gilar, R., and Castejón, J. L. (2012). A structural model of cognitive-motivational variables as explanatory factors of academic achievement in Spanish Language and Mathematics. *Anal. Psicol.* 28, 45–54.

Miranda, A., Colomer, C., Fernández, I., and Presentación, M. J. (2012). Executive functioning and motivation of children with attention deficit hyperactivity disorder (ADHD) on problem solving and calculation tasks. *Revista de Psicodidáctica* 17, 51–71.

Moenikia, M., and Zahed-Babelan, A. (2010). A study of simple and multiple relations between mathematics attitude, academic motivation and intelligence quotient with mathematics achievement. *Proc. Soc. Behav. Sci.* 2, 1537–1542. doi: 10.1016/j.sbspro.2010.03.231

Murayama, K., Pekrun, R., Lichtenfeld, S., and vom Hofe, R. (2013). Predicting long-term growth in students' mathematics achievement: the unique contributions of motivation and cognitive strategies. *Child Dev.* 84, 1475–1490. doi: 10.1111/cdev.12036

Namkung, J. M., Peng, P., and Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: a meta-analysis. *Rev. Educ. Res.* 89, 459–496. doi: 10.3102/0034654319843494

Pajares, F. (2008). "Motivational role of self-efficacy beliefs in self-regulated learning," in *Motivation And Self-Regulated Learning: Theory, Research And Applications*, eds D. Schunk and B. J. Zimmerman (New York, NY: Lawrence Erlbaum Associates), 111–141.

Passolunghi, M. C., Caviola, S., De Agostini, R., Perin, C., and Mammarella, I. C. (2016). Mathematics anxiety, working memory, and mathematics performance in secondary-school children. *Front. Psychol.* 7:42. doi: 10.3389/fpsyg.2016.00042

Peetsma, T., Hascher, T., van der Veen, I., and Roede, E. (2005). Relations between adolescents' self-evaluations, time perspectives, motivation for school and their achievement in different countries and at different ages. *Eur. J. Psychol. Educ.* 20, 209–223. doi: 10.1007/BF03173553

Peixoto, F., Sanches, C., Mata, L., and Monteiro, V. (2017). "How do you feel about math?": relationships between competence and value appraisals, achievement emotions and academic achievement. *Eur. J. Psychol. Educ.* 32, 385–405. doi: 10.1007/s10212-016-0299-4

Pérez-Fuentes, M. C., Núñez, A., Molero, M., Gázquez, J. J., Rosário, P., and Núñez, J. C. (2020). The Role of Anxiety in the Relationship between Self-efficacy and Math Achievement. *Psicología Educativa* doi: 10.5093/psed2020a7 [Epub ahead of print].

Pintrich, P. R. (2000). Multiple goals, multiple pathways: the role of goal orientation in learning and achievement. *J. Educ. Psychol.* 92, 544–555. doi: 10.1037/0022-0663.92.3.544

PISA (2018). *PISA Report. Asturias Results 2018*. Paris: PISA.

Pizzie, R. G., and Kraemer, D. J. M. (2017). Avoiding math on a rapid timescale: emotional responsivity and anxious attention in math anxiety. *Brain Cogn.* 118, 100–107. doi: 10.1016/j.bandc.2017.08.004

Regueiro, R., Suárez, N., Valle, A., Núñez, J. C., and Rosário, P. (2015). Homework motivation and engagement throughout compulsory education. *Rev. Psicodidáct.* 20, 47–73. doi: 10.1387/RevPsicodidact.12641

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

Rodríguez, S., González-Cabanach, R., Valle, A., Núñez, J. C., and González-Pienda, J. A. (2001). Approach goals, avoidance goals and multiple academic goals. *Psicothema* 13, 546–550.

Rosário, P., Costa, M., Núñez, J. C., González-Pienda, J., Solano, P., and Valle, A. (2009). Academic procrastination: associations with personal, school and family variables. *Span. J. Psychol.* 12, 118–127. doi: 10.1017/s1138741600001530

Rosário, P., Lourenço, A., Paiva, O., Rodrigues, A., Valle, A., and Tuero-Herrero, E. (2012). Prediction of mathematics achievement: effect of personal, socioeducational and contextual variables. *Psicothema* 24, 289–295.

Rosário, P., Núñez, J. C., Ferrando, P., Paiva, O., Lourenço, A., Cerezo, R., et al. (2013). The relationship between approaches to teaching and approaches to studying: a two-level structural equation model for biology achievement in high school. *Metacogn. Learn.* 8, 44–77. doi: 10.1007/s11409-013-9095-6

Rosário, P., Núñez, J. C., Vallejo, G., Cunha, J., Nunes, T., Mourão, R., et al. (2015). Does homework design matter? The role of homework's purpose in student mathematics achievement. *Contemp. Educ. Psychol.* 43, 10–24. doi: 10.1016/j.cedpsych.2015.08.001

Rosário, P., Salgado, A., Núñez, J. C., González-Pienda, J. A., Bernardo, A., Valle, A., et al. (2008). Test anxiety: associations with personal and family variables. *Psicothema* 20, 563–570.

Roskam, I., and Nils, F. (2007). Predicting intra-individual academic achievement trajectories of adolescents nested in class environment: Influence of motivation, implicit theory of intelligence, self-esteem and parenting. *Psychol. Belg.* 47, 119–143. doi: 10.5334/pb-47-1-119

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

Sternberg, R. J. (1993). *Sternberg Triarchic Abilities Test*. New Haven: Department of Psychology University of Yale.

Sternberg, R. J., Castejón, J. L., Prieto, M. D., Hautamäki, J., and Grigorenko, E. (2001). Confirmatory analysis of the sternberg triarchic abilities test in three international samples. *Eur. J. Psychol. Assessm.* 1, 1–16. doi: 10.1027//1015-5759.17.1.1

Stevens, T., Olivarez, A., Lan, W. Y., and Tallent-Runnels, M. K. (2004). Role of self-efficacy and motivation in mathematics performance across ethnicity. *J. Educ. Res.* 97, 208–221. doi: 10.3200/JOER.97.4.208-222

Suarez-Alvarez, J., Fernandez-Alonso, R., and Muñiz, J. (2014). Self-concept, motivation, expectations, and socioeconomic level as. *Learn. Individ. Differ.* 30, 118–123. doi: 10.1016/j.lindif.2013.10.019

Suárez-Pellicioni, M., Núñez-Peña, M. I., and Colomé, A. (2015). Math anxiety: a review of its cognitive consequences, psychophysiological correlates, and brain bases. *Cogn. Affect. Behav. Neurosci.* 16, 3–22. doi: 10.3758/s13415-015-0370-7

Tosto, M. G., Asbury, K., Mazzocco, M. M. M., Petrill, A., and Kovas, Y. (2016). From classroom environment to mathematics achievement: the mediating role of self-perceived ability and subject interest. *Learn. Individ. Differ.* 50, 260–269. doi: 10.1016/j.lindif.2016.07.009

Weiner, B. (1986). *An Attributional Theory Of Motivation And Emotion*. Londres: Springer-Verlag.

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., and Schiefele, U. (2015). "Development of achievement motivation and engagement," in *Handbook of Child Psychology and Developmental Science: Socioemotional Processes*, eds M. E. Lamb and R. M. Lerner (Hoboken, NJ: John Wiley & Sons Inc), 657–700.

Williams, K. E., White, S. L. J., and MacDonald, A. (2016). Early mathematics achievement of Boys and girls: do differences in early self-regulation pathways explain later achievement? *Learn. Individ. Differ.* 51, 199–209. doi: 10.1016/j.lindif.2016.09.006

Wormington, S. V., and Linnenbrink-Garcia, L. (2017). A new look at multiple goal pursuit: the promise of a person-centered approach. *Educ. Psychol. Rev.* 29, 407–445. doi: 10.1007/s10648-016-9358-2

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: historical, background, methodological developments, and future prospects. *Am. Educ. Res. J.* 45, 166–183. doi: 10.3102/0002831207312909

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.

The reviewer MC declared a past co-authorship with one of the authors PR to the handling Editor.

Copyright © 2020 Abín, Núñez, Rodríguez, Cueli, García and Rosário. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 28 May 2020
doi: 10.3389/fpsyg.2020.01100

Design, Implementation and Evaluation of an Emotional Education Program: Effects on Academic Performance

María-José Mira-Galvañ and Raquel Gilar-Corbi*

Departament of Developmental Psychology and Didactics, University of Alicante, Alicante, Spain

Background: In recent decades, the amount of research on social and emotional learning programs in schools has increased significantly, showing a great number of positive student outcomes, including greater ability to perceive, understand and manage emotions, better attitudes about self and others, less aggressive and/or disruptive behavior, higher levels of psychological well-being and improvement in academic performance among others. The purpose of this research was the design and implementation of the OKAPI emotional education program. A multidimensional program based on cooperative learning methodology.

Methods: 86 students of Primary Education, from 3rd to 5th grade (45 students in the experimental unit and 41 in the control group).

Results: The implementation of the OKAPI has a positive impact on academic achievement.

Conclusion: This program shows the convenience of incorporating programs that can be integrated into school life and can be applied by the teaching staff using both social emotional learning and school climate approaches.

Keywords: emotional education, academic performance, primary education, cooperative learning, quasi-experimental design

OPEN ACCESS

Edited by:
Michael S. Dempsey,
Boston University, United States

Reviewed by:
Antonio Valle,
University of A Coruña, Spain
Susana Rodríguez,
University of A Coruña, Spain

****Correspondence:***
María-José Mira-Galvañ
jose.galvan@ua.es

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

Received: *20 January 2020*
Accepted: *29 April 2020*
Published: *28 May 2020*

Citation:
Mira-Galvañ M-J and Gilar-Corbi R (2020) Design, Implementation and Evaluation of an Emotional Education Program: Effects on Academic Performance. Front. Psychol. 11:1100. doi: 10.3389/fpsyg.2020.01100

INTRODUCTION

Emotional education programs have been related, in the last decades, not only to the improvement of social competences such as the ability to perceive, manage or understand emotions (Extremera and Fernández-Berrocal, 2004; Lópes and Salovey, 2004; Brackett et al., 2006), but also with better results in academic performance (Petrides et al., 2004; Gil-Olarte et al., 2005; Palomera and Brackett, 2006; Collaborative for Academic Social and Emotional Learning [CASEL], 2007; Pérez and Castejón, 2007; Nasir and Masrur, 2010; Fernández-Berrocal et al., 2012), the increase in positive behaviors (Mestre et al., 2006), and higher levels of students' self-esteem (Salovey et al., 2002; Fernández-Berrocal et al., 2006), among others.

Unfortunately, the implementation of some social and emotional programs has not always yielded the expected results. This is due to the fact that the programs have not been fully incorporated into the school routine, or they are based only on direct instruction and without

full and continuous support for teachers and school staff (Durlak et al., 2011; Jones and Bouffard, 2012). There are bidirectional relations between children and contexts in which they learn and nurture their social, emotional and academic competence (Bronfenbrenner and Ceci, 1994). Students' social-emotional development is closely linked to the environment in which they develop. A positive school climate is required in order to support student learning and provide opportunities for social and emotional development. Both aspects have health benefits and a positive impact on students' life (Osher et al., 2014; Greenberg et al., 2017).

School climate and Social Emotional Learning are essential components of effective schools (Osher et al., 2004). This trends shows that the most effective programs create safe conditions to practice the skills learnt in the program, engage multiple members of the school community, provide professional development for teachers and school staff. Furthermore, teachers are in charge of their application. The current research seeks to prove the effectiveness of the OKAPI program, a multidimensional program that uses both school climate and social and emotional learning approaches and also incorporates active methodologies, such as cooperative learning, to encourage interaction among students.

Emotional Education

Emotional education is a continuous and permanent educational process that aims to promote emotional development as a complement to the cognitive development of the child. Emotions can facilitate or impede children's academic engagement and achievement (Zins et al., 2004; Gil-Olarte et al., 2006; Pérez and Castejón, 2007; Durlak et al., 2011; Blake et al., 2015). Effective mastery of social-emotional competences is associated with greater well-being and pro-social behavior (Greenberg et al., 2003). It can also reduce risk factors and prevent certain dysfunctions (Bisquerra, 2000, 2003).

According to Collaborative for Academic Social and Emotional Learning [CASEL] (2007), social and emotional learning programs (SEL) can be defined as the capacity to recognize and manage emotions, solve problems efficiently, and establish positive relationships with others. CASEL establishes five different competencies that reflect both intrapersonal and interpersonal spheres: self-awareness (the ability to recognize and comprehend one's own emotions, personal goals and values), self-management (the ability to regulate emotions and behaviors), social awareness (the ability to understand the perspective of others, no matter their cultural background), relationship skills (the ability to establish positive relationships) and responsible decision making.

In the past two decades, research on emotional education has shown that the most effective programs are the ones that deliver explicit lessons that teach social and emotional skills, provide opportunities for students to use those competences throughout the day (Weissberg et al., 2015) and have multi-year and multicomponent approaches (Zins et al., 2004).

In addition, Durlak et al. (2011) found that programs are likely to be effective if they are taught by teachers, involve the school community and are also evaluated during the process and at the end of the program. They also need to follow step-by-step SAFE criteria (sequenced, active, focused, and explicit): use a sequenced training approach, use active forms of learning, focus sufficient time on skill development, and have explicit learning goals (Durlak, 1997; Bond and Hauf, 2004).

There are three main ways of socializing emotions: through direct exposure and observation; making use of direct teaching and/or coaching and fostering opportunities for practice (Brenner and Salovey, 1997). That is why it is crucial teachers' emotional training. Teachers with high EI scores generate a more positive school climate (Extremera et al., 2003; Brackett et al., 2010) which helps the program implementation and gives more opportunities to practice the socials and emotional skills.

Cooperative Learning

From the perspective of social relationships, there are three different types of educational situations that take place in the classroom: individualistic, cooperative and competitive (Pujolàs, 2001; Johnson and Johnson, 2009; Arnaiz and Linares, 2010; Vega and Hederich, 2015). Students who work in a cooperative way can have higher levels of achievement because the collective production is higher in quantity and quality than individual production. It also allows students to develop their emotional intelligence through numerous socials interaction (Bisquerra et al., 2015) so they can develop social and interpersonal skills while they are working in teams (Díaz-Aguado, 2003; Linares, 2004).

The OKAPI program is based on a cooperative learning methodology. We define cooperative learning as: "a continuum of learners working together in a small group, so that everyone can participate in the collective task that has been clearly defined by the teacher. So students work together to learn and are responsible for one another's learning as well as their own" (Slavin, 1995).

In cooperative learning, students work in small and heterogeneous groups. It involves students working together to achieve common goals and requires the contribution of all group members in order to fulfill individual and group objectives. Students are responsible for their own learning process and, at the same time, are responsible of their peers' learning process. Peer interaction and peer tutoring are promoted. Knowledge is built collectively taking into account the classroom environment and social relationships.

Cooperative learning specifically contributes to the development of interpersonal intelligence, because the skills needed for social interaction are practiced on a daily basis and developed in an indirect way (Linares, 2004). In fact, there are important similarities between cooperative learning and training in social skills (León et al., 2014; Larraz et al., 2017). In cooperative learning students learn social skills through modeling, they practice those skills in real situations and they receive feedback concerning their behavior from their classmates (León et al., 2015). But the necessary socials skills for effective cooperative work need to be learned beforehand (Gillies and Haynes, 2011). So cooperative learning and social emotional programs are interdependent.

Over the past decade, cooperative learning has emerged as an one of the most effective approaches to classroom instruction. According to research, students who using cooperative learning methodology improve their academic achievement, have a greater number of positive social-skills and prosocial behavior, higher self-esteem and better motivation (Ovejero, 1993; Johnson et al., 1999; Pujolàs, 2001, 2005; Díaz-Aguado, 2003; Wilson-Jones and Caston, 2004).

OKAPI

The OKAPI Learning Environment program aims to create a positive classroom and school climate that enables the development of the socio-emotional competencies of all members of the educational community.

The program uses direct instruction to teach socio-emotional skills and it has different tools that are applied cross-sectionally and are incorporated into everyday routines. It also favors the creation of a positive environment where students, through interactions within the particular context, put into practice the socio-emotional competences. In order to implement the program, each class had to incorporate in their schedule 10 to 15 min during the morning to do an assembly, approximately 5 min to practice full attention after the recess and 30 min per week (3 times per month) to work in different activities related to the program.

The key components of the program respond to the Catalan acronym of OKAPI: O (Classroom Organization: routines, rules, operation, distribution of spaces, etc.), K (Cooperative Learning as a main methodology for the teaching-learning process) A (Positive classroom environment), P (Participation: Involvement of all members of the educational community) and I (Emotional intelligence: emotional education as a key point in the curriculum).

The main characteristics of the program can be summarized as follows:

Cooperative Learning

Cooperative learning has been used as a tool to improve the classroom climate. Every month, students do teambuilding activities and class building activities. These are specifically designed to stimulate peer interactions, promote social cohesion and to improve social skills. Cooperative learning was the main methodology used in the different subjects and it represents a basic tool for the program. Several cooperative learning strategies were used in the program such as Quiz, Quiz trade, Round Table or Rally Coach (Kagan, 1994) among others.

Positive Classroom Environment

Three aspects are emphasized to improve the classroom environment explicitly. First, effective and coherent plans to develop positive behaviors in a specific class and in the school. Second, strategies for conflict resolution: The VET technique (evaluates, explores, and transforms) to be used by teachers. And students put into practice the PiP technique (tool that encourages students to stop and think when they have a problem using questions to reflect on the situation) and MENU (tool to find strategies that guide students how to resolve conflicts with other people). Finally, the improvement of communication between families and schools using different strategies (student of the month, weekly self-report card, etc.).

Emotional Intelligence

Emotional competencies are taught through direct instruction, approximately two times per month in a 30 min class using the Bisquerra (2009) five block model: emotional conscience, emotional regulation, emotional autonomy, social competency, life skills and well-being. There were seven sessions of direct instruction. Also, after every recess, students perform full attention techniques and strategies for about 6 to 7 min (overall, at least 30 min per week).

Moreover, in the class schedule there is 10 to 15 min for classroom assembly, named in the OKAPI program as a "Start Circle," at the beginning of the day. In the Start Circle the teacher explains the planning of the day, students work with the mood meter (a technique used to label, manage and understand emotions) and it is a time where students can talk and share their thoughts and feelings. Also, every Friday during the "Start Circle," one word pertaining to emotions is introduced. The emotion is explained by the teacher, discussed in the assembly and associated to real situations. Students discuss that word with their families during the weekend and the following Monday, during the assembly it is explained again and introduced onto the mood meter.

Furthermore, in the language areas we use the strategy CiC (Tell and Share). After reading some stories, students, using a cooperative learning strategy, discuss the feelings of the character, consequences of their decisions and look for different alternatives.

Research Questions

This study seeks to prove the effectiveness of an emotional education and positive climate program based on cooperative learning methodology for primary school students. The purpose of this research is to analyze the contribution of the OKAPI program on academic achievement and the improvement of emotional competence in students.

The questions that guide this work are as follows: Can an emotional educational and positive climate program based on cooperative learning such as the OKAPI program improve the academic performance of students?

We hypothesize that we will observe statistically significant differences in academic performance and in social competence between the experimental and the control group.

METHODS

Participants

A non-probability convenience sampling was used to select the participants. Two different schools from Alicante Province participated in the current study. The selection of these schools were based on knowledge of the characteristics of both schools as well as their populations, and their willingness to participate in the study. Both schools have similar characteristics: they are

public schools, they have a similar size and they use the same textbooks. The total sample was composed of 86 students from two different primary schools.

Students from 3rd grade (9 to 10 years old), 4th grade (10 to 11 years old) and 5th grade (11 to 12 years old) participated in the current study. Overall, six different classrooms: three in the control group and three in the experimental group.

The experimental group consisted of 45 students: 15 from the third grade, 14 from the fourth grade and 16 from the fifth grade. In the control group there were 41 students: 14 from the third grade, 14 from the fourth grade and 13 from the fifth grade. In total, 61.7% were boys and 38.3% were girls.

Instruments

The following instruments were utilized in this study:

(1) Emotional Quotient Inventory: Young Version (EQ-i:YV) (Bar-On and Parker, 2000). This questionnaire measures the level of emotional and social functioning in children and adolescents providing an estimate of their underlying emotional and social intelligence. It consists of 30 items with values on a 5-point Likert-type scale and it evaluates five general factors of EI. The EQ-i:YV (s) has adequate internal reliability in every dimension: interpersonal scale ($a = 0.67$), intrapersonal scale ($a = 0.84$), stress management ($a = 0.84$), adaptability ($a = 0.83$) and the EQ-i total scale ($a = 0.77$). Emotional intelligence was measured through the student's self-report and through the teacher's evaluation of each of the students.

(2) Academic performance: This was evaluated using the average grades in the main subjects (Maths, Catalan language, Spanish language, Social science, and Natural science). These scores are the result of assessment tests designed by the teaching staff in accordance with the guidelines established by the competent authority in education.

Procedure

Prior to the implementation of the study, the researcher was given permission by the principals of each school, as well as teachers and parents of each and every participant in the study. We guaranteed confidentiality relating to all of the data obtained from both teachers and students. This study was carried out in accordance with the recommendations of the University of Alicante Ethics Committee. All participants gave informed written consent in accordance with the Declaration of Helsinki.

Teachers who would participate in the study received training during the previous school year. This training was part of an Innovation project named *Cooperative Learning as a tool to develop emotional competencies*. Teachers received conceptual training in Cooperative Learning, emotional intelligence and OKAPI components in 10 sessions, with a total of 20 h. Moreover, monthly supervision and training sessions during the intervention were implemented to ensure the accuracy and quality of the program.

The quasi experimental study was conducted over a period of 6 months, during a school year. The pre-test was administrated to the experimental and control group in November. The implementation of the Okapi program in the experimental group took place from December to May while the control group received normal classroom instruction. In June, the post-test was administered to both groups.

The OKAPI program was applied by the teachers although the person in charge of the research was in each classroom for at least 3 h per week in order to support teachers and carry out the more complex activities. At the same time, on a monthly basis, training sessions for teachers were conducted and also the teachers carried out self-evaluations as part of the program implementation.

Design

A quasi-experimental design with non-equivalent control-group and pre-test/post-test was used (Garcia Gallego, 2001). Quasi-experimental design was used because students weren't randomly assigned to classes by the researcher. The research was conducted over a period of 6 months. The interval between the pre-test and the post-test was 8 months.

Data Analysis

A robust statistical and methodological treatment based on the principles of the General Linear Model is used to analyze the data (Leon et al., 2003; Ganju, 2004).

The data were analyzed as follows. First, to ensure homogeneity between groups, a comparative analysis was carried out at the beginning of the research. The differences between the experimental and control group were assessed by using Student's *t*-test. This is used in order to know how significant the differences between the means of two groups are.

Second, to measure the effect of the intervention we used the GLM repeated measure. This statistical technique tests the main effects within and between subjects. Groups of related dependent variables are analyzed which represent different measures of the same attribute (Freeman, 1973; Vuchkov and Solakov, 1980). At least one of the factors is based on independent observations and another one is based on correlated observations. A multivariate variance analysis (MANOVA) and a repeated-measure univariate analysis of variance (ANOVA), in which the measures of dependent variables are treated as variables within subjects (intrasubjects) and groups are treated as variables between subjects, are done (Lavori, 1990; Stemmler, 1994; Keselman et al., 1998; Woodrow, 2014).

A comparison of means was performed to analyze if there were significant differences between dependent variables among control and experimental groups (Duby et al., 1977; Deleon and Atkinson, 1991). Null hypotheses are contrasted in terms of the effects of both intersubject and intrasubject factors, and the interactions between individual effects and factors are investigated (Gabriel and Hopkins, 1974).

Finally, we graphically represent the differences between the experimental and control groups in the pre-test and the post-test situations to interpret the differences. The analysis and graphical representation of the data were performed using SPSS v.20 with a license from the University of Alicante.

RESULTS

Initial Situation: Experimental and Control Groups

First, to analyze whether there were differences in EI levels (measured using the EQ-i:YV) in the two groups before the intervention, mean contrast for the independent samples was completed using independent *t* test (continuous-variables). The result showed that there were initial differences in some dimensions of the EQ-i:YV scale answered by the students. Results for the intrapersonal dimension shows differences between control and experimental groups [$t_{(80,74)} = -2.23$, $p = 0.03$]. The control group obtains higher scores than the experimental group. In the stress management dimension there were also initial differences. In this case, the control group also has higher scores [$t_{(84)} = 3.31$, $p = 0.00$]. The EQ-i:YV scale answered by the teachers referring to the students shows differences between the control and the experimental group in the intrapersonal dimension [$t_{(84)} = 3.84$, $p = 0.00$]. So that we can see that the control group has higher scores.

Analysis of the Effects of the Program

Results of a Box's *M* test did not show homogeneity of the variance-covariance matrix for the global average academic performance of the five subjects' scores used in the research. The value of The Box's *M* test was significant ($F = 3.02$; $gl = 1777031.77$; $p = 0.03$). The violation of this assumption has a minimum effect if the groups are approximately equal in size (Hair et al., 1999). Because there are only two levels, sphericity is assumed.

To evaluate the impact of the program on academic performance, the scores obtained by both groups were compared before and after the intervention. As can be seen in **Table 1**, the interaction between the evaluation time (pre-test and post-test) and the program is significant ($p \leq 0.05$).

There are significant differences between the students' average scores. Students in the experimental group slightly improve their academic performance once the intervention program has been completed. Nevertheless, the observed power (testing the null hypothesis of no effect) doesn't have a high value (0.69), so the size of the effect (η^2) is not as desired. The differences in the averages of the scores can be seen in the **Figure 1**.

To evaluate the effectiveness of the program in students' emotional competence, a GLM of repeated measures was also used. The effects of the intra-subject and inter-subject show that the effect of the interaction between the time of the evaluation (pre-test and post-test) and the implementation of the program is significant ($p \leq 0.05$) (**Table 2**). There are significant differences between the students' scores in the EQ-i: YV (S) as can be seen in **Figure 2**. However, the effect of the size (η^2) is not as high as desired, maybe it is because of the small size of the sample.

DISCUSSION

During the past two decades the interest in social-emotional learning (SEL) has increased exponentially. Schools and teachers are aware of the importance of social-emotional competence in students' cognitive development and their link to academic outcomes (Durlak et al., 2011). Despite that, SEL is not fully incorporated in every school. The lack of time and resources and the increasing demands on schools in terms of their requirements (diversity, inclusion, etc.) on a daily basis present an obstacle which is difficult to overcome.

According to Jones et al. (2017) in their review of effective SEL programs, it is not unusual that some schools, after implementing a SEL program, do not see the results they expected based on previous research. This could be due to the superficial implementation of the program in the school without being part of the school life.

OKAPI program has been designed to be integrated into the classes' routines, to be taught using the existing curriculum (Bear et al., 2017) and with an approach that is positive and sensitive to the school environment. The purpose of the program was to be a unique tool to help teachers to improve class climate, student well-being, social competence, and academic performance.

Several studies have found that SEL programs have better results if they are conducted by teachers (Durlak et al., 2011). That's why the design of the program, activities and strategies used were created as purposefully uncomplicated to facilitate their implementation.

The results of the study provide preliminary data supporting the effectiveness of the OKAPI program on students' academic performance and emotional competence. The profile graphs show the differences obtained by the groups (experimental and control) in the pre-test and post-test situations, allowing us to visualize the meaning of the differences in favor of the experimental group in the variables studied. In reference to academic performance, the students' average scores have risen in the experimental group from 7.2 points in the pre-test to almost 7.6 points on the post-test. It is a small improvement on the overall scores but it shows a tendency. In every subject evaluated the experimental group has increased their scores.

TABLE 1 | Summary of intra-inter subject univariate analysis of variance: Academic Performance.

	Font	Type III	*gl*	*F*	Significant	η^2 partial	Observed power
Academic Performance	Intra	1.08	1	4.61	0.03	0.05	0.56
	Intra × Inter	1.44	1	6.14	0.01	0.07	0.69
	Error Intra	19.71	84	0.23			
	Inter	0.16	1	0.05	0.82	0.00	0.05
	Error Inter	274.82	84				

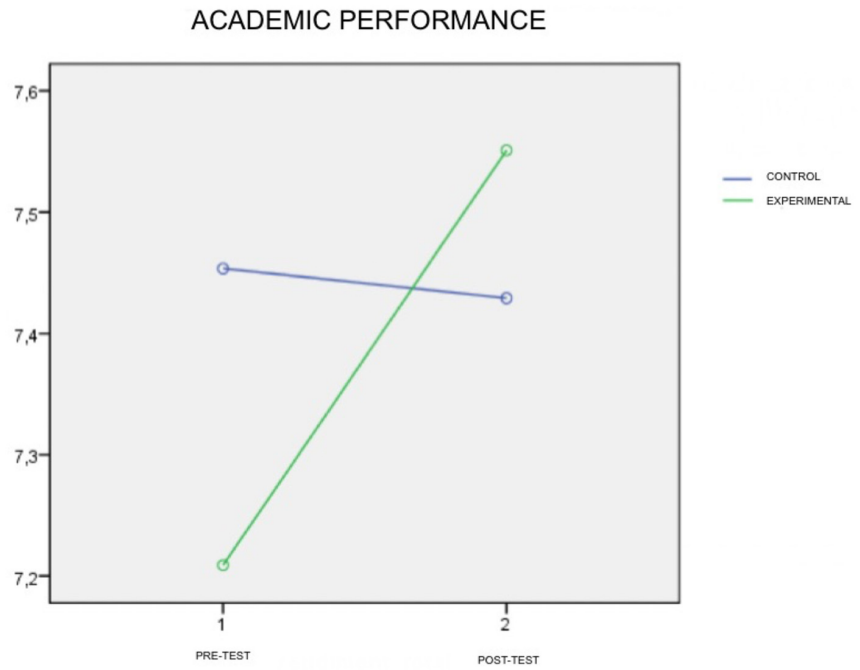


FIGURE 1 | Academic Performance scores of the experimental and control groups at the pretest (1) and posttest (2) times.

TABLE 2 | Summary of intra-inter subject univariate analysis of variance: Total EQ-i:YV(S).

	Font	Type III	*gl*	*F*	Significant	η^2 partial	Observed power
Total EQ-i:YV(S)	Intra	486.87	1	12.50	0.00	0.13	0.94
	Intra × Inter	184.89	1	4.75	0.03	0.05	0.57
	Error Intra	3270.79	84				
	Inter	37.59	1	0.59	0.44	0.00	0.12
	Error Inter	5314.87	84				

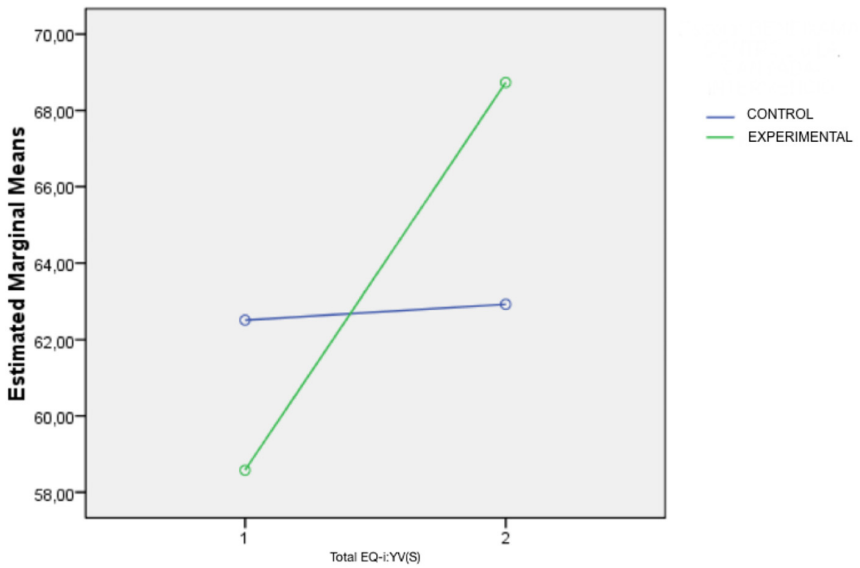


FIGURE 2 | Total EQ-i: YV(S) scores of the experimental and control groups at the pretest (1) and posttest (2) times.

The OKAPI program uses different strategies that have been associated with different studies to an improvement of academic performance such as the use of cooperative learning methodology (Johnson et al., 2014; Santos et al., 2018).

Cooperative learning offers a simple method to facilitate the daily SEL practice. When students work together they improve their social competence through the interaction with their peers. At the same time they perform positive personal relationships and live the experience of building their social emotional competence in real situations. We believe this supports the creation of warm and supportive ties which are linked to better school performance (Osher et al., 2018). Moreover, research points out that cooperative learning increases students' motivation and levels of self-esteem and self-efficacy (Johnson et al., 2014).

Several researchers have worked in programs with both social-emotional and school climate approaches which had a positive impact on school achievement (Cohen et al., 2009). The school and classroom environment has a role in shaping students' emotions and influences school performance (Divecha and Bracket, 2020). The circumstances in which students' learning take place matter (Suàrez et al., 2011). We agree with other studies that highlight the importance of developing a caring and supportive learning community in school (Issen, 2008). Strategies such as the Start Circle (the daily classroom meetings) or the different techniques used in the program to solve problems and manage their own emotions contribute to this objective.

Schools traditionally establish norms, lists of undesirable behaviors and consequences (Divecha and Bracket, 2020). To develop students' emotions we need to go beyond that and offer strategies to help them to transform those undesirable behaviors for others that will help the construction of a safe and positive environment.

This research analyses the impact of an emotional education program, that has been created with a positive classroom climate approach, on academic performance. Although it uses direct instruction to teach emotional competences, the main elements of the program are integrated during the routines of the school day. This aspect is essential to develop emotional competences and to facilitate the implementation of the program. The OKAPI program is aligned with research that highlights that emotional education programs should be integrated from a school-wide perspective and in collaboration with positive school climate initiatives (Jones et al., 2017).

Although the results of the program's effectiveness are positive, the effect size was relatively small. Further research is needed to determine the longer-term effect and it is also desirable to test it in larger samples. On the other hand, it would be interesting to use a standard test in order to obtain standardized date to the improvement in students' performance.

LIMITATIONS

This research has various limitations, including: the exclusive use of self-reporting as the method for collecting data (such methods are not very robust against possible bias in responses introduced by the subjects themselves), the impossibility of an accurate control over the variables in real situations in primary classrooms or the size of the sample used. Moreover, the participation of only two schools limits the extrapolation of results to other settings.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of Alicante Ethics Committee (Ref. UA-2015-07-06). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

M-JM-G carried out a theoretical review of the topic, did fieldwork, program implementation, and worked on quantitative methods. RG-C was director of the research and carried out the quantitative methods.

FUNDING

This research was supported by the Spanish Ministry of Economy and Competitiveness (EDU2015-64562-R).

REFERENCES

Arnaiz, P., and Linares, J. E. (coord) (2010). *Proyecto ACOOP. Proceso de Implementación del Aprendizaje Cooperativo en el aula de Educación Primaria. Proyecto de Innovación/Investigación Dentro del Convenio Universidad de Murcia y Consejería de Educación, Formación y Empleo*. Colombia:Convocatoria.

Bar-On, R., and Parker, J. D. (2000). *The Emotional Quotient Inventorv: Youth version (EQ-i:YV)*. North Tonawanda: Multi-Health Systems.

Bear, G. G., Yang, C., Mantz, L. S., and Harris, A. B. (2017). School-wide practices associated with school climate in elementary, middle, and high schools. *Teach. Teach. Educ.* 63, 372–383. doi: 10.1016/j.tate.2017.01.012

Bisquerra, R. (2000). *Educación Emocional y Bienestar*. Barcelona: Praxis.

Bisquerra, R. (2003). Educación emocional y competencias básicas para la vida. *Rev. Educ. Educ.* 21, 7–43.

Bisquerra, R. (2009). *Psicopedagogía de las Emociones*. Madrid: Síntesis.

Bisquerra, R., Pérez, J. C., and Garcia, E. (2015). *Inteligencia Emocional en Educación*. Madrid: Editorial Síntesis.

Blake, P., Piovesan, M., Montinari, N., Warneken, F., and Gino, F. (2015). Prosocial norms in the classroom: the role of self-regulation in following norms of giving. *J. Econ. Behav. Organ.* 115, 18–29. doi: 10.1016/j.jebo.2014.10.004

Bond, L. A., and Hauf, A. M. C. (2004). Taking stock and putting stock in primary prevention: characteristics of effective programs. *J. Prim. Prevent.* 24, 199–221. doi: 10.1023/B:JOPP.0000018051.90165.65

Brackett, M. A., Palomera, R., Mojsa-Kaja, J., Reyes, M. R., and Salovey, P. (2010). Emotion-regulation ability, burnout, and job satisfaction among

British secondary-school teachers. *Psychol. Sch.* 47, 406–417. doi: 10.1002/pits.20478

Brackett, M. A., Rivers, S., Shiffman, S., Lerner, N., and Salovey, P. (2006). Relating emotional abilities to social functioning: a comparison of self-report and performance measures of emotional intelligence. *J. Pers. Soc. Psychol.* 91, 780–795. doi: 10.1037/0022-3514.91.4.780

Brenner, E. M., and Salovey, P. (1997). "Emotional regulation during childhood: developmental, interpersonal, and individual considerations," in *Emotional Development and Emotional Intelligence: Educational Implications*, eds P. Salovey and D. J. Sluyter (New York, NY: Harper Collins), 168–195.

Bronfenbrenner, U., and Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: a biological model. *Psychol. Rev.* 101, 568–586. doi: 10.1037/0033-295X.101.4.568

Cohen, J., McCabe, E. M., Michelli, N. M., and Pickeral, T. (2009). School climate: research, policy, teacher education and practice. *Teach. Coll. Rec.* 111, 180–213.

Collaborative for Academic Social and Emotional Learning [CASEL] (2007). *Benefits of SEL: Meta-analysis[online]*. Chicago: (UIC) University of Illinois at Chicago.

Deleon, A. C. P., and Atkinson, A. C. (1991). Optimum experimental-design for discriminating between 2 rival models in the presence of prior information. *Biometrika* 78, 601–608. doi: 10.2307/2337029

Divecha, D., and Bracket, M. (2020). Rethinking school-based bulling prevention through the lens of social and emotional learning: a bioecological perspective. *Int. J. Bullying Prevent.* 2, 93–113. doi: 10.1007/s42380-019-00019-5

Díaz-Aguado, M. J. (2003). *Educación Intercultural y Aprendizaje Cooperativo*. Madrid: Pirámide.

Duby, C., Guyon, X., and Prum, B. (1977). The precision of different experimental designs for a random field. *Biometrika* 64, 59–66. doi: 10.1093/biomet/64.1.59

Durlak, J. A. (1997). *Successful Prevention Programs for Children and Adolescents*. New York, NY: Plenum.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., and Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Dev.* 82, 405–432. doi: 10.1111/j.1467-8624.2010.01564.x

Extremera, N., Fernández-Berrocal, P., and Durán, A. (2003). Inteligencia emocional y burnout en profesores. *Encuentros Psicol. Soc.* 1, 260–265.

Extremera, N., and Fernández-Berrocal, P. (2004). El papel de la inteligencia emocional en el alumnado: evidencias empíricas. *REDIE* 6,

Fernández-Berrocal, P., Alcaide, R., Extremera, N., and Pizarro, D. A. (2006). The role of emotional intelligence in anxiety and depression among adolescents. *Individ. Differ. Res.* 4, 16–27.

Fernández-Berrocal, P., Berrios-Martos, M. P., Extremera, N., and Augusto, J. M. (2012). Inteligencia emocional: 22 años de avances empíricos. *Psicol. Conduct.* 20, 5–13.

Freeman, J. (1973). Environment and High IQ, a Consideration of Fluid and Crystallized Intelligence. *Pers. Individ. Differ.* 4, 307–313. doi: 10.1016/0191-8869(83)90152-6

Gabriel, R. M., and Hopkins, K. D. (1974). Research Design and Analysis Clinic - Relative Merits of Manova, Repeated Measures Anova, and Univariate Anovas for Research Utilizing Multiple Criterion Measures. *J. Spec. Educ.* 8, 377–389. doi: 10.1177/002246697400800410

Ganju, J. (2004). Some unexamined aspects of analysis of covariance in pretest-posttest studies. *Biometrics* 60, 829–833. doi: 10.1111/j.0006-341X.2004.00235.x

Garcia Gallego, C. (2001). "Investiagación cuasiexperimental I: Diseños preexperimentales y diseños cuasiexperimentales con grupo de control no equivalente," in *Diseños de Investigación en Psicología*, eds D. S. Fontes de Gracia, C. G. Gallego, A. J. G. Trillo, M. C. R. Pérez-Llantada, and E. S. Sánchez (Madrid: UNED), 374–378.

Gillies, R. M., and Haynes, M. (2011). Increasing explanatory behaviour, problem-solving, and reasoning within classes using cooperative group work. *Instr. Sci.* 39, 349–366. doi: 10.1007/s11251-010-9130-9

Gil-Olarte, P., Guil, R., Mestre, J. M., and Núñez, I. (2005). "La Inteligencia Emocional como variable predictora del rendimiento académico," in *Psicología Social y Problemas Sociales. Psicología Ambiental, Comunitaria y Educación*, eds J. Romay and R. García (Madrid: Biblioteca Nueva), 351–357.

Gil-Olarte, P., Palomera, M. R., and Brackett, M. A. (2006). Relating emotional intelligence to social competence, and academic achievement among high school students. *Psicothema* 18, 118–123.

Greenberg, M., Domitrovich, C., Weissberg, R., and Durlak, J. (2017). Social and Emotional Learning as a Public Health Approach to Education. *Future Child* 27, 13–32. doi: 10.1353/foc.2017.0001

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., et al. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *Am. Psychol.* 58, 466–474. doi: 10.1037/0003-066X.58.6-7.466

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1999). *Análisis Multivariante*, Vol. 491. Madrid: Prentice Hall.

Issen, A. M. (2008). "Some ways in which positive affect influences decision making and problem solving," in *Handbook of Emotions*, 3rd Edn, eds L. Feldman, M. Lewis, and J. M. Haviland-Jones (New York, NY: Guilford Press), 548–573.

Johnson, D., Johnson, R., Roseth, C., and Seob, T. (2014). The relationship between motivation and achievement in interdependent situations. *J. Appl. Soc. Psychol.* 44, 622–633. doi: 10.1111/jasp.12280

Johnson, R. T., and Johnson, D. W. (2009). An educational psychology success story: social interdependece theory and cooperative learning. *J. Educ. Res.* 38, 365–379. doi: 10.3102/0013189X09339057

Johnson, R. T., Johnson, D. W., and Holubec, E. J. (1999). *El Aprendizaje Cooperativo en el Aula*. Buenos Aires: Paidós.

Jones, S. M., and Bouffard, S. (2012). Social and emotional learning in schools: from programs to strategies. *Soc. Policy Rep.* 23, 1–33. doi: 10.1002/j.2379-3988.2012.tb00073.x

Jones, S. M., Brush, K., Bailey, R., Brion-Meisels, G., McIntyre, J., Kahn, J., et al. (2017). *Navigating SEL from the Inside Out: Looking Inside & Across 25 Leading SEL Programs: A Practical Resource for Schools and OST Providers*. Cambridge, MA: Harvard University.

Kagan, S. (1994). *Cooperative Learning*. San Clemente, CA: Kagan Pub.

Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., et al. (1998). Statistical practices of educational researchers: an analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Rev. Educ. Res.* 68, 350–386. doi: 10.3102/00346543068003350

Larraz, N., Vazquez, S., and Liesa, M. (2017). Transversal Skills development through cooperative learning. *Train. Teach. Future* 25, 85–95. doi: 10.1108/oth-02-2016-0004

Lavori, P. (1990). Anova, manova, my black hen - comments on repeated measures. *Arch. Gen. Psychiatry* 47, 775–778. doi: 10.1001/archpsyc.1990.01810200083012

León, B., Felipe, E., Meno, S., and Iglesias, D. (2015). Habilidades Sociales en equipos de Aprendizaje Cooperativo en el contexto universitario. *Psicol. Conduct.* 23, 191–214.

Leon, S., Tsiatis, A. A., and Davidian, M. (2003). Semiparametric estimation of treatment effect in a pretest-posttest study. *Biometrics* 59, 1046–1055. doi: 10.1111/j.0006-341x.2003.00120.x

León, B., Gonzalo, M., and Polo, M. I. (2014). Aprendizaje cooperativo y acoso entre iguales. Infancia y Aprendizaje. *J. Study Educ. Dev.* 35, 23–35. doi: 10.1174/021037012798977494

Linares, J. E. (2004). *Programa de Desarrollo de Conductas de Cooperación en el Alumnado*. Murcia: Consejería de Educación y Cultura Región de Murcia.

Lópes, P. N., and Salovey, P. (2004). "Toward a broader education: social, emotional, and practical skills," in *Building School Success on Social and Emotional Learning*, eds J. E. Zins, R. P. Weissberg, M. C. Wang, and H. J. Walberg (New York, NY: Teachers College Press), 79–93.

Mestre, J., Guil, R., Lopes, P., Salovey, P., and Gil-Olarte, P. (2006). Emotional intelligence and social and academic adaptation to school. *Psicothema* 18, 112–117.

Nasir, M., and Masrur, R. (2010). An exploration of emotional intelligence of the students of IIUI in relation to gender, age and academic achievement. *Bull. Educ. Res.* 32, 37–51.

Osher, D., Cantor, P., Berg, J., Steyer, L., and Rose, T. (2018). Drivers of human development: how relationships and context shape learning and development. *Appl. Dev. Sci.* 24, 6–36. doi: 10.1080/10888691.2017.1398650

Osher, D., Dwyer, K., and Jackson, S. (2004). *Safe, Supportive and Successful Schools: Step by Step*. Longmont, CO: Sopris West Educational Services.

Osher, D., Kendziora, K., Spier, E., and Garibaldi, M. L. (2014). "School influences on child and youth development," in *Advances in Prevention Science Vol. 1: Defining Prevention Science*, eds D. Z. Sloboda and H. Petras (New York, NY: Springer), 151–170.

Ovejero, A. (1993). Aprendizaje cooperativo: una eficaz aportación de la psicología social a la escuela del siglo XXI. *Psicothema* 5, 373–391.

Palomera, R., and Brackett, M. A. (2006). Frequency of positive affect as a possible mediator between perceived emotional intelligence and life satisfaction. *Ansiedad Estrés* 12, 231–239.

Petrides, K. V., Frederickson, N., and Furnham, A. (2004). The role of trait emotional intelligence in academic performance and deviant behavior at school. *Pers. Individ. Differ.* 36, 277–293. doi: 10.1016/S0191-8869(03)00084-9

Pérez, N., and Castejón, J. L. (2007). La Inteligencia Emocional como predictor del rendimiento académico en estudiantes universitarios. *Ansiedad Estrés* 12, 121–131.

Pujolàs, P. (2001). *Atención a la Diversidad y Aprendizaje Cooperativo en la Educación Obligatoria*. Málaga: Aljibe.

Pujolàs, P. (2005). Grupos Cooperativos: el cómo, el porqué y el para qué del aprendizaje cooperativo. *Cuadernos Pedagogía* 345, 50–54.

Salovey, P., Stroud, L. R., Woolery, A., and Epel, E. S. (2002). Perceived emotional intelligence, stress reactivity, and symptom reports: further explorations using the trait meta-mood scale. *Psychol. Health* 17, 611–627. doi: 10.1080/08870440290025812

Santos, M. A., Ferraces, M. J., Godas, A., and Lorenzo, M. (2018). Do cooperative learning and family involvement improve variables linked to academic performance? *Psicothema* 30, 212–217. doi: 10.7334/psicothema2017.311

Slavin, R. E. (1995). *Cooperative Learning: Theory, Research, and Practice*, 2nd Edn. Boston, MA: Allyn & Bacon.

Stemmler, M. (1994). A nonparametrical evaluation of anova and manova designs using interaction structure-analysis. *Biometr. J.* 36, 911–925. doi: 10.1002/bimj.4710360803

Suàrez, N., Tuero, E., Bernardo, A., Fernández Alba, E., Cerezo, R., and González-Pienda, J. (2011). El fracaso escolar en Educación Secundaria: análisis del papel de la implicación familiar [School failure in secondary education: an analysis of the role of family involvement]. *Rev. Form. Prof. Invest. Educ.* 24, 49–64.

Vega, M. L., and Hederich, C. H. (2015). Impacto de un programa de aprendizaje cooperativo en el rendimiento académico en matemáticas y español en un grupo de estudiantes de 4° de primaria y su relación con el estilo cognitivo. *New Approach. Educ. Res.* 4, 90–97. doi: 10.7821/naer.2015.7.124

Vuchkov, I. N., and Solakov, E. B. (1980). The influence of experimental-design on robistness tot non-normality of the F-test in regression-Analysis. *Biometrika* 67, 489–492. doi: 10.1093/biomet/67.2.489

Weissberg, R., Durlak, J., Domitrovich, C., and Gullota, T. (2015). "Social and emotional learning: past, present, and future," in *Handbook of Social and Emotional Learning*, eds J. Durlak, C. Domitrovich, R. Weissberg, and T. Gullota (New York, NY: The Guilford Press), 3–19.

Wilson-Jones, L., and Caston, M. C. (2004). Cooperative learning on academic achievement in elementary African American males. *J. Instruct. Psychol.* 31:280.

Woodrow, L. (2014). "Anova, ancova and manova," in *Writing about Quantitative Research in Applied Linguistics*, ed. L. Woodrow (London: Palgrave Macmillan), 73–84. doi: 10.1057/9780230369955

Zins, J., Bloodworth, M., Weissberg, R., and Walberg, H. (2004). "The scientific base linking social and emotional learning to school success," in *Building Academic Succes on Social and Emotional Learning: What Does the Research Say?*, eds J. E. Zins, R. P. Weissberg, M. C. Wong, and H. J. Walberg (New York, NY: Teachers Collage Press), 3–22.

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.

Copyright © 2020 Mira-Galvañ and Gilar-Corbi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

CONCEPTUAL ANALYSIS
published: 03 June 2020
doi: 10.3389/fpsyg.2020.01037

Future Time Perspective and the Achievement of Optimal Best: Reflections, Conceptualizations, and Future Directions for Development

Huy P. Phan, Bing H. Ngu and Kelvin McQueen*

School of Education, University of New England, Armidale, NSW, Australia

Time is an interesting concept. For some cultural groups, time is an entity that exists only in the here and now, whereas for others it can be linear, emphasizing a person's past, present, and future. Many of us, while living in the "present moment," may also anticipate and project future goals, dreams, hopes, and ambitions. Indeed, from a positive point of view, future orientations are healthy and may direct one's focus, instill motivation and persistence, and mobilize the expenditure of effort. Existing research has provided empirical evidence to support the promotion and encouragement of a positive *future time orientation*. From an educational point of view, the study of time may be useful for calculating achievement, given that a student may use future time orientation to guide and direct his/her academic and/or non-academic future. One notable question for consideration, in this case, relates to the *importance of timespan* – that is, how far into the future should one project? There may be a significant difference between, say, a timespan that scopes a 6-month period as opposed to a timespan that scopes a 2-year period. By the same token, over the past few years we have delved into an interesting line of inquiry, namely, the *nature of optimal best* – for example, what facilitates and/or causes a person to achieve an optimal level of best practice in particular subject matter? Our theory of human optimization, consolidated and recently published in *Frontiers in Psychology*, provides an in-depth theoretical account of an underlying process, which we postulate could help explain the achievement of optimal best. Optimization, in this case, is intimately linked to a person's achievement of optimal best. We rationalize that within the context of academic learning, cognitive complexity of particular subject matter could serve as an important source of motivation in the anticipation and projection a student's extended future timespan. In this analysis, the extremely complex nature of a learning task or a suite of tasks may compel a student to consider a longer future timespan for successful completion. We also argue, in contrast, that the specific duration of a future timespan (for e.g., 6 months vs. 2 years) could play a significant role in the successful optimization of a student's state of cognitive functioning.

Keywords: time, future time orientation, extended timespan, optimal best, optimization, positive psychology

OPEN ACCESS

Edited by:
José Carlos Núñez,
Universidad de Oviedo Mieres, Spain

Reviewed by:
Emmanuel Manalo,
Kyoto University, Japan
Daryl Wayne Niedermoser,
University Psychiatric Clinic Basel,
Switzerland

****Correspondence:***
Huy P. Phan
hphan2@une.edu.au

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

Received: *29 August 2019*
Accepted: *24 April 2020*
Published: *03 June 2020*

Citation:
Phan HP, Ngu BH and McQueen K (2020) Future Time Perspective and the Achievement of Optimal Best: Reflections, Conceptualizations, and Future Directions for Development. Front. Psychol. 11:1037. doi: 10.3389/fpsyg.2020.01037

INTRODUCTION

Time is an interesting and mysterious concept. For Albert Einstein, space and time are merged inextricably into a four-dimensional space-time continuum. Buddhist philosophy holds the philosophical position that there is no past and there is no future; only the present moment. This belief means that we should not ponder past events, nor consider and plan for the future. Rather, in tandem with Eastern understandings of mindfulness and meditative practice (Nyanaponika, 1972; Master, 2010) everything is here and now. *We live in the present moment*. The authors appreciate this philosophical premise, but contend that our own personal beliefs, research development, and professional experiences differ from Buddhist philosophy and we acknowledge and recognize the importance for modern physics of time being linear. But in this study, the personal experience of time is the essence for consideration; like when people ask which nation will next win the next FIFA World Cup; when are we likely to experience another financial crisis; what will the next iPhone look like; and so on? When people pose questions like these, for example, then they are rejecting a viewpoint of life being lived in just the present moment.

Time perspective, commonly known as TP, is an interesting theoretical concept that has received considerable research interest over the years (e.g., Zimbardo and Boyd, 2008; Milfont et al., 2012; Janeiro et al., 2017). Consideration of time perspective, from our synthesis and review of the literature, can explain in part the relevance and personal significance of life experiences. Time perspective, in this case, considers time a non-singular entity (Zimbardo and Boyd, 1999) through which a person self-reflects upon his/her past experiences to shape their present moment, which in turn informs his/her future actions. The future, in this sense, indicates different possible outcomes, both positive and negative, depending on personal estimations of past and present life experiences. Flowing from this, the authors acknowledge that time, in terms of its continuity, has relevance, significance, and applicability in relation to the achievement of both educational and non-educational outcomes. A student's previous negative experiences in mathematics, for example, may give rise to his/her present state of anxiety when learning this academic subject, which would also be likely to deter him/her from choosing a mathematics-related career in future. In a similar vein, an adolescent's positive feelings when working with senior citizens may motivate him/her in future to enroll in a social work degree.

With reference to the study of TP, *future time orientation* is an interesting timepoint for analysis and examination. As individuals, both in academic and non-academic spheres, we are always curious and interested to know what the future holds. Will we achieve exceptional grades; will we have enough savings for retirement? In the context of schooling, in the case under consideration here and according to several commentators, future time orientation plays a meaningful role in motivating students to seek new frontiers and to strive for successful accomplishments (Lens et al., 2002; Simons et al., 2004a,b). With this in mind, one interesting element of enquiry is the *specific extended timespan into the future* that would be considered optimal. What is the most adequate timespan into the future that would yield productive and enriching outcomes for a person? There are individual variations in perceptions of future time, for example: a 2-week timespan into the future for a 4-year-old child is quite lengthy as they anticipate their birthday present (e.g., "What will I get for my birthday?"), as opposed to a 5-year timespan into the future for a 17-year-old teenager who wishes to become a medical doctor. From the motivational literature, anticipating and setting forth a specific timespan into the future is a valuable consideration. How briefly or deeply into the future we anticipate appears to influence our internal motivational state. It becomes motivational and proactive for a person to anticipate and set forth an extended timespan into the future for attaining positive yields.

One focus of enquiry to consider is the establishment of different pathways and means for encouraging the contemplation and projection of specific timespans into the future. At the same time, of course, there is the need of directing a person's focus toward actually seeking achievement of future ambitions. In considering both these needs, a possible line for development is motivating people to achieve *optimal best* (Martin, 2006; Liem et al., 2012; Phan and Ngu, 2019a). Optimal best, that is, the attempted maximization of a person's state of functioning, may enhance the active processes of human agency. We contend that striving to achieve optimal best in school subjects, for instance, could assist a person to project a specific and definitive timespan into the future. In a similar vein yet conversely, a developing a specific timespan for future achievement could help galvanize a person's focus and motivation to achieve optimal best.

We acknowledge that to date very little, if any, focus in the relevant literature has been given to considering the relationship between TP and the achievement of optimal best. Do the two theoretical concepts relate to each other and/or to what extent do these two theoretical concepts explain proactive human behaviors? This line of inquiry, theoretically and conceptually, is innovative for its positive nature, reflecting our recent research into the study of optimal best (e.g., Phan and Ngu, 2017a; Phan et al., 2019a,b,c). As a working hypothesis, we could argue that optimal best is necessarily reliant on a point of reference structured into the future. In other words, unlike a person's historical optimal best, which is past, optimal best as a positive outcome requires time for achievement, i.e., achievement in the future. In that case and as one consideration, how much time is needed to achieve optimal best may depend on the cognitive complexity of the subject matter. Furthermore, having an extended timespan into the future is, perhaps, a necessity for the achievement of optimal best. Overall then, we contend that a focus on TP within the context of optimal best is significant, contributing to advance the research into the optimization of optimal best.

THE IMPORTANCE OF TIME

Time, as we briefly described, is an interesting concept that has implications for society and individuals. Why does a local government need to plan ahead in terms of policy development? Why is a secondary school student being asked to reflect on his/her previous academic performance in mathematics? Why

does a footballer have to focus on her/his current state of physical functioning? These questions are significant for this study since they place emphasis on a linear trajectory and, more importantly, suggest potential interrelations between past, present, and future events. Some individuals, of course, may simply fixate upon past situations and experiences, whereas others may choose to ponder their future. This diversity is interesting as it acknowledges that to a large extent our lives are intricately linked to the nature of time. Erikson's (1968) *psychosocial theory of personality* with its eight stages, for example, places emphasis on a linear progression from one level to the next. For Erikson (1968) completion and success at a particular stage form the basis for continuing onto the next stage. This theory emphasizes the importance of time as a linear trajectory, as a pathway and pattern of development, where past experiences contribute to influencing the present moment and, likewise, the future.

It is interesting, however, that other viewpoints do not necessarily concur with the notion of the continuation of time. Let us delve into this positioning a bit. Our research development in *mindfulness* from an Eastern perspective (Hanh, 1976; Yeshe and Rinpoche, 1976; Goldstein and Kornfield, 1987; Loden, 1996; Kabat-Zinn, 2015) has offered a different viewpoint about the concept of time. That viewpoint is that time is a non-linear entity. Time is not thought of as continuous and there is no associated discourse of temporal linearity. Mindfulness, according to Taiwanese colleagues known to the authors who are practicing Buddhist nuns and monks, is concerned solely with the *present moment* – everything is here and now. Buddhist scholars, for example, believe that the past does not matter and, likewise, we do not plan for the future since we do not know what it holds. Life in the present moment is straightforward and has few complications. Thus, the personal experience of the present moment is said to reflect a state of *mindfulness*; a state of clearness, calmness, and enlightenment. This philosophical position, derived from Buddhist sutras (Yeshe and Rinpoche, 1976), places emphasis on concentrating on the essence of the context as the "here and now."

Indeed, we acknowledge that "situating" life in the present moment appears to be "simple" and straightforward. Why worry and stress about the future when one does not know what this future may hold? Similarly, why reflect upon and ponder the past when it is impossible to amend events, situations, and/or experiences? Buddhist teaching, in this sense, is related to the notion of a person living and leading a simple, healthy, and happy life. We acknowledge, however, that this viewpoint is largely incompatible with Western theorizations and understandings of temporal linearity. For example, Nuttin's (1964) article, entitled *"The future time perspective in human motivation and learning,"* provides an interesting reference to time: "A simple analysis of human behaviors calls attention to the fact that man [and woman], in his [/her] dealing with a given situation, is usually directed toward something which is not yet there, something which is still to come, something different, even something new. . ., are all oriented toward something ahead, something that they are looking for: their behavior is 'future bound'. . .." (p. 60). This conceptualizing of time emphasizes the importance of a linear trajectory that connects a person's past experiences, that is, his/her current state of functioning, into considerations about the future. Moreover, Nuttin (1964) proposes a mysterious and unknown nature to future specific timepoints.

We do not discount the possibility that everyone could, if so inclined, perceive time as something "singular" (i.e., limiting itself to the present moment). The benefit of such a viewpoint is its straightforwardness – that is, it encourages a person to free himself/herself from the complexities and perplexities of life and, from this, perhaps to live "a happy, healthy life in the present moment." Having said this, however, we acknowledge that in the modern world with its particular type of economic development and perpetuation of certain types of social arrangements, and the associated expectations of personal growth, require in many circumstances the rejection of this viewpoint and the acceptance, in contrast, of a viewpoint that time is continuous. The associated psychological state is that people expect to reflect, envisage, and anticipate different timepoints in their lifetimes. The study of the psychology of time by different researchers (e.g., Lewin, 1942; Nuttin, 1964; Mehta et al., 1972; Zimbardo and Boyd, 1999, 2008) is insightful as this line of inquiry makes a concerted effort to explain the differing patterns in cognition, motivation, and behavior of individuals in both educational and non-educational contexts. Mehta et al.'s (1972) writing provides a brief historical summary of other writers' understanding of time (e.g., Israeli, 1934; Frank, 1939; Lewin, 1942; LeShan, 1952; Fraisse, 1963; Nuttin, 1964). So, to propose a technical definition of what is time perspective, commonly known as TP, we can turn to Zimbardo et al. (1997) who define TP as "the manner in which individuals, and cultures, partition the flow of human experience into the distinct temporal categories of past, present, and future. The boundaries, extension, salience, and utilization of any of these categories may vary considerably as a function of learned preferences that become stabilized into a functional cognitive style, and also as a consequence of situational, structural, and task demands" (Zimbardo et al., 1997). Extending this, TP perhaps takes on its most important aspect, from a layperson's point of view, as the ability to anticipate future situations and events by reflecting on his/her past experiences (Lennings et al., 1998). This common-sense understanding proposes that a person's life experience and growth are "sequenced" in a linear trajectory. A student's past experience of racism at school, for example, is likely to shape his/her current thinking and behavior, which may then link with deliberations about future actions (e.g., avoid attending school). In contrast, likewise, a student's current enjoyment of Industrial Arts (e.g., woodwork) may determine his/her choice of a career choice later on.

The Nature of Future Time Perspective (FTP)

One notable timepoint that is of interest for discussion is the unknown future. Future time perspective, commonly known as FTP, is defined as "the timing and ordering of personalized future events" (Wallace, 1956). Nevertheless, there are differing definitions of FTP. For example: "FTP is the degree to which and the way in which the chronological future is integrated into the present life-space of an individual through motivational

goal-setting processes" (Husman and Lens, 1999); or "the totality of the individual's views of his psychological future and psychological past existing at a given time" (Lewin, 1951); or "a general concern for future events" (Kastenbaum, 1961); or "a general capacity to anticipate, shed light on and structure the future" (Gjesme, 1983); or "the present anticipation of future goals" (Simons et al., 2004b). These differing definitions, according to Seijts (1998), emphasize the complex nature of FTP. A point of commonality in these definitions is that FTP is related to and espouses the importance of a person's current envisaging of future events, situations, tasks, etc., in his/her present timespace – for example, while writing up this manuscript, we envisage and anticipate its completion in a few weeks from now.

Seijts (1998) definitional overview is interesting for its description of five major facets of FTP: (i) *extension*, which is concerned with the length of the future timespan that a person conceptualizes; (ii) *coherence* is the degree of organization of envisaged events in the future timespan; (iii) *density* is defined as the number of events that are expected in a person's future – that is, his/her goals, hopes, fears, and wishes; (iv) *directionality*, which is the extent to which a person perceives himself/herself to be moving forward from the present moment into the future; and (v) *affectivity* relates to the extent to which a person feels gratified by anticipated events. These five facets, in totality, play a central role in the operational functioning of FTP. For example, as existing research has shown, the extension of a person's FTP is closely related to his/her cognition and motivational patterns (Simons et al., 2004a,b). Indeed, developments in research over the past seven decades have produced clear and consistent evidence highlighting the significant impact of FTP on different types of adaptive outcomes (e.g., Zimbardo et al., 1997; de Bilde et al., 2011; Lens et al., 2012; Phan, 2015; Husman et al., 2016; Janeiro et al., 2017).

From a motivational perspective, FTP may operate to encourage a person to be purposive and self-regulated and to flourish in the course of academic learning and schooling (Zimbardo and Boyd, 1999; Barber et al., 2009; Janeiro et al., 2017). It is important, in this analysis, for a person to consider a specific future timespan in terms of envisaging different types of endeavors for accomplishment. This notion of varying extensions of future timespan may serve as a significant source of motivation (Simons et al., 2004a). In general, we may have either short or long (or deep) extended timespans into the future (Simons et al., 2004b). This variation of FTP (e.g., say, 4 weeks into the future vs. 2 years into the future), as the literature has shown, is intricately linked to a person's state of motivation, cognition, commitment, and behavior. Academically, for example, two fourth-year undergraduate students may aspire toward obtaining Ph.D. degrees in 5 years' time. One student, Chou, has a long FTP, whereas David, his best friend, has a short FTP. Consequently, the psychological distance toward this future goal of obtaining a Ph.D. is experienced differently by both students – Chou, in this case, is likely to perceive having a Ph.D. as being closer in time than David because of the latter's short FTP. For David, the same future goal may not be part of his life space. Obviously, there is a negative correlation between an individual's FTP extension and his/her perceived psychological distance toward the self-set future goal (in this case, obtaining a Ph.D. degree) (Moreas and Lens, 1991). What is of interest also, however, is that the length of a person's FTP does not affect his/her perceived psychological distance from the goal when this goal is set in the very near or very far future. In this sense, tomorrow or Friday is very near for all of us regardless of the various lengths of our FTPs, whereas 20 or 25 years from now may be chronologically too far away to matter.

Simons et al.'s (2004b) overview of FTP highlights an interesting theoretical aspect from De Volder and Lens' (1982) research – namely, the distinction between the *cognitive component* and the *dynamic component* in FTP. Consider the aforementioned example of Chou and his best friend, David. Chou, with his extended FTP, is likely to perceive his present behavior as being more instrumental as this would help him achieve a broad range of both immediate and future goals. This is the cognitive component. At the same time, Chou also values his present task-engagement more strongly because the anticipated value of the future goal of obtaining a Ph.D. is higher. This is the dynamic component. David's case, by contrast, may be somewhat different. Because of his relatively short FTP, David perhaps does not anticipate more-distant future goals and, consequently, his present actions are perceived as less instrumental and/or containing less utility.

In sum, from the forgoing discussion, extended timespans into the future play a significant role in helping to determine and explain individuals' motivational beliefs and patterns in learning, cognition, and behavior. In this analysis, the study of FTP is not simply limited to the notion of "setting forth future goals for accomplishment" This line of reasoning, we contend, is too narrow and does not reflect the complex nature of a future timespan. A timespan into the future, as we described it, may be relatively short and indicate a simple goal or focus for consideration – for example, what will a person get for her birthday next week? We would contend that it is instead of more value, both in academic and non-academic pursuits, for individuals to consider extended timespans. Individuals with extended FTPs, the literature strongly suggests and we concur, are more likely to be motivated to engage closely with their learning (De Volder and Lens, 1982). These individuals (the example of Chou) perceive their present behavior as being more instrumental in achieving immediate and future goals. Likewise, individuals with extended FTPs value their present task-engagement more strongly because the anticipated value of the future goal is higher (Simons et al., 2004b). The pervasive question of course, arising from this analysis, is the appropriate and optimal extension of FTP.

We contend that it is noteworthy for educators and researchers to develop and explore pathways, means, and/or opportunities that could encourage and foster a healthy extension of FTP. This line of inquiry reflects the importance and benefits of a purposive extended timespan into the future. We do not consider there to be much value in having a short future timespan (e.g., preparing for the class quiz scheduled for next Friday) as this "briefness" would not have any meaningful impact on an individual's growth. From the perspective of formal education, we believe that it is an important feat for a student to contemplate the different types of positive outcomes that could arise from him/her having an extended FTP. As noted previously, it is

probably more significant to focus on and think about university in 3 years' time than to think about the final exam at the end of this year. Contemplating something that is "deep" into the future is, we believe in concordance with the literature, motivational and may operate to guide and direct a person to plan and strive for successive accomplishments. Career advice and university open days in Australia, for example, play a central role in helping secondary school students think about their futures post-secondary school. Aside from this opportunity, what else is available to help students envisage an extended timespans into the future? Of course, interlocutors could have students write down and discuss both their short-term and long-term future plans. It is also plausible to use vicarious information (Bandura, 1986, 1997), such as observation and role modeling, to facilitate in-depth understanding and appreciation of extended FTPs. Watching and/or observing a credible model who is timely in decision making and who is successful with his/her future endeavors (an ideal example may be someone saying to students the following: "I was successful by the age of 24 to achieve [. . .]. I planned this when I was 18. . .") may, in this analysis, provide vicarious information that the individual, too, may succeed with his/her future.

Our research interest, in this case, relates to the development of a conceptualization that could inform and facilitate the extension of an appropriate timespan into the future. Specifically, our focus of inquiry makes attempts to consider the theory of optimization (Phan and Ngu, 2017a, 2019b; Phan et al., 2017, 2019c) as a means by which one could extend his/her FTP. This consideration, we contend, is significant as it places emphasis on the tenet of optimal achievement best (Phan and Ngu, 2017a; Phan et al., 2017, 2019c) and relating this optimal best to a person's extended FTP. One particular aspect that we rationalize and argue for is that *a deep timespan into the future is more advantageous than a short timespan*. We incorporate theoretical understandings of optimal achievement best to support this proposition – that is, for example, the cognitive complexity of optimal achievement best would closely associate with a person's specific timespan into the future. An extremely complex level of optimal achievement best would extend a person's timespan (say, 6 months), whereas a low complex level would shorten his/her timespan (say, 2 weeks).

BIDIRECTIONAL ASSOCIATIONS BETWEEN FTP AND OPTIMAL BEST

We rationalize the potential bi-directional interrelationship between FTP and the nature of optimal achievement best. This conceptualization, which we explore in detail in this section and subsequent sections, is innovative as it draws attention to two lines of research: (i) the stipulation of an extension in FTP in order to achieve optimal best; and (ii) the complexity of optimal best, which, we have asserted, closely aligns with an extension in FTP. Our conceptualization, in general, proposes the promotion and fostering of an extended FTP and the setting of an optimal best that has high cognitive complexity.

The Framework of Achievement Bests

Our first stipulation is that an extended FTP could coincide with and operate within the *process of optimization* in order to facilitate the achievement of optimal best. Optimization, extensively mentioned in the literature (e.g., Freund and Baltes, 1998; Fraillon, 2004; Ziegelmann and Lippke, 2007; Eguizábal et al., 2018), is an underlying process that serves to optimize an individual's state of functioning (Phan and Ngu, 2017a). Drawing from Fraillon's (2004) and Phan et al.'s (2017) seminal papers, we recently provided a comprehensive discussion of the *Framework of Achievement Bests*, which is a theoretical model that makes attempts to explain the process of optimization (Phan et al., 2019c). Optimization is like a "vehicle" that optimizes an individual's state of functioning from one timepoint (e.g., T_1) to that of another timepoint (e.g., T_2). The Framework of Achievement Bests emphasizes two levels of best practice: (i) *realistic best*, or *actual best* (Fraillon, 2004), concerned with "a person's current, realistic level of cognitive ability" (Phan et al., 2016) – for example, what is it that I am capable of at present in mathematics?; and (ii) *optimal best,* or *notional best* (Fraillon, 2004), concerned with "the maximum of a person's cognitive capability" (Phan et al., 2016) – that is, what is my optimal best in mathematics? A notable inquiry arising from this focus is this: what *causes* and/or *facilitates* a person to progress from realistic best, L_1, to optimal best, L_2? Specifically, the totality of the process of optimization involves the active operation of three major pathways (see **Figure 1**):

i. *Pathway A*: the activation and enactment of different types of agents: educational (e.g., an instructional design for effective learning: Ngu et al., 2014), psychological (e.g., personal belief in efficacy: Bandura, 1997), and psychosocial (e.g., the importance of teacher-student social relationship: Roorda et al., 2011), which then act as sources of "energy."
ii. *Pathway B*: the impact of energy on the stimulation of buoyancy of different types of psychological attributes – for example, intrinsic motivation, personal resolve, effective functioning, mental strength, and effort expenditure.
iii. *Pathway C*: the buoyancy of psychological attributes (e.g., mental strength) that may then arouse and sustain a person's state of functioning and its improvement from T_1 to T_2.

According to Phan et al.'s (2019c) theorization, the combination of Pathway A, Pathway B, and Pathway C constitutes an "optimizing effect," which the authors term "γ." Aside from this positing of γ, the authors also proposed as an "Index of Optimization" (IO) where this equates with the following: $\Delta_{(L2-L1)} \times \gamma$, where $\Delta_{(L2-L1)}$ = quantitative difference between L_1 and L_2. Phan et al.'s (2019c) theoretical model of optimization, capitalizing on previous research developments (Fraillon, 2004; Phan et al., 2016, 2017, 2019a,c), is innovative for proposing a *quantitative nature* – that is, the possibility of measuring and quantifying the process of optimization. Furthermore, this consideration emphasizes the importance of the *magnitude of optimization* – that is, how "much" optimization is needed to facilitate the achievement of L_2

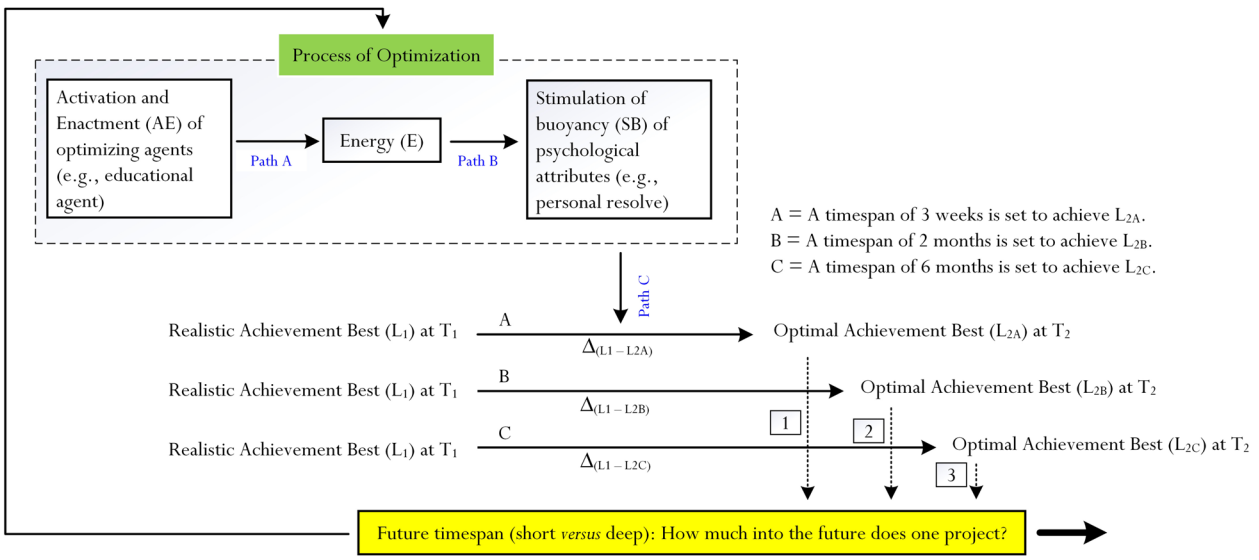


FIGURE 1 | Conceptual model for consideration. Note: this conceptualization is derived from Phan et al.'s (2019c) theoretical model of optimization. According to this theoretical model of optimization, the activation and enactment of different types of optimizing agents (e.g., educational agent being an appropriate instructional design) would create a perceived sense of positive energy (E), which would then stimulate the buoyancy of different psychological attributes (e.g., personal resolve). This process of optimization, in turn, would differ in terms of its magnitude (i.e., its strength: Phan and Ngu, 2019b), consequently as a result of the level of cognitive complexity of optimal best (e.g., $\Delta_{(L1\ -L2A)}$ is not as difficult, cognitively, as $\Delta_{(L1\ -L2B)}$, etc.). In this conceptualization, however, we propose that aside from optimization, a person's future timespan (i.e., short vs. deep) would correspond to the cognitive level of optimal best – denoted as path "1," path "2," path "3," etc. For example, the achievement of L_{2A} at T_2, which is a relatively simple modeling, would perhaps be associated with a short future timespan.

from L_1? This notion of magnitude, or strength, of optimization is interesting from our viewpoint since it places a focus on three major facets: (i) an individual's state of motivation, personal experience, and level of L_1; (ii) the quantitative and qualitative complexity of L_2; and (iii) the difference or range between L_1 and L_2. Consider this example of L_1 and L_{2A} where a secondary-school student is learning equation solving (Ngu and Phan, 2016):

- L_1: the student knows how to solve one-step linear equations, for example: $x + 10 = -4$, solve for x.
- L_{2A}: the student believes that he/she is able to solve multi-step linear equations, for example: $4\ (x + 5) = 3\ (x - 7)$, solve for x.
- L_{2B}: the student believes that he/she is able to solve quadratic equations, for example: $4\ (x + 8)^2 = 6$, solve for x.

As stated from the example, L_{2B} is more complex than L_{2A}, which in turn, is more complex than L_1. Hence, referring to our previous contention, it would require a greater magnitude of optimization to help in the achievement of L_{2B} from L_1 (i.e., $L_1 \rightarrow L_{2A}$) than of L_{2A} from L_1 (i.e., $L_1 \rightarrow L_{2B}$). As an example in school contexts, the magnitude of optimization (e.g., how much optimization is needed for me to understand something?) for understanding instructional material may involve a student's reliance on and/or use of different instructional designs and/or pedagogical practices (Ngu and Yeung, 2012; Ngu et al., 2014). Comparative pedagogical practices (e.g., the use multimedia information such as a YouTube video vs. an in-depth text), in this case, may inform a student's L_1 to L_2 differently.

In a similar vein, educationally, a native speaker of English may find it difficult to study and learn Chinese Mandarin as a foreign language. Hence, similar to the case of L_{2B}, a student learning Chinese Mandarin as a foreign language would require significant optimization, especially if he/she does not have prior experience of Chinese, intrinsic motivation, etc. With reference to FTP, in this analysis, the student may be counseled to envisage a long future timespan (e.g., 2 years to provide adequate time) for successful accomplishment. A short future timespan (e.g., 2 months), by contrast, we think would not provide an adequate timeframe for the student to achieve optimal best in learning Chinese. By the same token, in terms of educational resources for the purpose of optimization, we would expect to find the student relying on multimedia presentations, personal one-on-one scaffolding, etc. Hence, as a point for consideration, let us consider two different scenarios.

- L_1: the student has elementary knowledge of Chinese Mandarin, which enables him/her to have a basic conversation with another person (Note: in this case, for the sake of argument, assume that the student can read and write about 50 Chinese characters).
- L_{2A}: the student believes that he/she is able to read and write at least 150 Chinese characters. At the same time, L_{2A} may also stipulate the combination of different Chinese characters to have new meanings.
- L_{2B}: the student believes that he/she is able to read and write at least 300 Chinese characters, with some of these being sophisticated in nature (e.g., the character of 孝, xiào, which

means filial piety). L_{2B} may also stipulate the combination of different Chinese characters to have new meanings.

Similar to our previous example of mathematics, progressing from L_1 to L_{2A} is easier than achieving L_{2B} from L_1 for Chinese Mandarin learning. Knowing how to read and write 300 Chinese characters, some of which are sophisticated in nature, would require much more time (e.g., a future timespan of 2 years for accomplishment) and a sustained effort at optimization (e.g., the reliance on many multimedia sources).

Optimization, in its totality, is positive in nature and reflects the paradigms of *positive psychology* (Seligman and Csíkszentmihályi, 2000; Seligman et al., 2009) and *motivation* (Franken, 2007). The enactment and/or the operational functioning of optimization is postulated to result in and to account for the achievement of optimal best, whereas inactive optimization is likely to give way to ineffective learning and sub-optimal achievements. Optimization, in this case, is more than just testament of an association between two variables (Phan et al., 2019c). A positive correlation (r), in this case, is somewhat limited and does not explain and/or reflect the full scope of the process of optimization. The proposition of an "optimizing effect," denoted as γ, as Phan et al. (2019c) propose, is interesting and may depict and explain the complex nature of achievement of optimal best. Empirical research into the underlying mechanisms of optimization is somewhat limited to date with the exception of a few correlational inquiries that we have undertaken (e.g., Phan et al., 2018, 2019a,b, 2020; Phan and Ngu, 2019b). One possible reason for this limitation, we contend, is related to the "appropriateness" of a specific methodology that could assist the accurate measurement and assessment of optimization – we refer to this as "methodological appropriateness."

An Extended FTP and the Achievement of Optimal Best (FTP → OB)

How does an extended FTP assist in the achievement of an optimal level of best practice? In this analysis, we rationalize that a deep or long timespan into the future may act in tandem with, and/or form part of, the process of optimization (Fraillon, 2004; Phan and Ngu, 2017a; Phan et al., 2017, 2019c). Our conceptualization, illustrated in **Figure 1**, shows a person's FTP acting as a psychological agent of change. This conceptualization, developed from our recent revision (Phan et al., 2019c) of the theory of optimization (Phan and Ngu, 2017a; Phan et al., 2017), shows the potential impact of FTP on the process of optimization.

As shown in **Figure 1**, we propose that a deep future timespan would correspond to the achievement of a complex level of optimal best – denoted in **Figure 1**, in this case, as path "1," path "2," and path "3." With reference to this proposition, let us consider three different possibilities: (i) the achievement of L_{2A} from L_1; (ii) the achievement of L_{2B} from L_1; and (iii) the achievement of L_{2C} from L_1. As depicted visually, the difference between L_1 and L_{2C} ($L_1 - L_{2C}$) is "larger" than the difference between L_1 and L_{2B} ($L_1 - L_{2B}$) and this difference, in turn, is larger than the difference between L_1 and L_{2A} ($L_1 - L_{2A}$). Referring back to our previous discussion, the complexity of L_{2C} is greater than that of L_{2B} and L_{2A} (and the complexity of L_{2B} is greater than L_{2A}). The achievement of L_{2C} therefore would require "more" time and optimization (i.e., the magnitude of optimization is relatively high). Our rationalization posits that envisaging and planning a deep future timespan would facilitate and encourage the striving of a more complex level of optimal achievement best. For example, a Year 11 student envisaging a timespan into the future about university life would, perhaps, strive to achieve more difficult and ambitious endeavors at present (e.g., obtaining consistent A and A^+ grades for different subjects). Envisaging a shorter future timespan, from our point of view, would negate a student's motivation and their desire to seek out complex endeavors for accomplishment (e.g., personal contentment with C and C^+ grades for different subjects). As such, a relatively non-complex cognitive level of optimal best would require a lesser period of time for accomplishment.

As Simons et al. (2004b) explain, individuals with long or deep FTPs set goals that are situated in the distant future, whereas individuals with short FTPs set most of their goals in the near future (p. 122). Moreover, we would contend that an extended FTP is indicative of personal maturity, thoughtful deliberation, and well-measured ambitions. From this understanding, an investment in a long or deep FTP is noteworthy and in terms of academic pursuits, for example, a deep FTP could help students anticipate the following: (i) consider spending a longer period of time seeking assistance by consulting with others and/or utilizing different resources (e.g., going to the library); (ii) consider spending a longer period of time on personal reflection; and (iii) plan, organize, and develop a myriad of objectives and goals to assist in achieving a complex optimal best. A purposively extended FTP is more likely to direct a student's attention, cognition, and motivation toward achieving specific complex optimal bests in life. Having a deep FTP is advantageous as this extended timespan (e.g., I need to spend the next 2 months studying this) would act as a source of energy, guiding the student's behavior, personal resolution, and personal belief to recognize that anticipating long futures is beneficial and not wasteful.

We argue that conceptualization of an extended future timespan is not wasteful but may enable a student to consult others, engage in different types of cognitive strategies, and work on problems that could help improve his/her optimal level of best practice. As Lens et al.'s (2012) writing suggests, many of us have long FTPs whereas others, in contrast, may have relatively short FTPs. Those with short FTPs envisage and set most of their goals in the near future. They do not take into consideration what will come later on in life. Coupled with previous descriptions (Simons et al., 2004b; Zimbardo and Boyd, 2008; Lens et al., 2012), we contend that a short FTP is ineffective as this would negate the motivation, personal belief, and achievement of complex optimal best. In other words, a short FTP is intrinsically linked with a modest level of cognitive complexity. What this means then, of course, is that we do not expect a person with a short future timespan to achieve successfully a high level of cognitive complexity – for example, a relatively short future timespan of 2–3 weeks is unlikely, in this case, to be sufficient for preparing a student to plan, organize, and accomplish a

complex level of optimal best. By contrast, however, a deep future timespan is more meaningful and advantageous as this would provide sufficient time (i.e., duration) for a person to reflect, contemplate, plan, seek help, etc (Zimbardo and Boyd, 2008; Husman et al., 2016).

Moreover, we contend that an extended future timespan is important as this duration provides adequate time to help facilitate and enact the process of optimization. In this sense, enactment of optimization is not instantaneous, but rather requires some timeframe for development, depending on the level of cognitive complexity of a given task. Something that is easy, for instance, would not require too much optimization – and hence, a shorter period of time in this case would be needed. Aside from this understanding, we also theorize that future timespan could fundamentally relate to a person's *perceived sense of subjective task value* for learning (e.g., how important is this task?), and his/her *expectation to succeed* at the given task. In this analysis, with reference to this postulation, there are two considerations:

i. Positive task values (e.g., a student perceives that learning algebra is important for her/his future mathematics-related career plan), acting as a psychological agent, would motivate a person to persist and, likewise, to expend more effort into his/her studying and learning experiences (Eccles et al., 1983; Wigfield, 1994; Wigfield and Eccles, 2000). Valuing a particular task for significant personal reasons, we suggest, necessarily calls for an investment of time, planning, effort, resources, etc. In contrast, of course, a perceived minimal value of a learning task would not act as an optimizing agent of change.
ii. Expectation of achieving a complex optimal best would, likewise, act as a psychological agent in the process of optimization. A high level of expectation to succeed, which, in this case, a person is confident of achieving would energize and motivate the expenditure of time, effort, personal resolve, etc. A low level of expectation to achieve complex optimal best, by contrast, would deter the process of optimization.

Moreover, a perceived sense of subjective value and high expectation to succeed in the achievement of optimal best would, from our point of view, necessary require consideration of an appropriate future timespan. In this analysis, we would expect to find an association between a deep future timespan and a high level of subjective value and expectation – for example, a student who values Chinese Mandarin (i.e., perceives its importance for his/her future plans) and expects to succeed is likely, in this case, to recognize that a prolonged future timeframe would be needed to accomplish the task. A low level of expectation and the perception of low subjective value, by contrast, would not necessarily equate to a deep future timespan – in other words, a low level of expectation and low subjective value would instead correspond with a short future timespan. Time would be considered a minor factor as there would be a perception of its minimal value, reinforced by a low expectation of success and, hence, from the student's point of view, envisaging a deep future timespan would be non-logical to the point of irrelevancy. Having said this, however, we also acknowledge the importance of cost-benefit factors (e.g., time spent vs. intensity – or anticipations of either or both) with reference to a person's perceived subjective interests and well-being. Does envisaging a deep future timespan for the sake of achieving optimal best seem justified in all cases, especially when one considers the potential cost involved (e.g., the amount of effort, personal resolve, etc.)?

Complexity of Optimal Best and Its Impact on an Extended FTP (OB → FTP)

How does optimal best in a subject matter assist in the development of a deep FTP? Unlike our first stipulation, we rationalize that in this case, a person's striving to achieve a complex optimal level of best practice would construct a deep timespan into the future. As existing research inquiries have noted, we all differ in our FTPs – some of us have long and deep FTPs, while others have short FTPs. Moreover, those of us who have long FTPs tend to set most of our future goals in the distant future, whereas those with short FTPs set future goals in the near future (Simons et al., 2004b). A near future can be next month whereas, by contrast, a distant future may consist of a timespan that is 5–7 years from now.

Common-sense thinking about human nature, suggests, of course, that we all have different timespans. Why is it that some of us have long FTPs and not others? Motivation, academic capability, personal ethos and philosophical belief, and/or confidence, in this case, may account for individual variations in the setting of a particular timespan. A pervasive question is, how do we encourage students to develop and sustain deep FTPs? Chou, from our previous example, may envisage a future timespan of 5–7 years for the completion of his Ph.D. Another student, by contrast, may not have this ambition and instead project a future timespan of a couple of months maximum. This example of disparity in future timespans has implications for educators, stakeholders, and policymakers. In particular, let us explore the topical theme of optimal best (Fraillon, 2004; Martin, 2006; Phan et al., 2016) and consider how this feat could serve as a means to encourage and promote an extended FTP.

Achievement of optimal best is subjective and, indeed, reflects a person's personal best (Fraillon, 2004; Martin, 2006; Liem et al., 2012; Phan et al., 2016). Optimal best in a academic subjects, for instance, is not static but may improve over time, as a result of cognitive maturity, increasing effort expenditure, in-depth understanding, and personal growth. For example, at present, a Year 8 student may indicate that her optimal best in "essay composition" is the capacity to write a scholarly 2000-word essay about the life of Erik Erikson. By the time this student is in first-year university, she would reflect and realize that her previous optimal best is now somewhat "low" in terms of cognitive capability. No doubt, at this stage of cognitive maturity, composing and writing a 2000-word essay about Erikson's life would be perceived as being easy. As has been discussed previously (Phan and Ngu, 2017a), a person's optimal best at the present time (i.e., L_2) eventually becomes prior cognitive

experience (i.e., L_1) and forms part of his/her repertoire of knowledge. What this entails is that there is no definitive limit to a person's optimal best. However, an "unrealistic" level of optimal best would not be conducive, giving rise to inaccurate results.

To date, to our knowledge, there has been very little study of the importance of timespan with reference to the added factor of a person's optimal achievement best. This line of inquiry, for us, is innovative and stipulates the notion that cognitive complexity of optimal best could act as a catalyst to encourage and facilitate the setting of deep FTPs. We propose that for a high level of cognitive complexity of optimal best to develop requires a certain amount of time into the future for its development. In terms of schooling, for example, we can consider different levels of cognitive complexity (van Merriënboer et al., 2003; van Merriënboer and Sweller, 2005) – for example: (i) a Year 8 student, Thomas, striving to achieve a maximum-scoring football season; (ii) a Year 9 student, Melissa, seeking to overcome a health issue; or (iii) a group of Year 12 students wanting to enroll in a postgraduate medical degree program. Clearly, some tasks and activities are less complex and require only a modest range of focus, motivation, persistence, and effort expenditure, whereas others are more complex and demand much more time, effort, persistence, etc. A deep future timespan, we suggest, is healthy and as such, it is sufficient to encourage individuals to consider a more complex optimal best – for example, within the context of academic learning, a student may choose to strive to achieve an "A" grade for algebra (Phan et al., 2020).

Achieving optimal best is a personal endeavor. In school contexts, there is a difference between the problem of "$4(x + 5) = 3(x - 7)$, solve for x" and the problem of "$4(x + 8)^2 = 10$, solve for x" for L_2. Likewise, there is a difference between knowing how to compose a two-page essay and knowing how to write a 100-page thesis dissertation in Chinese Mandarin. The level of cognitive complexity of L_2, as we explained, is dependent on a number of factors, such as a student's current level of understanding and knowledge and their state of motivation. Our argument is that setting a complex optimal best is a desirable feat, as this could stimulate and encourage a student to envisage an extended timespan into the future. This rationalization is logical as more time and effort would be needed for a student to achieve a complex level of optimal best. How long would it take a person to solve the "Collatz Conjecture" problem? How much time does a senior citizen need in order to achieve and experience a state of self-actualization? Likewise, for an extremely knowledgeable student, how much time would he/she need to achieve a moderate level of cognitive complexity of optimal best? These questions, for us, emphasize the important need for a person *to contemplate and strive for complex optimal bests* in different academic subjects, given this feat would then form the premise for him/her to consider a correspondingly appropriate future timespan. In essence, this postulation considers the potential influence of the cognitive complexity of optimal best in facilitating a person's projection of a future timespan.

Achieving optimal best, we contend, may also reflect a person's perceived sense of subjective task value for the given task. Successful achievement of optimal best in an academic subject would highlight the perceived interest, appreciation, and/or placement of importance that the person has for the subject itself – for example, the student sees this learning as relevant for their future career plans. Some task that is relatively simple, by contrast, could instead indicate a non-meaningful subjective value and insignificant concern for the subject matter. Indeed, a relatively simple task would not hold any significance and, more importantly, a person would not have much expectation about its importance. This consideration differs somewhat from the discussion in the preceding section, where we argued that the perceived subjective task value and high expectation would act as psychological agents (Phan et al., 2017, 2019c) in the optimization of achievement best. Here, in this section, we postulate that cognitive complexity of optimal best could influence the perception of the relevance, significance, and/or importance of a learning task. Something that is perceived as being difficult for optimal achievement is likely to instill a strong sense of belief in its subjective value. In the context of schooling, for example, the creation of high cognitive complexity in a topical theme for optimal best achievement would convey to students the message that the subject matter itself is of value. Such discourse (i.e., cognitive complexity of optimal best → value), in turn, would connote a specific future timespan for consideration. In contrast, of course, cognitive simplicity of the subject matter would indicate subjective perceptions of unimportance and insignificance.

In Summary

The preceding sections have provided theoretical and conceptual accounts of the interrelationship between FTP and optimal achievement best. Our conceptualization, we contend, indicates a potential cyclic system, as shown in **Figure 2**. This cyclical system depicts the potential effect of a complex level of optimal achievement best on a deep extended future timespan and, in turn, this deep extended future timespan is observed as positively influencing the complex level of optimal achievement best. What does this conceptualization depict? We can consider the following:

- A deep timespan into the future is advantageous and beneficial, providing subjective grounding and opportunities for the successful enactment of optimization in order to facilitate in the striving of complex optimal bests (i.e., deep extended future timespan → complex optimal best).
- Achieving complex optimal bests requires adequate effort and time and, hence, a deep timespan into the future for a person to envisage what would be needed (i.e., complex optimal best → deep extended future timespan).

The research into FTP is clear and consistent (Lewin, 1951; Nuttin, 1964; Gjesme, 1983; de Bilde et al., 2011; Taylor and Wilson, 2019), while the study of the nature of optimal achievement best is still progressing (e.g., Phan et al., 2018, 2019a,b). One notable line of inquiry regarding optimal achievement that is of interest, at present, is related to its methodological account (Phan et al., 2019c) – for example, how

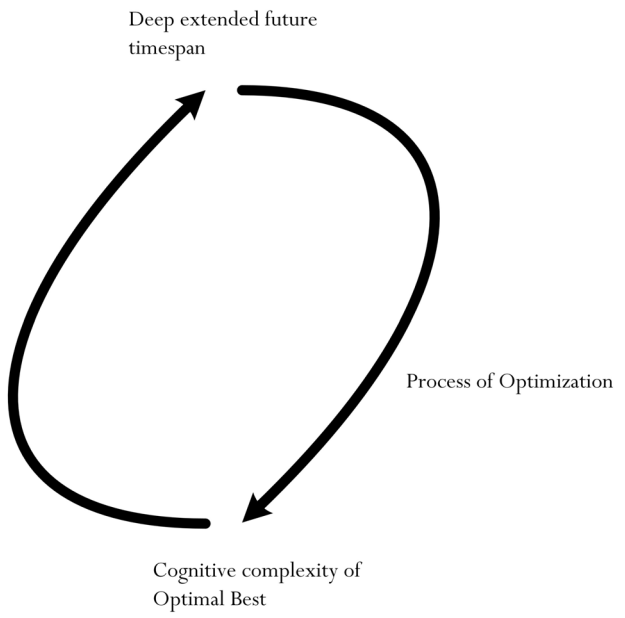


FIGURE 2 | Cyclic relationship between extended timespan and cognitive complexity of optimal best.

do we measure, assess, and evaluate the true nature of optimal best? To do so our research has established a conceptualization that attempts to integrate two independent research inquiries into one holistic model. This consideration is innovative as it places emphasis on the positive interrelation of the potential for achieving optimal achievement best and extended future timespans. Arising from our description is a desirable proposition for development: *a high level of cognitive complexity of optimal best* and *a deep, meaningful future timespan* (**Figure 2**).

CONCLUSION AND FUTURE DIRECTIONS

Time is a mysterious concept. We cannot go into the past and change our past behaviors, feelings, experiences, etc. What we can do, though, is reflect upon and use past experiences to inform the present moment and set future goals for accomplishment. The future is unknown and, in many cases, we will never truly know what the future holds. Motivational research into the proactivity of human agency has led to extensive research inquiries emphasizing the interconnections between a person's time experiences and his/her performance outcomes, academically and non-academically. Future time perspective is an interesting psychological concept for study, given that the future is always unknown and uncertain. One interesting aspect for development, in particular, is related to the promotion of the projection of deep, meaningful future timespans. Deep FTPs are healthy as they facilitate a heightened state of personal resolve, motivation, and effort expenditure.

It is not always easy for a person to spend time anticipating his/her future. Many of us, in this case, are content with our lives in the present moment. Thinking about future trajectories often appears to people as being somewhat meaningless, especially in light of day-to-day work demands and personal commitments, etc. By the same token, of course, it is uncertain whether a person's anticipation of future goals, aspirations, etc., would be accurate or realistic enough to have a credible meaning. Despite this plausible criticism, we have proposed from an educational point of view that a focus on extended future time goals is a useful point for encouragement. We usually have to make plans anyway (e.g., "What I would like to do in the next 12 months"), organize our time schedules, work through our finances, etc. On this basis, anticipation of a positive future should be encouraged as this would direct a person's focus and commitment toward something that is tangible. Accordingly, we need to consider the stipulation of an appropriate timespan. How "much" time into the future should one think about, anticipate, and/or project? As existing research (e.g., Simons et al., 2004a,b; Zimbardo and Boyd, 2008) has shown, an extended FTP is more healthy and robust than a short FTP, as this directs and motivates a person to work harder toward achieving his/her future goals. At the same time, a deep FTP may reveal and reflect a person's state of seriousness, degree of commitment, and ambition. Having said this, of course, we also raise the issue whether a deep extended FTP could give rise to negative outcomes. How much time into the future should one anticipate before that timespan causes problems?

Our discussion has provided an in-depth examination of how FTP could relate to the nature of optimal achievement best. Specifically, within the context of educational application, our proposition emphasizes two major interrelated paths for further development: (i) whether variations in a person's future timespan and, more importantly, increasing temporal distances of FTP (say, from 5 months to 5 years) motivate or detract from the achievement of cognitively complex optimal bests; and (ii) whether the types of complex optimal bests found in different academic subjects necessarily impact upon the extent of projection of FTPs. This proposal for a line of inquiry is both conceptual and philosophical, but also, we feel, requires empirical data to validate and affirm its value. What we have established so far is a preliminary inquiry; theoretically derived from existing independent research in the areas of FTP and optimal achievement best. There is a need for researchers to develop appropriate methodological designs that could measure, assess, and validate these proposed lines of inquiry (the nature of the relationships: Extended FTP → Complex OB and Complex OB → Extended FTP). For example, in relation to our proposed conceptualization, we suggest it would be worthwhile for researchers to develop different types of cognitive tasks, learning activities, etc., and to explore how these associate with a person's projection of a future timespan. Is there a direct correlation, which could/would then result in a statistically-derived equation, pattern, etc., for general application?

When we refer to time then, methodologically, research in this area necessarily would rely upon the collection of longitudinal data. An extended FTP places emphasis on a time duration or a time period into the future and, as such, it is sufficient to say that there is a "time difference," denoted as $\Delta_{(T1-T2)}$,

between now (T_1), the present moment, and a particular point into the future (T_2). It also needs to be kept in mind that the difference between T_1 and T_2 would not necessarily reflect either regular and constant development or at some time-point become instantaneous. **Figure 3** summarizes our description and shows a simple methodological design, which consists of the collection of data on two separate occasions, T_1 and T_2. As an aside, it can be noted that in terms of the issues raised in this paper the difference between T_1 and T_2 is, indeed, analogous with an extended FTP. A positive quantitative difference between T_1 and T_2, we contend, should be reflect a realistic human-centered timespan (e.g., 1 year from now). What does this methodological proposal mean in terms of optimal achievement best?

The current authors recently published a detailed theoretical account of the process of optimization, where we specifically delved into the issues of methodology – i.e., how do we measure and assess optimization (Phan et al., 2019c)? One recommendation that we proposed included the use of the same measure of best practice (e.g., cognitive competence test) on multiple occasions (e.g., CCT-T_1 and CCT-T_2, where CCT = cognitive competence test) in order to provide a proxy indicator of the enactment of optimization. From **Figure 3**, we now propose that in order to attain an understanding and affirmation of the actual achievement of a goal at T_2 (e.g., achieving optimal best in the topic of algebra), we would need to collect data on two occasions – say, now, T_1, and 1 year from now, T_2. We contend that it is inadequate and somewhat limited, in this case, simply to measure and collect data at T_1. In other words, a measurement of achievement at the conclusion of an extended timespan (e.g., the achievement of a future goal at T_2), alone, is limited as we would not know what actually occurred at T_1 that may have led to achievement at T_2. In a similar but converse vein, as we argued in our recent article (Phan et al., 2019c), a true indication of a person's optimal achievement best still would require its measure at T_2 and not just at T_1.

Validating the cyclic system shown in **Figure 2** would require the use of multi-wave panel data (Marsh and Yeung, 1997, 1998). This longitudinal methodological design could provide a grounding for researchers to establish the following patterns, say: T_1 extended future timespan → T_2 optimal achievement best; and T_1 optimal achievement best → T_2 extended future timespan (note: "→" = prediction). Evidence obtained could establish a cause-and-effect model for further experimental manipulation (Phan and Ngu, 2017a). In a similar vein, personal commitment, or a lack thereof, and what this means for a person's subsequent achievement and fulfillment of future goals is a correlative line of inquiry. This line of inquiry is interesting and was raised by one of our reviewers for consideration. Indeed, personal commitment may give rise to a student's personal resolve to strive for both short-term and long-term successes (Phan et al., 2020).

We acknowledge from the literature that having a deep future timespan is valuable, as this instills a sense of motivation and guides and directs a person's cognition and behavior toward the future goal(s) (e.g., a student's striving to enroll in medical school). At the same time, of course, a deep future timespan may reflect a person's hopes and ambitions, as well as his/her mental fortitude, to achieve different types of long-term goals. Having said this, however, we do have some reservations, which we previously described, regarding the "depth" of future timespan (e.g., 5 years in duration). For example, in the context of schooling, how does a student remain autonomous, independent, and/or disciplined enough to sustain his/her state of motivation, and/or to remain on task over a 3-year period? Uncertainties, personal circumstances, and extraneous influences may, individually and/or in combination, act to derail, negate and demotivate a student from maintaining his/her state of motivation and discipline to remain on task. When this is the case, we would not necessarily expect the student to achieve and fulfill his/her future goals. On this basis, what would educators and/or researchers have to understand to counter the problem of sustaining motivation, focus, discipline, etc.?

In a recent study, Oyama et al. (2018) proposed a term known as the "Hemingway effect," which is defined as "a positive effect of not completing a task" (p. 8). According to the authors' rationalization, "there are certain conditions, [when] not completing a task can actually enhance people's motivation to engage in the task – to complete or continue it" (p. 8). The *Hemingway effect*, we contend, may associate with our previous concern regarding the sustaining and continuation of a student's motivation and discipline to achieve future long-term goals. Inthis analysis, we postulate that the Hemingway effect could act to guide a student's state of cognition, motivation, and/or behavior over the course of time. The results of Oyama et al.'s (2018)

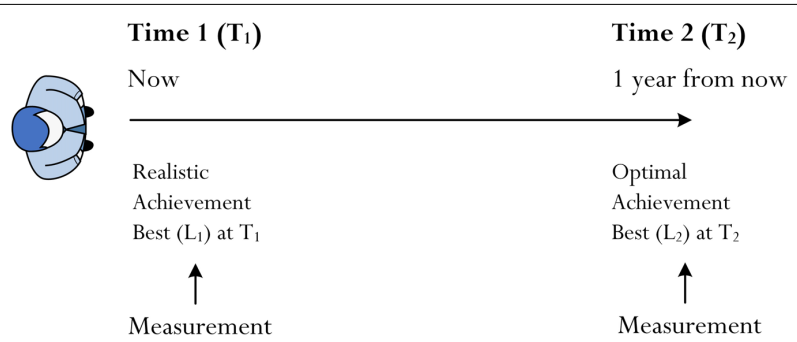


FIGURE 3 | Proposed methodological design.

study, interestingly, indicate that there are two conditions by which the Hemingway effect would occur: (i) explaining and clarifying for a student what is needed to complete the unfinished task(s); and (ii) a student's perception of closeness to completing that task. It would be of interest for researchers to consider the potential of the Hemingway effect to guide the way students and educators might envisage a deep future timespan to achieve optimal best. In other words, researchers could explore whether and/or to what extent a student's perception of closeness to completing a task helps to guide his/her focus of attention and discipline, as well as sustaining his/her state of motivation.

Finally, in tandem with Phan et al.'s (2019c) recent publication, we contend that the issue of measurement and assessment of optimal best practice is currently somewhat inconclusive. Most particularly, how does an educator measure and/or determine the complexity of a person's optimal best? This question is even more poignant when we incorporate the complicating factor of future timespans. Most particularly in this case, how would we determine the extent of accuracy or usefulness of a person's projection of his/her future timespan? In a similar vein and the subject at the core of this article, how could we accurately measure that the complexity of optimal best is aligned with a person's consideration of deep future timespan – that is, does achieving more complex levels of optimal best require more time, as suggested by our proposition, while achieving less complex levels of optimal best require shorter time periods? As have been discussed before, it is also of interest to consider the potential association between the cognitive complexity of optimal best and future timespan with reference to a person's determination and/or perceived sense of task-value for learning – in particular, the quest for researchers to design and develop a robust methodological design that could measure and assess a person's determination and/or perceived sense of task-value when optimal best is achieved and/or when a specific future timespan is set.

AUTHOR CONTRIBUTIONS

HP and BN contributed equally to the conceptualization, articulation, and write up of this manuscript. KM contributed to the revision of this manuscript.

ACKNOWLEDGMENTS

We would like to extend our appreciation and gratitude to the three reviewers for their insightful comments.

REFERENCES

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *J. Soc. Clin. Psychol.* 4, 359–373. doi: 10.1521/jscp.1986.4.3.359

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York, NY: W. H. Freeman & Co.

Barber, L. K., Munz, P. G., Bagsby, M. J., and Grawitch, C. (2009). When does time perspective matter? Self-control as a moderator between time perspective and academic achievement. *Personal. Ind. Diff.* 46, 250–253. doi: 10.1016/j.paid.2008.10.007

de Bilde, J., Vansteenkiste, M., and Lens, W. (2011). Understanding the association between future time perspective and self-regulated learning through the lens of self-determination theory. *Learn. Instr.* 21, 332–344. doi: 10.1016/j.learninstruc.2010.03.002

doi: 10.1037/0022-3514.42.3.566De Volder, M., and Lens, W. (1982). Academic achievement and future time perspective as a cognitive-motivational concept. *J. Pers. Soc. Psychol.* 42, 566–571. doi: 10.1037/0022-3514.42.3.566

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

Eguizábal, S. E., Berodia, ÁI., Portilla, J. B., and Ponce, B. (2018). Optimization model for school transportation design based on economic and social efficiency. *Trans. Policy* 67, 93–101. doi: 10.1016/j.tranpol.2018.01.015

Erikson, E. H. (1968). *Identity: Youth and Crisis*. New York, NY: Norton.

Fraillon, J. (2004). *Measuring Student Well-being in the Context of Australian Schooling: Discussion Paper*. Carlton South, VIC: The Australian Council for Research.

Fraisse, P. (1963). *The Psychology of Time*. New York, NY: Harper and Row.

Frank, L. K. (1939). Time perspectives. *J. Soc. Philosophy* 4, 293–312.

Franken, R. E. (2007). *Human Motivation*. Belmont, CA: Thomson Wadsworth.

Freund, A. M., and Baltes, P. B. (1998). Selection, optimization, and compensation as strategies of life management: correlations with subjective indicators of successful aging. *Psychol. Aging* 13, 531–543. doi: 10.1037//0882-7974.13.4.531 doi: 10.1037/0882-7974.13.4.531

Gjesme, T. (1983). On the concept of future time orientation: considerations of some functions' and measurements' implications. *Int. J. Psychol.* 18, 443–461. doi: 10.1080/00207598308247493

Goldstein, J., and Kornfield, J. (1987). *Seeking the Heart of Wisdom: The Path of Insight Meditation*. Boston, MA: Shambhala Publications, Inc.

Hanh, T. N. (1976). *Miracle of Mindfulness*. Boston, MA: Beacon.

Husman, J., Hilpert, S. K., and Brem, C. (2016). Future time perspective connectedness to a career: the contextual effects of classroom knowledge building. *Psychol. Belgica* 56, 210–225. doi: 10.5334/pb.282

Husman, J., and Lens, W. (1999). The role of the future in student motivation. *Educ. Psychol.* 34, 113–125. doi: 10.1207/s15326985ep3402_4

Israeli, N. (1934). Measurement of the outlook on the future. *Psychol. Bull.* 31, 643–646.

Janeiro, I. N., Duarte, A. M., Araújo, M., and Gomes, A. I. (2017). Time perspective, approaches to learning, and academic achievement in secondary students. *Learn. Ind. Diff.* 55, 61–68. doi: 10.1016/j.lindif.2017.03.007

Kabat-Zinn, J. (2015). Mindfulness. *Mindfulness* 6, 1481–1483. doi: 10.1007/s12671-015-0456-x

Kastenbaum, R. (1961). The dimensions of future time perspective, an experimental analysis. *J. Gen. Psychol.* 65, 203–218. doi: 10.1080/00221309.1961.9920473

Lennings, C. L., Burns, G., and Cooney, N. (1998). The profiles of time perspective and personality: developmental considerations. *J. Psychol.* 132, 629–641. doi: 10.1080/00223989809599294

Lens, W., PaixÃO, D., Herrera, A., and Grobler, P. (2012). Future time perspective as a motivational variable: content and extension of future goals affect the quantity and quality of motivation. *Jpn. Psychol. Res.* 54, 321–333. doi: 10.1111/j.1468-5884.2012.00520.x doi: 10.1111/j.1468-5884.2012.00520.x doi: 10.1111/j.1468-5884.2012.00520.x

Lens, W., Simons, J., and Dewitte, S. (2002). "From duty to desire: the role of students' future time perspective and instrumentality perceptions for study motivation and self-regulation," in *Academic Motivation of Adolescents*, eds F. Pajares and T. Urdan (Greenwich, CT: Information Age Publisher), 221–245.

LeShan, L. L. (1952). Time perspective and social class. *J. Abnor. Soc. Psychol.* 47, 589–592.

Lewin, K. (1942). *Time Perspective and Morale. Civilian morale. G. Watson*. Boston, MA: Houghton Mifflin, 103–124.

Lewin, K. (1951). *Field Theory in the Social Sciences: Selected Theoretical Papers*. New York, NY: Harper.

Liem, G. A. D., Ginns, P., Martin, A. J., Stone, B., and Herrett, M. (2012). Personal best goals and academic and social functioning: a longitudinal perspective. *Learn. Instr.* 22, 222–230. doi: 10.1016/j.learninstruc.2011.11.003

Loden, G. A. T. (1996). *Meditations on the Path to Enlightenment*. Reading, MA: Addison-Wesley.

Marsh, H. W., and Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: structural equation models of longitudinal data. *J. Educ. Psychol.* 89, 41–54. doi: 10.1037/0022-0663.89.1.41

Marsh, H. W., and Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: gender differences in the development of math and English constructs. *Am. Educ. Res. J.* 35, 705–738. doi: 10.3102/00028312035004705

Martin, A. J. (2006). Personal bests (PBs): a proposed multidimensional model and empirical analysis. *Br. J. Educ. Psychol.* 76, 803–825. doi: 10.1348/000709905X55389

Master, S. Y. (2010). *The Dharma Drum Lineage of Chan Buddhism: Inheriting the Past and Inspiring the Future*. Taipei: The Sheng Yen Education Foundation.

Mehta, P., Sundberg, N. D., Rohila, P. K., and Tyler, L. E. (1972). Future time perspectives of adolescents in India and the United States. *J. Cross-Cult. Psychol.* 3, 293–302. doi: 10.1177/002202217200300307

Milfont, T. L., Wilson, J., and Diniz, P. (2012). Time perspective and environmental engagement: a meta-analysis. *Int. J. Psychol.* 47, 325–334. doi: 10.1080/00207594.2011.647029

Moreas, A., and Lens, W. (1991). *The Motivational Meaning of the Individual Future Time Perspective (Internal Research Report)*. Belgium: Research Center for Motivation and Time Perspective.

Ngu, B. H., and Phan, H. P. (2016). Unpacking the complexity of linear equations from a cognitive load theory perspective. *Educ. Psychol. Rev.* 28, 95–118. doi: 10.1007/s10648-015-9298-2

Ngu, B. H., and Yeung, A. S. (2012). Fostering analogical transfer: the multiple components approach to algebra word problem solving in a chemistry context. *Contem. Educ. Psychol.* 37, 14–32. doi: 10.1016/j.cedpsych.2011.09.001

Ngu, B. H., Yeung, A. S., and Tobias, S. (2014). Cognitive load in percentage change problems: unitary, pictorial, and equation approaches to instruction. *Instruct. Sci.* 42, 685–713. doi: 10.1007/s11251-014-9309-6

Nuttin, J. R. (1964). The future time perspective in human motivation and learning. *Acta Psychol.* 23, 60–83.

Nyanaponika, T. (1972). *The Power of Mindfulness*. San Francisco, CA: Unity Press.

Oyama, Y., Manalo, E., and Nakatani, Y. (2018). The Hemingway effect: how failing to finish a task can have a positive effect on motivation. *Think. Skills Creat.* 30, 7–18. doi: 10.1016/j.tsc.2018.01.001

Phan, H. P. (2015). The Impact of FTP on career choices: situating within a social cognitive perspective. *High. Educ. Res. Dev.* 34, 368–382. doi: 10.1080/07294360.2014.958069

Phan, H. P., and Ngu, B. H. (2017a). "Positive psychology: the use of the Framework of Achievement Bests to facilitate personal flourishing," in *Well-Being and Quality of Life*, ed. A. A. V. Boas (Rijeka: Intech: Open Science), 19–33.

Phan, H. P., and Ngu, B. H. (2017b). Undertaking experiments in social sciences: sequential, multiple time series designs for consideration. *Educ. Psychol. Rev.* 29, 847–867. doi: 10.1007/s10648-016-9368-0

Phan, H. P., and Ngu, B. H. (2019a). "Expanding on the theoretical concept of 'optimization' for effective learning: establishing empirical evidence from an Eastern sociocultural perspective," in *Asian Education Miracles*, eds A. D. Liem and S. T. Hong (Oxon, NY: Routledge), 222–240. doi: 10.4324/9781315180625-14

Phan, H. P., and Ngu, B. H. (2019b). *Teaching, Learning and Psychology*. Docklands: Oxford University Press.

Phan, H. P., Ngu, B. H., and Alrashidi, O. (2018). Contextualised self-beliefs in totality: an integrated framework from a longitudinal perspective. *Educ. Psychol.* 38, 411–434. doi: 10.1080/01443410.2017.1356446

Phan, H. P., Ngu, B.-H., and Williams, A. (2016). Introducing the concept of optimal best: theoretical and methodological contributions. *Education* 136, 312–322.

Phan, H. P., Ngu, B. H., and Yeung, A. S. (2017). Achieving optimal best: instructional efficiency and the use of cognitive load theory in mathematical problem solving. *Educ. Psychol. Rev.* 29, 667–692. doi: 10.1007/s10648-016-9373-3

Phan, H. P., Ngu, B. H., and Yeung, A. S. (2019c). Optimization: in-depth examination and proposition. *Front. Psychol.* 10. doi: 10.3389/fpsyg.2019.01398

Phan, H. P., Ngu, B.-H., Wang, H.-W., Shih, J.-H., Shi, S.-Y., and Lin, R.-Y. (2019a). Achieving optimal best practice: an inquiry into its nature and characteristics. *PLoS ONE* 14:e0215732. doi: 10.1371/journal.pone.0215732

Phan, H. P., Ngu, B. H., Lin, R.-Y., Wang, H.-W., Shih, J.-H., and Shi, S.-Y. (2019b). Predicting and enhancing students' positive emotions: an empirical study from a Taiwanese sociocultural context. *Heliyon* 5:e02550. doi: 10.1016/j.heliyon.2019.e02550

Phan, H. P., Ngu, B. H., Shih, J.-H., Lin, R.-Y., Shi, S.-Y., and Wang, H.-W. (2020). Validating newly developed 'optimizing' concepts: the importance of personal resolve, effective functioning, and academic striving. *Educ. Psychol.* 40, 448–472. doi: 10.1080/01443410.2019.1693507 doi: 10.1016/j.heliyon.2019.e02550

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., and Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: a meta-analytic approach. *Rev. Educ. Res.* 81, 493–529. doi: 10.3102/0034654311421793

Seijts, G. H. (1998). The importance of future time perspective in theories of work motivation. *J. Psychol.* 13, 154–168. doi: 10.1080/00223989809599156

Seligman, M., and Csíkszentmihályi, M. (2000). Positive psychology: an introduction. *Am. Psychol.* 55, 5–14.

Seligman, M., Ernst, R. M., Gillham, J., Reivich, K., and Linkins, M. (2009). Positive education: positive psychology and classroom interventions. *Oxf. Rev. Educ.* 35, 293–311. doi: 10.1080/03054980902934563

Simons, J., Dewitte, S., and Lens, W. (2004a). The role of different types of instrumentality in motivation, study strategies, and performance: know why you learn, so you'll know what you learn!. *Br. J. Educ. Psychol.* 74(Pt 3), 343–360. doi: 10.1348/0007099041552314

Simons, J., Vansteenkiste, M., Lens, W., and Lacante, W. (2004b). Placing motivation and future time perspective theory in a temporal perspective. *Educ. Psychol. Rev.* 16, 121–139. doi: 10.1023/b:edpr.0000026609.94841.2f

Taylor, J., and Wilson, J. C. (2019). Using our understanding of time to increase self-efficacy towards goal achievement. *Heliyon* 5:e02116. doi: 10.1016/j.heliyon.2019.e02116

van Merriënboer, J., Kirschner, P. A., and Kester, L. (2003). Taking the load off a learner's mind: instructional design for complex learning. *Educ. Psychol.* 38, 5–13. doi: 10.1207/s15326985ep3801_2

van Merriënboer, J., and Sweller, G. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educ. Psychol. Rev.* 17, 147–177. doi: 10.1007/s10648-005-3951-0

Wallace, M. (1956). Future time perspective in schizophrenia. *J. Abnor. Soc. Psychol.* 52, 240–245. doi: 10.1037/h0039899

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: a developmental perspective. *Educ. Psychol. Rev.* 6, 49–78. doi: 10.1007/bf02209024

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

Yeshe, L., and Rinpoche, L. Z. (1976). *Wisdom Energy: Basic Buddhist Teachings*. Somerville, MA: Wisdom Publications.

Ziegelmann, J. P., and Lippke, S. (2007). Use of selection, optimization, and compensation strategies in health self-regulation – Interplay with resources and successful development. *J. Aging Health* 19, 500–518. doi: 10.1177/0898264307300197

Zimbardo, P. G., and Boyd, J. N. (1999). Putting time in perspective: a valid, reliable individual-differences metric. *J. Pers. Soc. Psychol.* 77, 1271–1288. doi: 10.1037/0022-3514.77.6.1271

Zimbardo, P. G., and Boyd, J. N. (2008). *The Time Paradox: The New Psychology of Time That Will Change Your Life*. New York, NY: Free Press.

Zimbardo, P. G., Keough, K. A., and Boyd, J. N. (1997). Present time perspective as a predictor of risky driving. *Personal. Ind. Diff.* 23, 1007–1023. doi: 10.1016/s0191-8869(97)00113-x

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.

Copyright © 2020 Phan, Ngu and McQueen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 09 December 2020
doi: 10.3389/fpsyg.2020.515531

Neural Network Model for Predicting Student Failure in the Academic Leveling Course of Escuela Politécnica Nacional

Iván Sandoval-Palis[1], David Naranjo[1], Raquel Gilar-Corbi[2] and Teresa Pozo-Rico[2]*

[1] Departamento de Formación Básica, Escuela Politécnica Nacional, Quito, Ecuador, [2] Department of Developmental Psychology and Didactics, University of Alicante, Alicante, Spain

The purpose of this study is to train an artificial neural network model for predicting student failure in the academic leveling course of the Escuela Politécnica Nacional of Ecuador, based on academic and socioeconomic information. For this, 1308 higher education students participated, 69.0% of whom failed the academic leveling course; besides, 93.7% of the students self-identified as mestizo, 83.9% came from the province of Pichincha, and 92.4% belonged to general population. As a first approximation, a neural network model was trained with twelve variables containing students' academic and socioeconomic information. Then, a dimensionality reduction process was performed from which a new neural network was modeled. This dimension reduced model was trained with the variables application score, vulnerability index, regime, gender, and population segment, which were the five variables that explained more than 80% of the first model. The classification accuracy of the dimension reduced model was 0.745, while precision and recall were 0.883 and 0.778, respectively. The area under ROC curve was 0.791. This model could be used as a guide to lead intervention policies so that the failure rate in the academic leveling course would decrease.

Keywords: neural network, predictive modeling, student success, academic leveling course, learning analytics, academic performance

OPEN ACCESS

Edited by:
José Jesús Gázquez, University of Almería, Spain

Reviewed by:
Tulio Ramírez, Universidad Católica Andrés Bello, Venezuela
Johanna Vásquez, National University of Colombia, Colombia

****Correspondence:***
Iván Sandoval-Palis
ivan.sandoval@epn.edu.ec

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

Received: *28 November 2019*
Accepted: *10 November 2020*
Published: *09 December 2020*

Citation:
Sandoval-Palis I, Naranjo D, Gilar-Corbi R and Pozo-Rico T (2020) Neural Network Model for Predicting Student Failure in the Academic Leveling Course of Escuela Politécnica Nacional. Front. Psychol. 11:515531. doi: 10.3389/fpsyg.2020.515531

INTRODUCTION

The concern of universities about the quality of the educational service they offer has triggered several and continuous evaluation processes to detect the underlying problems and act in this regard (Sandoval et al., 2019). The problems identified through these evaluation processes include several aspects of the education system; nevertheless, one of the most serious is the high rate of student failure in university education, which is significantly higher during the first year of studies. For instance, in South America during the last decade, about 50% of students failed their university studies during their first year (Parrino, 2010). The results of several studies provide evidence that student failure is influenced by an interaction of various factors which are decisive throughout the academic process (Amaya et al., 2015; Montoya Gutiérerz, 2016; Amo and Santelices, 2017; Lara et al., 2017).

In Ecuador, several studies have been performed to identify factors influencing student failure. In this context, factors such as family monthly income, type of school, type of housing, and even gender have been identified as components that intervene in the phenomenon of student failure (Sandoval et al., 2018). At the same time, the government and universities have proposed affirmative action policies to help students overcome the difficulties triggered by the influence of these factors. Therefore, identifying these factors and analyzing their influence on students' academic performance is an important process to be performed in order to early identify at-risk students and, consequently, implement corrective actions in the educational process (Di Caudo, 2015; Sandoval et al., 2019).

Artificial neural networks (ANNs) have optimal features to conduct this type of analysis due to its excellent prediction and classification performance. An artificial neural network is a reticular computer system that learns from experience by self-modifying its connections; in this way, data prediction can be estimated based on a wide range of information (Cao et al., 2018; Zhou et al., 2018; Bouwmans et al., 2019).

Artificial neural networks are helpful tools used for data analysis with which functional relationships between variables can be found and modeled. Indeed, they allow exploring relationships or models that otherwise could not be found, for example, by using traditional statistical procedures (Bouwmans et al., 2019). In addition, due to the fact that they are a type of machine learning algorithm, ANNs have advantages over traditional statistical methods when applied to studies in which input data are incomplete or ambiguous by nature. They also have good performance when studying non-linear problems or data with a lot of "noise" and can be applied even without meeting theoretical assumptions related to traditional statistic because ANNs decode the information implicit in the data (Cao et al., 2018).

Currently, ANNs are widely applied to solve prediction and classification problems in areas as diverse as Meteorology and Spectroscopy (Timoshenko et al., 2018; Lee and Kim, 2019). The application of ANN for studying academic performance has gained significance in recent years not only because of its higher performance but also because of the findings regarding factors that influence the educational process (Vandamme et al., 2007; Marbouti et al., 2016; Baars et al., 2017; Mason et al., 2018; Figueiredo et al., 2019).

The purpose of this study was to develop a neural network model for predicting student failure in the academic leveling course of the Escuela Politécnica Nacional of Ecuador based on academic and socioeconomic information.

METHODS

Participants

The participants in this study were 1308 higher-education students from the Escuela Politécnica Nacional of Ecuador whose characteristics are shown in **Table 1**.

TABLE 1 | Characteristics of the study population.

Variable	Distribution			
Gender	63.2% male			
	36.8% female			
Ethnicity	93.7% mestizo	0.9% Mulatto		
	2.7% indigenous	0.3% black		
	1.1% white	0.1% Montubio		
	1.0% afro-descendant	0.2% other		
Province of origin	83.9% Pichincha	1.2% Carchi	0.5% Chimborazo	0.1% Santa Elena
	3.1% Tungurahua	1.1% Imbabura	0.5% Sucumbíos	0.1% Zamora Chinchipe
	2.8% Cotopaxi	1.0% El Oro	0.3% Guayas	0.1% Orellana
	2.4% Santo Domingo de los Tsáchilas	0.5% Azuay	0.3% Manabí	
	1.4% Esmeraldas	0.5% Bolívar	0.2% Loja	
Population segment	92.4% general population			
	5.6% affirmative action			
	1.5% territorial merit			
	0.5% high-performance group			
Regime	35.9% Sierra			
	64.1% Costa			
Academic performance	31.0% passed			
	69.0% failed			

Measures

The variables to predict (levpass) states whether a student failed the academic leveling course while the following twelve variables were considered as predictors:

1. Application score (appscore): The score achieved by the students in the university application exam. This exam is graded between 400 and 1000 points. The higher the score, the higher the student's performance in the exam. The application score does not consider students' high school GPA.
2. Vulnerability index (vulnind): This index shows the relative socioeconomic vulnerability of a student. It is rated over 1000 points. The higher the index, the lower the socioeconomic vulnerability. The index is calculated from the information stated by the students in a socioeconomic survey during the university application process.
3. Gender (gender): student's gender.
4. Population segment (popsgmnt): It is a way of classifying students according to their academic performance and socioeconomic characteristics. There are four types of population segments:

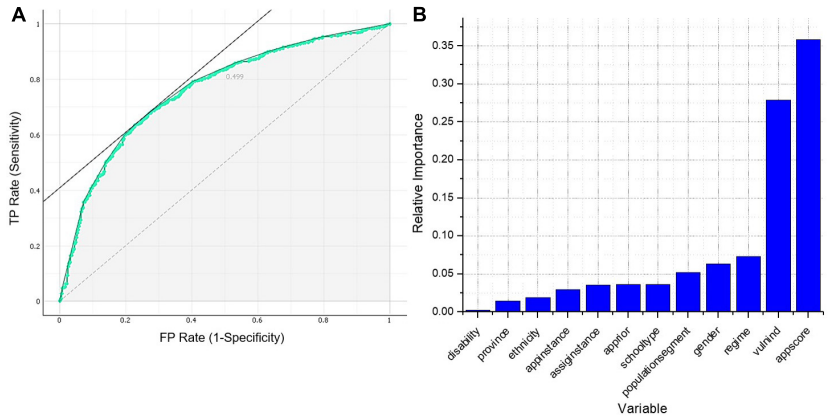


FIGURE 1 | (A) ROC curve of neural network 48–39–1. **(B)** Garson's weighted importance of the variables of neural network 48–39–1.

- High-performance group (HPG): This group is formed by the applicants best scored in the university application exam.
- Territorial merit: This group is formed by the best graduates of the educational institutions whether these are public, municipal, or fisco-misional schools.
- Affirmative action: This group is formed by the applicants in a situation of vulnerability which considers their socioeconomic situation, disability, territoriality (provinces with a lower rate of access to higher education), and other conditions of vulnerability.
- General population: This group is formed by the students who do not belong to any of the other population segments.

5. Application priority (appprior): The priority of the chosen degree made by the student during the university application process.
6. Application instance (appinstance): The instance in which a student applied to the university. It can be the first, the second, or the third instance.
7. Assignation instance (assiginstance): The instance in which the student was assigned a place in the university. It can be the first, the second, the third, the fourth, or the fifth instance of assignment of places.
8. School type (schooltype): The type of school a student comes from. It can be public, municipal, private, lay private, lay, religious private, fisco-misional, or foreign.
9. Regime (regime): The academic regime in which a student enters the university. It can be Costa or Sierra. The Sierra regime is analogous to the autumn semester, which runs from September to February, whereas the Costa regime is analogous to the spring semester, which runs from March to August.
10. Province (province): the province a student comes from.
11. Ethnicity (ethnicity): The ethnic group to which a student belongs. This information is self-stated by each student.
12. Disability (disability): This variable states whether a student has a disability.

Procedures

Data were collected from the existing computer records in the administration of the Escuela Politécnica Nacional of Ecuador with permission granted by the academic staff of the Institution. The data provided by the institution were anonymous.

Artificial Neural Network Modeling

In this study, an ANN model was used to predict student failure from their academic and socioeconomic information.

First, data were partitioned into training (70% of cases), validation (15% of cases), and testing (15% of cases) sets by a random sampling process. Then, an ANN model was trained on the training data set. The hyperparameters of the ANN model were one hidden layer, logistic activation function, Adam optimization algorithm, learning rate of 0.0001, and 4000 as the maximum number of iterations. Different ANN models with varying numbers of neurons in the hidden layer were validated. The number of neurons in the hidden layer was defined according to Eq. (1):

$$\frac{2}{3}NIL + NOL \leq NHL \leq 2NIL \tag{1}$$

where NHL is the number of neurons in the hidden layer, NIL is the number of neurons in the input layer, and NOL is the number of neurons in the output layer.

Once the model reached the maximum fit on the validation set, the ANN model was tested on the testing set in order to measure its performance (Teoh et al., 2006).

Dimensionality Reduction

As a first approximation, the ANN model was trained with the twelve variables containing students' academic and socioeconomic information which resulted in a complex model. Therefore, in order to reduce the dimensions of the model, three tests were performed: Garson test, ANOVA, and chi-squared test.

From Garson test, the relative importance of each of the variables considered in the model was determined, whereas, from ANOVA, it was determined whether there was a significant difference in the application score and the vulnerability index between the students who failed and those who passed the leveling course. On the other hand, the chi-squared test was performed to determine whether categorical variables were independent of each other. For ANOVA and chi-squared test, a significance level of $p = 0.05$ was chosen. Consequently, only those variables that had a relative importance greater than 5% and with a p-value less than 5% were chosen to train a new model. This dimension reduced model was trained according to the criteria stated in the preceding section (Vandamme et al., 2007; Helal et al., 2018).

Model Performance Evaluation Criteria

Although the procedure used to train the models allows obtaining a first approximation of their performance, this procedure is very susceptible to problems of overfitting on the training and validation sets. Therefore, the general performance of the models was determined by cross-validation with $k = 10$, which minimizes the effects of overfitting and selection bias (Cawley and Talbot, 2010; Juba and Le, 2019). The performance of the model was evaluated through the following indicators:

- Accuracy: the ratio of the total number of correct predictions to the total number of predictions.
- Precision: the ratio of positive class predictions that actually belong to the positive class.
- Recall: the ratio of positive class predictions made out of all positive cases.
- Area under the receiver operating characteristic (ROC) curve, which provides an aggregate measure of performance across all possible classification thresholds.

Data analysis and modeling were performed in SPSS 22, Orange 3.22.0, and RStudio Version 1.2.1335.

RESULTS

The ANN trained with the twelve predictor variables was modeled on an architecture of forty-eight neurons in the input layer and one neuron in the output layer. Each of the neurons in the input layer corresponds to each of the possible categories in the predictor variables. The highest performance for this model was achieved with thirty-nine neurons in the hidden layer; hence, the resulting architecture of this ANN can be written as 48–39–1. The classification accuracy for this model was 0.732, while precision and recall were 0.833 and 0.789, respectively. **Figure 1A** shows the ROC curve; the resulting value for the area under this curve was 0.757.

The results of the Garson test are presented in **Figure 1B**. Application score, vulnerability index, regime, gender, and population segment were the five variables that showed a relative importance greater than 5%, where the application score and vulnerability index resulted to be the most important for the model. Also, these variables explained more than 80% of the model.

TABLE 2 | Chi-squared test and ANOVA's p-values of the twelve predictor variables.

Variable	Levpass	Gender	Popsgmnt	Regime	Ethnicity	Apprior	Appinstance	Assiginstance	Schooltype	Province	Disability
Appscore	0.000										
Vulnind	0.001										
Levpass	–										
Gender	0.007	–									
Popsgmnt	0.034	0.377	–								
Regime	0.000	0.136	0.000	–							
Ethnicity	0.144	0.063	0.005	0.037	–						
Apprior	0.000	0.469	0.198	0.000	0.587	–					
Appinstance	0.001	0.018	0.005	0.006	1.000	0.001	–				
Assiginstance	0.000	0.030	0.028	0.000	0.991	0.000	0.000	–			
Schooltype	0.001	0.026	0.596	0.000	0.000	0.156	0.001	0.017	–		
Province	0.089	0.195	0.000	0.000	0.007	0.958	0.976	1.000	0.000	–	
Disability	0.180	0.127	0.403	0.000	0.353	0.043	0.649	0.103	0.027	1.000	–

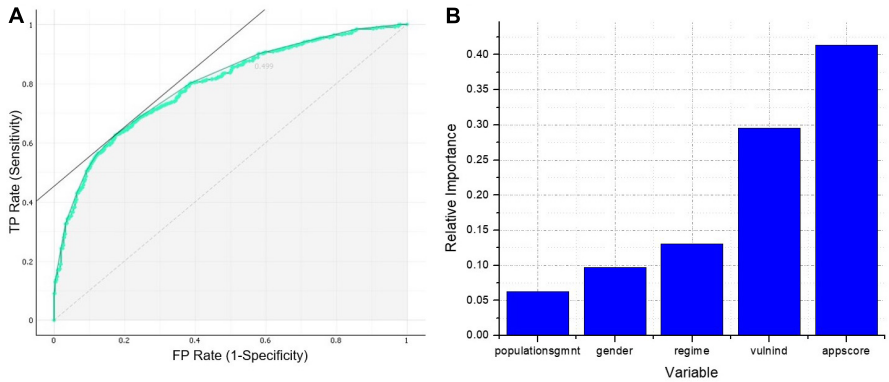


FIGURE 2 | (A) ROC curve of neural network 7–4–1. **(B)** Garson's weighted importance of the variables of neural network 7–4–1.

Chi-squared test and ANOVA's *p*-values are presented in **Table 2**. The application score and the vulnerability index of the students who passed the leveling course were significantly different from those who failed. Also, it is noted that disability, province, and ethnicity are not only variables with less relative importance according to the Garson test but also independent of student failure according to the chi-squared test. Additionally, regime is not independent of application priority, application instance, assignation instance, school type, and province, then, when including regime in the new model, the effect of the aforementioned five variables will be considered indirectly.

According to the results presented above, a dimension reduced model was trained with the variables application score, vulnerability index, regime, gender, and population segment.

The dimension reduced ANN model had seven neurons in the input layer and one neuron in the output layer. The highest performance for this model was achieved with four neurons in the hidden layer, which resulted in a 7–4–1 architecture. The classification accuracy for this model was 0.745, while precision and recall were 0.883 and 0.778, respectively. Also, the area under the ROC curve, shown in **Figure 2A**, was 0.791. The dimension reduced model was not only far simpler than the initial one but also higher on classification performance.

The results of the Garson test for the dimension reduced model are presented in **Figure 2B**, which indicates that all the five chosen variables had a relative importance greater than 5%, and, again, the two most important were application score and vulnerability index.

DISCUSSION

The purpose of this study was to identify as early and reliable as possible students that might fail the academic leveling course. Even though both models did not achieve a classification accuracy higher than 0.800, the dimension reduced model could be used as a guide to lead intervention policies so that the high rates of failure in the academic leveling course would decrease (Marbouti et al., 2016).

There are multiple reasons why both models could not classify correctly failing students. First, from a theoretical perspective, it is not possible to accurately model student failure because more factors could influence this phenomenon and it is almost impossible to include them all in the model. Therefore, any model will have limitations, resulting in misclassifying problems (Vandamme et al., 2007; Marbouti et al., 2016). Another reason is that biased and unbalanced data can significantly influence the classification performance of a model. However, for machine learning algorithms, omitting variables that have bias and balancing data should be performed cautiously (Cawley and Talbot, 2010). For instance, in this study, the variables ethnicity, province, and population segment have a considerable bias since 93.7% of the students self-identifies as mestizo, 83.9% comes from Pichincha, and 92.4% belongs to general population. Nevertheless, although ethnicity and province were not considered in the dimension reduced model, population segment was actually included to train this model. This is because, according to historical information from the Escuela Politécnica Nacional, students of affirmative action are more likely to fail the leveling course in comparison to students from other population segments. Thus, omitting variables that have bias and balancing data should only be done to decrease the bias in the sample data and make them more closely represent the real population. In fact, if population segment had not been included in the model, the omitted-variable bias would have occurred (Cawley and Talbot, 2010).

On the other hand, the variables that describe students' behaviors were not considered in this study. This information is of utmost importance since aspects such as previous motivation and the attitude with which students face their studies might be decisive when defining student success. However, variables that describe behavior are very likely to vary over time, so the model might present different results depending on the stage in which it is used (Marbouti et al., 2016; Helal et al., 2018; Mason et al., 2018). This last statement raises the question of whether an early diagnosis of student failure is effective.

Identifying students at risk of failing as early as possible is crucial. Nevertheless, there is important information that is generated as the semester elapses. Student academic performance

during the semester might be the best indicator of student success, but predictions based only on this information could provide results when there was not enough time to help students in need of academic support. Furthermore, predictions made too early are likely to be somewhat inaccurate (Helal et al., 2018; Yang and Li, 2018). Thus, in future research, finding the optimal time to utilize a prediction model should be an important point to focus on. Besides, this whole process should be kept as simple as possible so that time and resources would be optimized, and that is one of the most important reasons why the dimension reduced model turns out to be more attractive in terms of applicability (Mason et al., 2018).

The variables considered in the dimension reduced model not only describe socioeconomic as well as academic factors but also confirm their historical influence on student failure. Thus, for example, according to historical data, women have a higher failure rate compared to men, and students with a lower application score and in a situation of vulnerability tend to fail the leveling course. Also, students from the Costa regime are more likely to fail the leveling course in comparison to students from the Sierra regime (Sandoval et al., 2018, 2019). In this context, interventions by the government and the university should aim to mitigate students' economic difficulties, through financial aid or the strengthening of scholarship programs. On the other hand, regarding academic factors, efforts should focus on offering academic support before the academic leveling course as well as peer tutoring programs during the semester so that student failure and even dropout rates would decrease.

CONCLUSION

This study took the first step in identifying factors that influence student failure in the academic leveling course of the Escuela Politécnica Nacional of Ecuador. A dimension reduced artificial neural network was modeled from five variables containing students' academic and socioeconomic information. The variables used for training this model were application score, vulnerability index, regime, gender, and population segment. The model correctly classified 74.5% of the students that actually failed the leveling course, then, even though the model does not reach the maximum classification performance, it could be used as a guide to lead intervention policies, such as financial aid or academic support, so that the high failure rates in the academic leveling course would decrease.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

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 for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

IS-P was director of the research and carried out the quantitative methods. DN worked on quantitative methods. RG-C and TP-R carried out a theoretical review of the topic. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by Escuela Politécnica Nacional.

ACKNOWLEDGMENTS

The authors thank the Escuela Politécnica Nacional for the funding of the project PIJ-17-12: "Implementación de un Curso Preparatorio Piloto para Estudiantes de Grupos Vulnerables – Política de Cuotas."

REFERENCES

Amaya, Y., Barrientos, E., and Heredia, D. (2015). Student Dropout Predictive Model Using Data Mining Techniques. *IEEE Latin Am. Trans.* 13, 3127–3134. doi: 10.1109/tla.2015.7350068

Amo, C., and Santelices, M. V. (2017). *Trayectorias universitarias: MÁS QUE PERSISTENCIA O DESERCIÓN. Congresos CLABES.* Panama: Technological University of Panama. Available online at: http://revistas.utp.ac.pa/index.php/clabes/article/view/1676/2412

Baars, G. J. A., Stijnen, T., and Splinter, T. A. W. (2017). A Model to Predict Student Failure in the First Year of the Undergraduate Medical Curriculum. *Health Profess. Educ.* 3, 5–14. doi: 10.1016/j.hpe.2017.01.001

Bouwmans, T., Javed, S., Sultana, M., and Jung, S. K. (2019). Deep neural network concepts for background subtraction: A systematic review and comparative evaluation. *Neur. Net.* 117, 8–66. doi: 10.1016/j.neunet.2019.04.024

Cao, W., Wang, X., Ming, Z., and Gao, J. (2018). A review on neural networks with random weights. *Neurocomputing* 275, 278–287. doi: 10.1016/j.neucom.2017.08.040

Cawley, G. C., and Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *J. Machine Learn. Res.* 11, 2079–2107.

Di Caudo, M. (2015). Política de cuotas en Ecuador: me gané una beca para estudiar en la Universidad. *Ponto Vírgula Revista de Ciências Soc.* 1, 196–218. doi: 10.47212/tendencias_vii_2019_14

Figueiredo, J., Lopes, N., and García-Pĕalvo, F. J. (2019). Predicting student failure in an introductory programming course with multiple back-propagation. *ACM Int. Conf. Proc. Ser.* 2019, 44–49. doi: 10.1145/3362789.3362925

Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J., et al. (2018). Predicting academic performance by considering student heterogeneity. *Knowl. Based Syst.* 161, 134–146. doi: 10.1016/j.knosys.2018.07.042

Juba, B., and Le, H. S. (2019). Precision-Recall versus Accuracy and the Role of Large Data Sets. *Proc. AAAI Conf. Artific. Intel.* 33, 4039–4048. doi: 10.1609/aaai.v33i01.33014039

Lara, H. O., Silva, J. S., Galeano, M. O., Carreño, C. C., and Ariza, A. B. (2017). *Estudio factores asociados a la deserción estudiantil en la universidad minuto de dios de la sede virtual ya distancia. Congresos CLABES*. Available online at: http://revistas.utp.ac.pa/index.php/clabes/article/view/1691 (accessed on Nov 15, 2017).

Lee, D., and Kim, K. (2019). Recurrent neural network-based hourly prediction of photovoltaic power output using meteorological information. *Energies* 12:215. doi: 10.3390/en12020215

Marbouti, F., Diefes-Dux, H. A., and Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Comput. Educ.* 103, 1–15. doi: 10.1016/j.compedu.2016.09.005

Mason, C., Twomey, J., Wright, D., and Whitman, L. (2018). Predicting Engineering Student Attrition Risk Using a Probabilistic Neural Network and Comparing Results with a Backpropagation Neural Network and Logistic Regression. *Res. High. Educ.* 59, 382–400. doi: 10.1007/s11162-017-9473-z

Montoya Gutiérerz, G. (2016). *Estudio Factores Asociados Al Abandono Temprano De La Educación Superior. Congresos CLABES*. Available online at: https://revistas.utp.ac.pa/index.php/clabes/article/view/1055 [Nov 3, 2016].

Parrino, C. (2010). *Deserción en el primer año universitario . Dificultades y logros. X Coloquio Internacional Sobre Gestión Universitaria En América Del Sur*. 1–19. https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/96620/PARRINO.pdf

Sandoval, I., Sánchez, T., Naranjo, D., and Jiménez, A. (2019). "Proposal of a mathematics pilot program for engineering students from vulnerable groups of Escuela politécnica Nacional," in *Proceedings of the LACCEI International Multi-Conference for Engineering, Education and Technology, 2019-July(August)*, (Boca Raton: LACCEI), doi: 10.18687/LACCEI2019.1.1.387

Sandoval, I., Sánchez, T., Velasteguí, V., and Naranjo, D. (2018). Factores Asociados Al Abandono En Estudiantes De Grupos Vulnerables. Caso Escuela Politécnica Nacional. *Congr. CLABES* 2018, 132–141.

Teoh, E. J., Tan, K. C., and Xiang, C. (2006). Estimating the number of hidden neurons in a feedforward network using the singular value decomposition. *IEEE Trans. Neur. Net.* 17, 1623–1629. doi: 10.1109/TNN.2006.880582

Timoshenko, J., Anspoks, A., Cintins, A., Kuzmin, A., Purans, J., and Frenkel, A. I. (2018). Neural Network Approach for Characterizing Structural Transformations by X-Ray Absorption Fine Structure Spectroscopy. *Phys. Rev. Lett.* 120:225502. doi: 10.1103/PhysRevLett.120.225502

Vandamme, J. P., Meskens, N., and Superby, −F. (2007). Predicting Academic Performance by Data Mining Methods. *Educ. Econ.* 15, 405–419. doi: 10.1080/09645290701409939

Yang, F., and Li, F. W. B. (2018). Study on student performance estimation, student progress analysis, and student potential prediction based on data mining. *Comput. Educ.* 123, 97–108. doi: 10.1016/j.compedu.2018.04.006

Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., et al. (2018). Graph Neural Networks: A Review of Methods and Applications. *arXiv* 22:1812.08434.

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.

Copyright © 2020 Sandoval-Palis, Naranjo, Gilar-Corbi and Pozo-Rico. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership