Technical University of Munich
TUM School of Social Sciences and Technology

Gender Differences in STEM Motivational Outcomes:

Exploring the Role of School-Based Interventions and Teacher Constructive Support

Kaley Lesperance

Complete reprint of the dissertation approved by the TUM School of Social Sciences and Technology of the Technical University of Munich for the award of the

Doktorin der Philosophie (Dr. phil).

Chair: Prof. Dr. phil. Doris Lewalter

Examiners:

1. Prof. Dr. phil. Doris Holzberger
2. Prof. Dr. Frank Reinhold

The dissertation was submitted to the Technical University of Munich on 05.09.2023 and accepted by the TUM School of Social Sciences and Technology on 19.11.2023.

Acknowledgments

I want to express my sincere gratitude to everyone who supported me during my dissertation and doctoral journey.

Firstly, I would like to thank my doctoral supervisor, Prof. Dr. Doris Holzberger, for her continued support since the beginning of my doctoral studies. I am supremely grateful for her guidance, expertise, motivation, and belief in my capabilities. My skills as a researcher have grown immensely under her mentorship, and I am so glad to have had this experience.

Next, I would like to thank my fellow co-authors on my papers, namely Prof. Dr. Jan Retelsdorf and Prof. Dr. Jasmin Decristan, for their help and feedback. I feel lucky to have had the opportunity to learn from their expertise and experience. I would especially like to thank Prof. Dr. Sarah Hofer not only for being a co-author, but also for being my mentor, and always being available to support and offer advice.

Many thanks also to Prof. Dr. Frank Reinhold, for taking the time to grade my dissertation, and to Prof. Dr. Doris Lewalter, for chairing my grading committee.

I want to thank all of my colleagues on the SUF team at the TUM: Anna Hartl, Janina Täschner, Lisa Ziernwald, Yvonne Holzmeier, Simon Munk, and Sara Moore. Working with you has been a pleasure, and I feel lucky to have such supportive and kind-hearted colleagues. A special thanks to Janina Täschner for her valuable feedback on my dissertation and for being such a great office mate. I would also like to thank the student research assistants, especially Aiymbubu Dzhaparkulova and Olga Viri Huerta Salinas, for their impressive dedication and help with my projects.

Importantly, I would also like to thank the people outside of academia, without whom I could not have gotten through the last five years. A huge thank you to my friends Erik, Queralt, Lana, Cafer, Lenn, Almir, and the rest of the Burger Group, for always making me laugh and sharing such fun times. An extra special thank you to Raven Rinas, for sharing this Ph.D. experience with me and being a shoulder to lean on through this process.

A heartfelt thank you to my partner, Niklas, for being the most supportive and patient partner I could have imagined, and always making sure I knew how much he believed in me.

Finally, I would like to thank my family. To my sisters, Kristine and Kourtney, thank you for being my best friends and keeping me sane when I felt overwhelmed. Thank you to my parents for always supporting me, letting me know you are proud of me no matter what, listening to me talk about my challenges with the dissertation (even if they did completely understand what I was talking about), and always encouraging me, even from a distance. Lastly, thank you to my grandparents, Boma and Bompa, for being so interested in and supportive of my work.

Included Publications

This dissertation is cumulative, and includes two research studies. Both studies underwent peer-reviews in the respective international journals. The first study has already been accepted and published, while the second study has been accepted and is currently in press. The author of this dissertation is the first author of both papers and was responsible for the study design, development of theoretical content, data preparation and analysis, and coordination of the entire publication process.

The first study was a meta-analysis of the gender-specific effects of school-based interventions:

Lesperance, K., Hofer, S., Retelsdorf, J., & Holzberger, D. (2022). Reducing gender differences in student motivational-affective factors: A meta-analysis of school-based interventions. *British Journal of Educational Psychology*, *92*(4), 1502-1536.

Alongside the publication process, the first author was responsible for preparing the manuscript, including literature review, data collection and preparation, and analysis, and writing the original draft of the paper (65%). The co-authors, Prof. Dr. Sarah Hofer (15%), Prof. Dr. Doris Holzberger (15%), and Prof. Dr. Jan Retelsdorf (5%), contributed to the improvement and revision of the study by providing feedback on the theory, methodology, analyses, and interpretation of results.

The second study investigated the moderating role of teacher constructive support for the relationship between gender and motivational outcomes:

Lesperance, K., Decristan, J., & Holzberger, D. (2023). The role of teacher constructive support for gender differences in motivational outcomes in secondary school mathematics. *International Journal of Gender, Science and Technology, 15*(3).

Besides overseeing the publication process, the first author was responsible for the conceptualization of the study, development of the theoretical background, data preparation and analysis, and writing the initial manuscript draft (70%). The co-authors, Prof. Dr. Jasmin Decristan (15%) and Prof. Dr. Doris Holzberger (15%) contributed to the paper by providing support and feedback regarding theoretical underpinnings, refinement of the research questions, methodological issues, and critical review of the manuscript.

Abstract

Persistent gender differences exist in student motivational outcomes in science, technology, engineering, and mathematics (STEM) fields, with female students often exhibiting lower levels of these outcomes compared to their male counterparts. This may result in fewer females pursuing STEM academically or professionally. Based on two empirical studies, this dissertation examines two key areas that could play a role in addressing these differences in primary and secondary school: school-based interventions and teacher constructive support. School-based interventions specifically target motivational outcomes, while teacher constructive support is a daily academic influence on student motivation. A joint exploration of these two areas offers a holistic perspective on confronting gender differences in motivational outcomes.

The first study (Study I) used meta-analysis to investigate whether school-based interventions in STEM subjects have higher effects for females than males on student motivational outcomes in primary and secondary school. A systematic database search yielded 71 effect sizes from 20 studies. Effect sizes were aggregated and analyzed for gender-specific effects, with context factors such as subject, school level, intervention target, and intervention strategy included as potential moderators. While results showed significant positive effects for students overall, no significant differences were found between males and females overall. Moderator analyses showed a significant impact of interventions using psycho-social strategies and methods specifically targeting females, but not for the school level or subject. This research suggests that while school-based interventions have positive effects on motivational outcomes for all students, no gender-specific overall effects are currently evident. The need for more targeted approaches to reducing gender differences using school-based interventions is discussed.

The second study (Study II) explored the moderating role of teacher constructive support on the relationship between student gender and motivational outcomes in mathematics lessons. Self-report data from a sample of $N = 1{,}116$ German students in lower secondary mathematics classrooms was used to analyze student self-concept, self-efficacy, and interest in mathematics lessons, as well as student perceptions of constructive support, separated into the facets of instructional support and social-emotional support. Results showed that females reported significantly lower levels of motivational outcomes than males. The latent interaction models revealed that while teacher constructive support positively affected students' motivational outcomes, it did not significantly moderate the relationship between gender and motivational outcomes. The results underscore the importance of teacher constructive support for student motivational outcomes. Additionally, they highlight the need for further research into theoretically meaningful moderators of the relationship between gender and motivational outcomes in STEM subjects.

Overall, this dissertation contributes to research aiming to reduce gender differences in STEM motivational outcomes. It sheds light on how both targeted interventions and the fundamental role of constructive support may be relevant for gender differences in motivational outcomes. It aims to offer directions for future research on how schools can help close the motivational gender gap in STEM subjects.

Zusammenfassung

In den Fächern Mathematik, Informatik, Naturwissenschaften und Technik (MINT) zeigen Schülerinnen oft niedrigere motivationale Merkmale als Schüler. Dies kann dazu führen, dass weniger Frauen MINT-Fächer akademisch verfolgen und auch in beruflichen Kontexten unterrepräsentiert sind. Anhand zweier empirischer Studien werden in dieser Dissertation zwei wichtige Aspekte untersucht, die bei der Reduktion dieser Geschlechterunterschiede eine Rolle spielen können: schulbasierte Interventionen und konstruktive Unterstützung der Lehrkräfte. Schulbasierte Interventionen wirken gezielt und zielen auf spezifische motivationale Merkmale ab, während konstruktive Unterstützung durch Lehrkräfte die Schülermotivation unspezifischer und über einen längeren Zeitraum beeinflussen kann. Die Untersuchung beider Aspekte bietet eine ganzheitliche Perspektive auf die Bewältigung von Geschlechterunterschieden bei motivationalen Merkmalen.

Die erste Studie (Studie I) untersuchte im Rahmen einer Meta-Analyse, ob schulbasierte Interventionen geschlechtsspezifische Auswirkungen auf motivationale Schülermerkmale in MINT-Fächern in Grund- und Sekundarschulen haben. Nach einer systematischen Datenbankrecherche wurden 71 Effektgrößen aus 20 Studien einbezogen, wobei Kontextfaktoren wie Fach, Schulstufe, Interventionsziel und Interventionsstrategie zusätzlich als potenzielle Moderatoren analysiert wurden. Die Ergebnisse zeigten insgesamt positive Effekte für alle Schüler*innen. Es gab jedoch insgesamt keine signifikanten geschlechtsspezifischen Unterschiede. Moderatoranalysen zeigten eine signifikante Wirkung von Interventionen, die psychosoziale Strategien und Methoden verwenden, die speziell auf Mädchen abzielen. Für Schulstufe oder Fach wurden hingegen keine Effekte gefunden. Die Notwendigkeit zielgerichteter Ansätze zur Verringerung von Geschlechterunterschieden durch schulbasierte Interventionen wird erörtert.

Die zweite Studie (Studie II) untersuchte die moderierende Rolle konstruktiver Unterstützung der Lehrkraft im Zusammenhang zwischen Geschlecht und motivationalen Schülermerkmalen. Selbstberichtsdaten von $N = 1.116$ deutschen Sekundarschüler*innen wurden analysiert. Untersucht wurden Selbstkonzept, Selbstwirksamkeit und Interesse am Mathematikunterricht sowie die Schülerwahrnehmungen der konstruktiven Unterstützung. Letztere wurde in fachlich-inhaltliche und sozial-emotionale Facetten differenziert. Die Ergebnisse zeigten, dass Schülerinnen niedrigere motivationale Merkmale angaben als Schüler. Die latente Interaktions-Strukturgleichungsmodelle ergaben, dass konstruktive Unterstützung durch Lehrkräfte zwar positive direkte Effekte auf die motivationalen Merkmale der Schüler*innen hatte, jedoch die Geschlechter-Motivations-Beziehung nicht signifikant moderierte. Implikationen hinsichtlich der Bedeutung der konstruktiven Unterstützung für motivationale Schülermerkmale sowie der Bedarf an weiterer Forschung zu theoretisch bedeutsamen Moderatoren der Beziehung zwischen Geschlecht und Motivation werden diskutiert.

Diese Dissertation bietet Einblicke in die Verringerung von Geschlechterunterschieden bei motivationalen Merkmalen in MINT-Fächern und betont sowohl gezielte Interventionen als auch die Rolle der konstruktiven Unterstützung. Ziel ist es, Wege aufzuzeigen, wie Schulen die Geschlechterunterschiede in MINT-Fächern hinsichtlich der Motivation überbrücken können.

Table of Contents

1 Introduction

"Strengthening women's participation in Science, Technology, Engineering and Mathematics (STEM) fields is not only a matter of equal opportunities and social justice, but also crucial to meet pressing societal challenges like the twin green and digital transitions." (European Commission, 2022, p. 1)

The gender[1] gap in science, technology, engineering, and mathematics (STEM) domains is a persistent issue that, despite various efforts from policymakers and educational institutions, continues to exist in many countries worldwide (World Economic Forum, 2023). In the current labor market, technological advancement and digitalization have dramatically increased the need for skilled workers in STEM sectors (Fatourou et al., 2019). However, in the European Union specifically, women account for only 41% of total employment in science and technology careers, and only 35% of higher graduates in STEM-related disciplines are women (European Commission, 2021). This issue seems to begin before individuals enter higher education or the job market. Research shows that, on average, females consistently display lower levels of STEM motivation and career aspirations than males, a trend that can already be seen during the school years (OECD, 2013, 2016). When viewed in light of the above quote, this trend is problematic for both individual students, who may not be reaching their full potential, and for society as a whole. Therefore, identifying strategies to address the problem of gender differences in STEM and evaluating their impact is crucial for reducing these differences.

When examining gender differences in STEM subjects, research has typically looked at performance-related outcomes, such as exam scores or aptitude tests (Cimpian et al., 2020). However, many studies indicate that male and female students hardly differ in STEM performance outcomes (Bloodhart et al., 2020; Hyde & Linn, 2006; Reilly et al., 2019; Stoet & Geary, 2018). Where research does find female students consistently lag behind their male peers in STEM subjects is in regard to their motivational outcomes (Meece et al., 2006). Motivational outcomes refer to a variety of constructs connected to student motivation and engagement in the classroom, including constructs such as interest, intrinsic and extrinsic motivation, self-concept, self-efficacy, and enjoyment (P. K. Murphy & Alexander, 2000). These motivational outcomes are extremely important for students' academic trajectories, as research shows that they are positively linked to future educational choices, and play a significant role in what careers students will choose to pursue later in life (Kim & Pekrun, 2014; Pintrich, 2003b). In STEM subjects such as science and mathematics, motivational-affective outcomes are especially influential for whether or not students decide to pursue studies and careers in those subjects (Rosenzweig & Wigfield, 2016). It is therefore concerning that gender differences consistently emerge in these outcomes in STEM subjects. These gender differences arise as early as primary school (Eccles et al., 1993; Master et al., 2021). In science and math, for example, female students tend to display lower self-efficacy,

[1] The current dissertation classifies gender according to a binary conceptualization based on biological sex (male and female). In the existing literature on gender differences in STEM, this is the primary conceptualization used, and therefore necessary and appropriate for the current dissertation (European Commission, 2021). However, it is also important to state that other conceptualizations of gender exist, including non-binary conceptualizations, and these should be acknowledged and supported. Gender identity is a complex, individual experience and may or may not correspond to sex assigned at birth (European Institute for Gender Equality, 2021).

more negative perceptions, and less value for these subjects than males (Else-Quest et al., 2010; P.H. Miller et al., 2006). This trend seems to intensify and become more stable as students get older (Jacobs et al., 2002) and persists well into higher education (Dasgupta & Stout, 2014).

Researchers have explored why these differences between male and female students' motivational outcomes arise in STEM subjects. Gender differences in various motivational outcomes emerge as individuals develop and engage with their social environment (M.-T. Wang & Degol, 2013). Gender stereotypes, which are fixed beliefs about aptitudes and characteristics attributed to males and females (Eagly & Steffen, 1984), are acquired by children from significant social actors in their lives, such as parents, teachers, and peers (Smith & Farkas, 2022; Tiedemann, 2000). There are various proposed mechanisms of how individuals learn and acquire these gender stereotypes, such as through model learning, reinforcement of gender-conventional behavior, differential treatment of boys and girls, or explicit expression of gender-stereotypical expectations (Gunderson et al., 2012; Heyder et al., 2019). Once acquired, these stereotypes can influence how individuals process and classify information, as well as their behavior, choices, and attitudes (C. L. Martin & Halverson, 1981). One of the most widely cited and empirically supported explanations for how these gender stereotypes relate to motivational outcomes in education is Eccles' expectancy-value theory of motivation. The expectancy-value theory postulates two main factors contributing to an individual's motivation for engaging and persisting in specific tasks: expectancy for success in the task and perceived value of the task (Eccles & Wigfield, 2020). Gender stereotypes influence these expectancy-value beliefs, which can in turn affect females' motivational outcomes and lead to gender disparities in certain areas, such as STEM (Eccles et al., 1993).

While not a new issue, gender differences in STEM motivational outcomes have garnered attention from politicians, educators, and researchers in recent years (Encinas-Martín & Cherian, 2023; European Commission, 2022). The primary concerns are understanding why these differences arise and investigating what can be done to reduce them (Chavatzia, 2017). Research on reducing gender differences in specific fields has been approached from many perspectives. Various strategies have been assessed at different levels in the educational research domain, ranging from large-scale curriculum reforms to specific intervention programs that target student motivational outcomes (Kong et al., 2020; Roberts, 2014). Much of this research emphasizes classroom-level strategies or changes that teachers can initiate and implement (van den Hurk et al., 2019). As students spend most of their academic hours in classrooms, these settings are essential for addressing gender differences. In this context, two significant gaps emerge from the existing research. First, there are various primary studies that evaluate school-based interventions targeting motivational outcomes in STEM overall (Wigfield & Wentzel, 2007). However, this pool of primary studies alone does not allow for a comprehensive assessment of which interventions are most effective in addressing gender differences in STEM motivational outcomes. In order to both further intervention research and provide evidence-based recommendations for educators, a clear understanding of the most effective intervention strategies is necessary. Second, while targeted interventions are important to assess, the intrinsic classroom dynamics that also impact student motivation are often overlooked. Though less overt, these factors might also play crucial roles in the gender differences observed in STEM motivational outcomes. Deepening the understanding of this relationship would highlight the significance of constructive support from educators, particularly in potentially enhancing the motivational outcomes of female students in STEM subjects. This dissertation precisely addresses these gaps, focusing on the potential of school-based interventions and the role of teacher constructive support, which is a crucial aspect of classroom dynamics in regard to student motivational outcomes.

School-based interventions are novel methods that differ from regular instruction, are used in the classroom or other school contexts, and can vary in terms of strategies, length, and implementation (Lin-Siegler et al., 2016). Numerous primary studies have shown that various school-based interventions can effectively promote motivational and affective outcomes in students on a general level (Durlak et al., 2011; Lazowski & Hulleman, 2016). Interventions have also been explicitly evaluated in the context of targeting gender differences in STEM motivational outcomes (Rosenzweig & Wigfield, 2016), but the results have been ambiguous. An unresolved question remains concerning the overall gender-specific effects of these interventions in STEM subjects, especially concerning their efficacy for females, who commonly report lower motivational outcomes in these subjects. In order to gain a clear understanding of which intervention strategies, if any, are most effective in reducing gender differences in student motivational outcomes, a systematic aggregation and analysis of these studies is essential. However, to date, such a thorough review remains notably absent from the literature.

Although school-based interventions offer a promising avenue, they often overlook the persistent, intrinsic aspects of students' classroom experiences that can impact motivational outcomes. Many factors contribute to the overall classroom environment; however, among these, teacher constructive support stands out as one of the most influential for student motivational outcomes (Cornelius-White, 2007). Constructive support encompasses not only assistance with learning and content but also emotional and personal encouragement (Praetorius et al., 2014). Constructive support has been shown to have positive effects on various student motivational-affective factors, including interest, self-concept, and self-efficacy (Fauth et al., 2014). Some studies have also found that constructive support especially benefits certain disadvantaged groups (Curby et al., 2009; Decristan et al., 2022). Despite this evidence, the role of constructive support in the intersection of gender and STEM motivational outcomes remains underinvestigated. Scholars have emphasized the need for a nuanced exploration of the relationship between student motivational outcomes, constructive support from teachers, and gender differences (Rueger et al., 2008).

Taken together, the main research goal of the current dissertation was to fill the aforementioned research gaps and thereby provide a well-rounded investigation of potential avenues for reducing gender differences in student motivational outcomes in STEM. To this aim, two empirical studies were conducted to evaluate the effectiveness of both novel, targeted strategies (school-based interventions) and influential aspects of everyday classroom experiences (teacher constructive support). The first study (Study I) used meta-analysis to evaluate school-based interventions in STEM subjects and whether they had differing effects on male and female students' motivational outcomes. Study I also examined whether school level, subject, intervention strategy, and other contextual factors might play a role in the magnitude of these intervention effects. The second study (Study II) aimed to evaluate whether constructive support moderated the effect of gender on student motivational outcomes in mathematics and the direct effects of both gender and constructive support on student motivational outcomes. In the following chapters, fundamental theoretical perspectives and prior empirical research on gender differences in STEM motivational outcomes, school-based interventions, and constructive support are presented. This is followed by an overview of the present research questions, extensive summaries of the included empirical studies, and a general discussion of the findings, strengths, limitations, and implications of the current research project.

2 Theoretical Background

The following chapters discuss the research background on gender differences in STEM motivational outcomes. First, in the current chapter (Chapter 2), theoretical definitions of motivational outcomes, as considered in this dissertation, are outlined (Section 2.1). Next, research on the role motivational outcomes play in student choices and achievement in STEM subjects is presented (Section 2.2). Gender differences in STEM motivational outcomes are then discussed, including current evidence for gender differences and empirical perspectives on the root of these gender differences (Section 2.3). The next chapter (Chapter 3) then examines research on reducing these gender differences in STEM motivational outcomes, with an emphasis on school-based interventions (Section 3.1) and teacher constructive support (Section 3.2), which are the focus of the current dissertation.

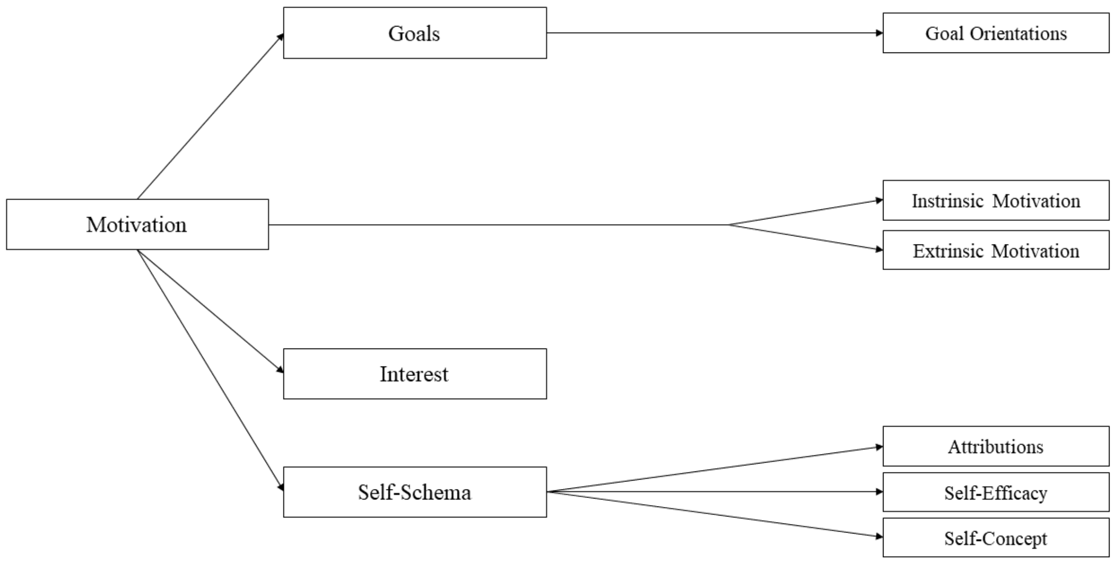
2.1 Motivational Outcomes in STEM Education

2.1.2 Theoretical Definitions of Motivational Outcomes in Education

The construct of motivational outcomes represents a complex and multifaceted topic within educational research. As a general term, motivation encompasses a broad spectrum of student outcomes and is conceptualized within a range of theoretical frameworks. The extensive body of research dedicated to academic motivation has given rise to diverse models that define and categorize its fundamental components and associated factors. In a review of terminologies and concepts pertaining to academic motivation, Murphy and Alexander (2000) used existing theories to identify four overarching categories and definitions of motivational constructs. The first category is motivation itself, which is described as the mechanisms or processes responsible for guiding, energizing, and sustaining actions. Motivation can be further separated into intrinsic motivation, which is the drive to pursue a task because it is inherently rewarding, and extrinsic motivation, which is the drive to pursue a task to obtain external rewards or fulfill external expectations. The second category is goals, which are what students want to achieve and can be broken down into goal orientations, which refer to a set of behavioral intentions that dictate the ways in which students undertake and engage in learning activities. The third motivational category is interest, which is an individual's positive, energized orientation toward an activity or task. The last category is self-schema, defined as an individual's personal understanding of themselves that characterizes how they perceive and engage with certain tasks and activities. Some specific examples of self-schemas are self-efficacy, which is an individual's beliefs regarding their ability to learn and perform specific tasks, and attributions, which are causal connections that students make to explain why successes or failures occurred. Self-concept, while not featured in the review by Murphy and Alexander, is another important form of self-schema and can be defined as a person's perception of themselves, including both cognitive and affective aspects (Bong & Clark, 1999). An adapted visualization of these motivational outcomes, as conceptualized by Murphy and Alexander (2000), can be seen in Figure 1.

Figure 1

Adapted overview of motivational outcomes as conceptualized by Murphy and Alexander (2000)



Although not exhaustive, this overview provides a comprehensive list of some of the most important constructs associated with motivation in academic contexts across various theoretical models of motivation (Cook & Artino Jr., 2016). However, some motivation researchers argue that many of these models neglect to consider that motivation also includes an affective or emotional component (Pintrich, 2003b). Students experience various emotional responses in educational contexts that function as a component of their motivation (Pekrun & Linnenbrink-Garcia, 2012). These emotional components can therefore also be encompassed under the umbrella of motivational outcomes. Some emotional components that are especially relevant to academic motivation include enjoyment, anxiety, or boredom regarding a particular task or activity (Pekrun et al., 2002). These motivational outcomes share several common characteristics. Research has shown that motivational outcomes are malleable in educational contexts, meaning they can change throughout a student's academic trajectory (Heckman & Kautz, 2013). As individuals develop, motivational outcomes naturally change. When students progress through their education, the personal factors and drivers that motivate them may evolve (Wigfield et al., 2006). Additionally, changes in the learning environment and interactions with others, such as teachers, parents, or peers, in educational contexts can also lead to natural changes in motivational outcomes (Wigfield & Wagner, 2005). Motivational outcomes can also be deliberately changed in order to promote or foster positive levels of motivation (Hidi & Harackiewicz, 2000). Another characteristic to note is that while certain motivational outcomes, for example, academic self-concept, can also be conceptualized as generic measures stable across various contexts when assessing motivational outcomes regarding specific subjects, it is essential to contextualize them within that domain (Michel et al., 2022). Motivational outcomes often pertain specifically to the academic subject in question, such as reading or mathematics (Hornstra et al., 2016). Students can, and often do, have different levels of motivational outcomes in different subjects, as these outcomes are tightly connected to the activity in question (Green et al., 2007; Marsh et al., 2001). Therefore, it is very common for a student to have high levels of motivational outcomes in one subject, but low levels in a different subject.

2.1.2 The Role of Motivational Outcomes in STEM

In assessing the factors that predict students' intentions toward studying a particular subject or pursuing a specific career, motivational outcomes emerge as highly influential factors in this process (Rottinghaus et al., 2002; Yeung & McInerney, 2005). Empirical studies suggest that students with strong interest, perceived competence, and motivation in a subject are more inclined to pursue it compared to peers with lower levels of these outcomes (Bandura et al., 2001; Parker et al., 2014). Various motivational outcomes have been shown to predict career aspirations and study choice, including self-concept (Korhonen et al., 2016), self-efficacy (Tang et al., 2008), interest (Päßler & Hell, 2012), and perceived value of the subject (Eccles & Wang, 2016). While this trend is pervasive across numerous fields of study, it is especially pronounced in the STEM domain. Empirical findings have consistently underscored that the decision-making process related to engagement in STEM disciplines is not a spontaneous choice. Rather, deciding to participate in STEM fields is a longitudinal process that begins before students enter university (Almeda & Baker, 2020; Maltese & Tai, 2011). While various factors influence the decision to pursue STEM education, such as familial influences, cultural backgrounds, and school effects, motivational outcomes are particularly salient in this complex decision-making process (M.-T. Wang & Degol, 2013). A wide range of motivational outcomes are strongly associated with students' STEM career aspirations and intent to study a STEM field after secondary school (Aschbacher et al., 2014; Lauermann et al., 2017; Watt et al., 2012). For example, Ahmed and Mudrey (2019) found that in a sample of 5,611 fifteen-year-old students, self-concept, enjoyment, and value in science all strongly and significantly predicted STEM career aspirations, even after controlling for socioeconomic status and prior science achievement. Blotnicky and colleagues (2018) also found that students in Grades 7 and 9 with higher mathematics self-efficacy and interest in technical and scientific activities were more likely to consider a career in STEM. From as young as eight years old, students' perceived value and self-beliefs in STEM subjects already predict their study choice and future career path (van Tuijl & van der Molen, 2016). These motivational outcomes remain relevant throughout students' academic trajectories in STEM fields. Research demonstrates that students with sustained higher levels of motivational outcomes are not only more likely to pursue STEM, but are also more inclined to persist and complete their STEM degrees (Cromley et al., 2015; Mau, 2003; Musu-Gillette et al., 2015). Given these compelling findings, it is imperative for educators and policymakers to prioritize and foster these motivational outcomes in students.

2.1.3 Summary

To summarize, motivational outcomes in educational contexts encompass a diverse range of constructs, including intrinsic and extrinsic motivation, goal orientations, interest, self-schema, and affective components. These motivational outcomes are vital factors for understanding students' choices and behaviors, offering insights into their persistence in specific domains and guiding their future career trajectories (Guo et al., 2015). Research indicates that higher levels of these motivational outcomes are pivotal in STEM subjects, as they substantially contribute to students' sustained engagement in STEM studies and impact their ultimate decision to pursue STEM degrees and careers (Wille et al., 2020). Consequently, understanding the disparities and nuances in these motivational outcomes between students within the STEM domain is crucial for research on gender equality in STEM subjects.

2.2 Gender Differences in STEM Motivation

2.2.1 Past and Current Research on Gender Differences in STEM Motivation

While it is evident that motivational outcomes play a significant role in shaping students' engagement and choices in STEM, it is equally important to understand how these motivational outcomes differ based on gender and how such differences might account for discrepancies in the representation of males and females in STEM fields. The issue of gender differences in STEM fields came to light in research as early as the 1970s, with researchers pointing out a trend of discrepancies between males and females in certain fields (Cole & Cole, 1973; Epstein, 1970). These early studies noted that there seemed to be a pattern of fewer females in proportion to males pursuing careers in areas such as engineering, medicine, and the natural sciences. Since then, policymakers, educators, and researchers have dedicated a great deal of effort to understanding why these discrepancies exist. Early on, some academics argued that these discrepancies might be due to biologically inherent differences in male and female cognitive abilities in mathematics and physical science (Benbow, 1988; Bock & Kolakowski, 1973; Thomas & Kail, 1991). Modern research has primarily debunked this notion, consistently showing negligible performance differences between males and females in subjects like mathematics and science. In the latest PISA findings from 2018, from a sample of 79 countries, boys only marginally surpassed girls in average mathematics performance, while in science, girls marginally exceeded the average performance of boys (OECD, 2020a). Numerous research syntheses have also demonstrated that males and females, on average, exhibit comparable levels of achievement in STEM subjects such as mathematics and science (Else-Quest et al., 2010; Hyde et al., 2008; S.M. Lindberg et al., 2010; O'Dea et al., 2018; Voyer & Voyer, 2014). From the presented evidence, it becomes clear that the underrepresentation of females in STEM cannot merely be attributed to differences in performance or ability. Consequently, recent research has focused on non-cognitive factors, such as motivational outcomes, that might better explain why fewer females than males choose to pursue STEM studies, despite their performance being equal to or even better than their male peers. Further research revealed significant differences in motivational outcomes, with female students consistently displaying lower levels of motivational outcomes, on average, than their male counterparts in STEM subjects (M.-T. Wang & Degol, 2017). Given the importance of student motivational outcomes for achievement and persistence in STEM education (Simon et al., 2015), these differences in motivational outcomes provide a more plausible explanation for the gender gap in STEM areas.

Research has shown that starting in primary school, boys and girls already begin to display some differences in many motivational outcomes in subjects such as mathematics and science, two core subjects in primary school (Conlon et al., 2023; McGuire et al., 2022). Regarding mathematics, by early to mid-primary school, girls display lower levels of mathematics self-concept and higher levels of mathematics anxiety than boys (Levine & Pantoja, 2021; Lindberg et al., 2013). As early as first grade, girls report more negative attitudes towards mathematics, less enjoyment, and lower competency self-beliefs in mathematics (Cvencek et al., 2021; Eccles et al., 1993). Female students in primary school have been shown to rate their mathematics capabilities lower than boys, even when their mathematics achievement was equal to boys (Herbert & Stipek, 2005; Heyder et al., 2019). Girls between 8 and 9 years of age are also less likely to attribute success in mathematics to their own high ability (Dickhäuser & Meyer, 2006) and are generally less likely to report thinking that they excel in mathematics (Freedman-Doan et al., 2000). This trend continues into secondary school. Results from large-scale, international studies such as PISA and TIMSS show that, on average, female students between ages 14 and 16 report lower self-efficacy in mathematics and less perceived value in mathematics as a subject (Else-Quest et al., 2010; Ghasemi & Burley, 2019; OECD, 2013). Various studies have found that girls in secondary school consistently display less interest, confidence, enjoyment, and fewer ambitions for mathematics than boys (Ganley & Lubienski, 2016; Guo et al., 2015; Preckel et al., 2008; Skaalvik et al., 2015).

Gender differences in motivational outcomes are also evident when comparing male and female students in science subjects. Results from PISA 2015, which emphasized science assessment, revealed that across all 72 participating countries, male students, on average, displayed a higher intrinsic motivation for science than females (OECD, 2016). The same results also showed that male students expressed greater interest and enjoyment in science compared to their female counterparts. In high school, girls' value and interest in science have been shown to be lower than boys' (Shumow & Schmidt, 2013). Girls are also less likely to aspire to science careers or view science as an important subject (Lupart et al., 2004; Riegle-Crumb et al., 2012; Schoon & Eccles, 2014). As students navigate through their secondary education and are presented with opportunities to participate in more specialized science courses, gender differences in certain disciplines become more apparent. Female students consistently report less interest, motivation, and future study aspirations in physics (Adams et al., 2006; Buccheri et al., 2011; Card & Payne, 2021; Heilbronner, 2012; Jansen et al., 2014) and computer science (Beyer, 2014; Kadijevich, 2000; Master et al., 2021). However, the one exception to this pattern is in biology, where gender differences in motivational outcomes are not as pronounced. Some studies suggest that girls exhibit more interest, motivation, and self-efficacy in this subject than boys (Baram-Tsabari & Yarden, 2011; P.H. Miller et al., 2006; Patall et al., 2018). This has been attributed to the association of biology and life sciences with health-related careers that emphasize helping others and caregiving, tasks traditionally viewed as characteristic of females (Su & Rounds, 2015).

2.2.2 Understanding Gender Differences in STEM Motivation

Given the substantial body of evidence highlighting gender differences in student motivational outcomes within STEM disciplines, recent research endeavors in this field have focused primarily on elucidating the origins of these differences in motivation. Even though males and females display comparable achievement in science and mathematics tasks, differences in motivational outcomes still persist in these subjects. Motivational outcomes are multifaceted and develop in relation to an interplay of various social, cultural, and psychological determinants (Pintrich, 2003a). This is equally true when it comes to gender differences in these outcomes. Wang and Degol (2017) conducted a comprehensive review of explanations offered by researchers in recent decades, spanning disciplines such as psychology, sociology, and economics, to account for the observed gender differences in STEM motivation outcomes. Their conclusions highlighted that, amidst the myriad of research perspectives, sociocultural factors consistently emerged as significant contributors to these gender differences in motivational outcomes. Foremost among these sociocultural factors are gender stereotypes. Gender stereotypes are defined as “structured sets of beliefs about the personal attributes of men and women” (Ashmore & Del Boca, 1979, p. 222). Gender stereotypes encompass preconceptions about the interests, roles, and competencies of males and females. Within the realm of education, these stereotypes predominantly reflect perceptions regarding the innate capabilities and aptitudes of boys and girls (Bian, 2022). Traditional stereotypes about gender abilities also extend to STEM domains, such as the belief that boys are better than girls at mathematics or that men are more interested in science than women (Master, 2021). Research indicates that children as young as six years old already internalize and express beliefs supporting these traditional gender stereotypes in areas such as mathematics, science, and technology (Cvencek et al., 2011; Kurtz-Costes et al., 2008; Master et al., 2021; D.I. Miller et al., 2018). Children start absorbing gender stereotypes early on, drawing primarily from influential agents in their social environment, such as parents, peers, and teachers (Kollmayer et al., 2018). Various explanations exist for the mechanisms through which children acquire these gender stereotypes from their social environment, such as through the reinforcement of gender-stereotypical behavior, model learning, differential treatment of males and females, or communication of gendered expectations (Gunderson et al., 2012; Heyder et al., 2019). For example, some researchers have found that parents and teachers judge girls’ mathematics ability as lower than boys’, even when they have received similar grades (Eccles, 2015; Lubienski et al., 2013; Tiedemann, 2000). Parents have also been shown to encourage boys more than girls to take mathematics and science courses (Tenenbaum, 2009), and to engage in more discussions about science and mathematics with boys (Crowley et al., 2001). Beyond the personal social environment, gender stereotypes are also very present in broader society and the media, and children can learn them from these sources as well (Ellemers, 2018; Ward & Grower, 2020). Media portrayals of STEM professionals often endorse traditional gender stereotypes and can also influence students’ perceptions of STEM (Olsson & Martiny, 2018; Steinke, 2017).

These gender stereotypes, once learned, can affect how individuals process and categorize information and can also influence their choices, behaviors, and beliefs. Eccles’ (2000) expectancy-value theory provides a widely supported explanation for how these learned gender stereotypes can lead to differences in the decisions individuals make and the ways in which they perceive themselves. The expectancy-value theory proposes that whether or not someone engages in a particular task or pursues a certain goal is determined by their expectations for success and their perceived value for that task or goal (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000). In other words, individuals are more likely to feel motivated for tasks they believe they can succeed in and consider important or meaningful. In this context, gender stereotypes have the ability to influence not only the perceived value of a task for an individual but also their self-beliefs, attitude, and perceived competence in that task, which in turn impacts their expectations of success (Eccles, 1994). In the case of STEM subjects, gender stereotypes that STEM is a male domain create a perception that males are more capable in these subjects and that STEM subjects are more relevant for males (M.-T. Wang & Degol, 2013). These experiences and stereotypes that an individual has involving

STEM-related activities, combined with their own gender identity, will therefore shape their self-beliefs, interest, perceived value, and, ultimately, long-term aspirations for those activities (Eccles et al., 1993). This will consequently lead to differences in students' decisions regarding pursuing studies or careers in STEM fields (Jacobs et al., 2005). If a female child grows up learning from her family, teachers, and broader social environment that math, science, and other STEM subjects are masculine domains, she is likely to perceive herself as less competent and have a lower value for those subjects throughout her academic career, leading to lower motivational outcomes, and subsequently lower aspirations for STEM (Eccles et al., 1993; Meece et al., 2006; M.-T. Wang et al., 2015).

2.2.3 Summary

The previous section thoroughly examines the subject of gender differences in STEM motivation. It explains early ideas of inherent cognitive differences between males and females while presenting modern evidence that these differences stem not from discrepancies in ability but rather from differences in motivational outcomes. From primary school through secondary education, girls consistently show lower levels of motivational outcomes in STEM subjects than boys (M.-T. Wang & Degol, 2013). Attempting to understand the development of these motivational differences has been a key focus of recent research. One commonly cited mechanism is that these discrepancies arise due to gender stereotypes learned at a young age (Kurtz-Costes et al., 2008). These stereotypes have a powerful effect on how individuals perceive themselves and their capabilities. Eccles' expectancy-value theory explains the complex relationships between gender stereotypes, motivational outcomes, and subsequent choices and preferences (Wigfield & Eccles, 2000). The conclusion is that these gendered preconceptions, which frame STEM as a male domain, may cause female students to underestimate their competence and value in these areas. The consequence is lower motivational outcomes in STEM and fewer females aspiring to STEM careers (Ahmed & Mudrey, 2019). Investigating strategies to reduce these gender differences in STEM motivational outcomes is therefore paramount to increasing female participation in STEM fields.

2.3 Reducing Gender Differences in STEM Motivation

Understanding why these differences exist between male and female students is crucial, but it is only one part of the solution. Given the importance of motivational outcomes in shaping student aspirations and future career choices, it is essential not just to investigate the origins of these differences but also to explore strategies that could potentially reduce these gender differences with the hope of consequently increasing female participation in STEM areas. Addressing this issue has been a long-standing goal in research on gender and STEM. The past decades have seen greater demand from researchers and practitioners alike for more studies investigating possible methods to reduce gender differences in STEM motivational outcomes (Dasgupta & Stout, 2014). Suggested avenues forward range from changes on the curricular or school level (e.g., Hübner et al., 2017; LaForce et al., 2019) to extracurricular programs in informal learning settings (e.g., Levine et al., 2015; Stringer et al., 2020), to family and community initiatives (e.g., Gumaelius et al., 2016; Rozek et al., 2015). The most prominent focus in educational research tends to be identifying effective changes within the classroom environment, as it is where students live out most of their educational experiences and therefore has a marked impact on their motivational outcomes (Tollefson, 2000).

The current dissertation examines two different potential approaches for reducing gender differences in student motivational outcomes in STEM, in an effort to address multiple research gaps in the literature. First, this dissertation assesses a targeted approach in the form of school-based interventions. School-based interventions are strategies implemented in classroom settings which are intended to directly affect certain student outcomes. These interventions can be considered a targeted approach because they are designed with the purpose of altering student outcomes and tested empirically using experimental research designs (Hsieh et al., 2005). Various school-based interventions have been tested in recent decades by educational researchers, and therefore, many primary studies have examined the effects of these interventions on student motivational outcomes (Hecht et al., 2019; Hulleman & Barron, 2016; Lazowski & Hulleman, 2016). However, many of these interventions have not examined the gender-specific effects of the strategies used, and no comprehensive overview exists that examines whether these school-based interventions that target motivational outcomes could have implications for gender differences in STEM subjects. Second, as another potential factor, this dissertation investigates teacher constructive support, an important aspect of the daily classroom environment. Contrary to school-based interventions, constructive support is not a targeted approach but rather an inherent characteristic of the classroom environment and a component of the regular interactions between students and teachers (Klieme et al., 2009). More importantly, constructive support is also strongly related to student motivational outcomes, and has been shown to have significant effects on these motivational outcomes overall (Patrick et al., 2007). However, to date, research has not investigated whether high levels of constructive support have differential effects on male and female motivational outcomes in STEM subjects. Combining these two distinct avenues presents a holistic examination of potential mechanisms to address gender differences in STEM motivational outcomes. Together, they encompass a broad spectrum of the educational landscape, capturing both deliberate intervention strategies and the subtleties of everyday classroom dynamics. A refined understanding of both these areas offers invaluable insights for researchers, policymakers, and educators.

2.3.1 School-Based Interventions

With regard to education-based strategies intended to promote and foster student motivational outcomes, one of the most frequently studied methods is school-based interventions. School-based interventions are a popular tool in educational research and have been used in many different contexts to change, promote, or support various student outcomes, including physical health, conduct behavior, and academic achievement (Rathvon, 2008). A school-based intervention refers to a manipulation in a classroom environment that differs from regular instruction, and is designed to (positively) change certain student outcomes (Hsieh et al., 2005). With respect to motivational outcomes, many studies have evaluated various school-based interventions that aim to positively foster student motivational outcomes and have demonstrated that these interventions can indeed strengthen student motivational outcomes overall (Lin-Siegler et al., 2016; Wigfield & Wentzel, 2007). For example, in a meta-analysis of 74 motivation interventions in educational settings, Lazowski and Hulleman (2016) found that the interventions were effective, with an average mean effect size of Cohen's $d = 0.49$. For individual interventions, they found that most of the effect sizes were moderate to large ($d = 0.27 - 1.28$). Durlak and colleagues (2011) also conducted a meta-analysis on 213 school-based interventions that targeted student social, emotional, and motivational outcomes. Their findings indicated that, in general, students who took part in these interventions reported significant improvements in social, emotional, and motivational outcomes compared to control groups. When it comes to STEM, there have been a sparse number of reviews that have aggregated the effects of interventions conducted specifically in STEM subjects. For instance, van den Hurk and colleagues (2019) performed a systematic review of interventions aimed at enhancing interest and persistence in STEM subjects. They found positive effects of some interventions but could not assess overall effects from systematic review alone. Similarly, Rosenzweig and Wigfield (2016) also found in their systematic review of 52 studies that some interventions in STEM showed positive effects on motivational outcomes, but others showed mixed or even negative effects. This systematic review was the only research synthesis on school-based interventions to date that examined whether the interventions had gender-specific effects. They found that of the included studies, only 13 examined gender-specific effects. Six of the studies found that the interventions had gender-specific effects, while the other seven did not find any gender-specific effects. These effects, however, could not be evaluated statistically due to the qualitative nature of systematic reviews. The existing literature demonstrates a clear need for a meta-analysis of school-based interventions in STEM that can conclusively evaluate the overall effect sizes.

The manipulations used in intervention studies on motivational outcomes vary widely but can be grouped into broader subcategories. One category is psycho-social interventions. Psycho-social interventions use strategies that do not necessarily teach educational content but instead focus directly on student motivational outcomes by targeting the subjective psychological processes that underlie these outcomes, and attempting to change them for the better (Walton, 2014). These interventions utilize various strategies based on well-established psychological theories (Yeager & Walton, 2011). For instance, interventions centered on cultivating a “growth mindset” are designed to instill in students the understanding that their skills and abilities are not static or predetermined, and that with consistent effort, dedication, and the right strategies, they can enhance and strengthen their talents (Yeager & Dweck, 2020). Growth mindset interventions seek to foster this perception in students, promoting a more adaptive and resilient approach to learning challenges (Burnette et al., 2023). Another example is attributional retraining interventions, which aim to change students’ beliefs about the reasons for their failures and successes (Walton & Wilson, 2018). By reinforcing the idea that students are in control of their outcomes and dispelling the attribution of failure to causes outside their control, these interventions can influence students’ self-beliefs and motivations (Robertson, 2000). Another example is self-efficacy interventions, where students are taught various skills, such as goal setting, or given productive feedback, in order to increase their feelings of competency, confidence, and motivation for completing tasks (Schunk & Ertmer, 2000). While not exhaustive, these examples represent the central strategies behind psycho-social interventions. In contrast, a different category of interventions focuses on instructional processes within the classroom rather than directly targeting motivational and emotional factors. The main goal of these instructional interventions has typically been to increase achievement; however, they have also been shown to positively impact students’ motivational outcomes. For instance, the use of active learning strategies, such as inquiry-based learning or project-based learning, allows students to play a central role in their own learning process (Almulla, 2020). In such frameworks, students actively engage with the material, often exploring real-world questions or problems, which leads to a deeper understanding and connection with the subject matter. This heightened sense of ownership and relevance in the learning process has been linked to improved engagement, motivation, and other academic outcomes (Bruder & Prescott, 2013; Kokotsaki et al., 2016). Additionally, instructional methods that incorporate technology, such as interactive online lessons or digital games, have also been shown to positively influence students' motivational outcomes (Erhel & Jamet, 2013; S.-K. Wang & Reeves, 2007). Technology-enhanced instructional methods can captivate student attention and foster motivation, interest, and engagement by making learning more interactive and aligned with the technology they use in their everyday lives (Lieberman et al., 2009).

Considering the motivational deficits that female students tend to display in STEM subjects it is plausible that females may respond more strongly to interventions targeting motivational outcomes in STEM, even when this is not the primary goal of the intervention. However, given the considerable success of these various interventions for motivational outcomes overall, researchers have also investigated whether school-based interventions could be a valuable tool in reducing gender differences in student motivational outcomes in areas such as STEM. These gender-targeted interventions aim to challenge the gender stereotypes held by students in certain subjects in order to reduce the effects of these stereotypes on motivational outcomes. For instance, some studies have investigated whether exposing students to role models or mentors who defy traditional gender roles, such as a female engineer or female programmer, could boost female student motivational outcomes in STEM subjects (González-Pérez et al., 2020; Morgenroth et al., 2015). By introducing students to individuals in non-traditional gender roles, the goal is to challenge the gender stereotypes they may hold and reduce the impact of these stereotypes on their motivational outcomes (Shin et al., 2016). Another strategy involves targeting students' beliefs about the value of STEM. According to Eccles (1994), students are more likely to engage and invest effort in a subject if they perceive it as personally valuable. In the context of STEM, female students may not find much value in these subjects as they do not align with their gender identity. Making the subject matter personally relevant to female students may alter their perception of the subject's value, thereby increasing the likelihood of their engagement, interest, and motivation (Casad et al., 2018; Gaspard et al., 2015).

Beyond the specific strategies of interventions, contextual and participant characteristics may also play a role in determining effectiveness. Research has explored how student grade levels influence the success of interventions targeting motivational outcomes (Wigfield & Wentzel, 2007). Although motivational outcomes can be influenced at any age, evidence suggests that they are more malleable and easier to foster in younger students (Gutman & Schoon, 2013; Snyder et al., 2019). As children transition into adolescence, motivational outcomes like self-perceptions, interest, and engagement often naturally decline, which may pose challenges to promoting them (Eccles et al., 1993; Jacobs et al., 2002). Some meta-analyses suggest stronger intervention effects in younger age groups (Lazowski & Hulleman, 2016; Unrau et al., 2018), but conclusions remain tentative, leaving grade-level effects underexplored. School subject is another contextual factor to consider when evaluating intervention effects on motivational outcomes. Given that motivational outcomes vary by subject and are intricately linked to each specific domain (Bong & Clark, 1999), the effects of interventions on these outcomes might differ across subjects. Additionally, there is some evidence that gender differences in motivational outcomes vary between subjects and may be more pronounced in some STEM subjects than others. Large-scale international studies such as PISA or the Trends in Mathematics and Science Survey results showed that while, on average, girls had lower self-efficacy and self-concept than boys in both mathematics and science, the differences were much wider in mathematics (OECD, 2015; Reilly et al., 2019). The duration of the intervention may also moderate its effects; however, findings are mixed, and this potential moderator is relatively underexamined in studies on motivational outcomes. While a small number of reviews have found slight moderator effects of intervention duration on other student outcomes (Durlak et al., 2011; Hattie et al., 1996), other reviews have found no effects of intervention duration (Dignath et al., 2008; van Eerde & Klingsieck, 2018). Collectively, these characteristics represent potential contextual factors that may influence the efficacy of these school-based interventions. Assessing these moderators, alongside the overall gender-specific effects of the interventions, is vital to addressing the research gap in this area.

2.3.2 Teacher Constructive Support

While school-based interventions are promising methods for reducing gender differences in motivational outcomes in STEM, they are often supplemental to regular instruction and usually not an integral part of students' day-to-day life in the classroom. In addition to investigating intervention strategies, it is also crucial to identify which aspects of everyday classroom experiences might contribute to fostering higher levels of motivational outcomes for female students in STEM subjects. The classroom is influenced by a multitude of diverse factors that shape students' experiences. These include elements such as teaching methods, feedback and assessment practices, peer interaction, and classroom resources (Urdan & Schoenfelder, 2006). However, one of the most important factors in the classroom involves the various interactions between the student and the teacher (Cadima et al., 2010). The teacher is the most important figure in the classroom in relation to the student, and the various ways that students and teachers interact during and outside of instruction greatly impact a variety of student outcomes (Anderman et al., 2011; Burić & Kim, 2020).

Of these various classroom characteristics, one aspect that is especially strongly linked to student motivational outcomes is teacher constructive support (Cornelius-White, 2007). Constructive support is a heavily researched construct in the literature on instructional quality and is also commonly referred to as supportive climate (Klieme et al., 2009) or positive climate (Burić & Kim, 2020). Constructive support refers to the quality of social interactions between teachers and students in classroom settings, characterized by elements of interest, respect, warmth, and emotional connection (Praetorius et al., 2014). Providing constructive support means accommodating individual student needs, giving encouragement, creating a safe environment to ask questions, and intentionally building positive student relationships (Fauth et al., 2014; Praetorius et al., 2018). Given that constructive support encompasses such an extensive range of characteristics, it can be further divided into two distinct components: instructional support and social-emotional support (Hamre & Pianta, 2005). Instructional support encompasses teachers' dedication to student learning and providing the support necessary for instructional challenges. Teachers provide instructional support by actively scaffolding student learning opportunities, offering productive and specific feedback, and fostering concept development and higher-order thinking skills (Pianta et al., 2012). On the other hand, social-emotional support refers to teachers' sensitivity, concern, and responsiveness toward students on a personal, emotional level. Examples of social-emotional support include displaying empathy and reassurance for student difficulties, effectively addressing misbehaviors, and establishing safe emotional dynamics with students (Mashburn et al., 2008). These two distinct facets are empirically distinguishable. However, they are strongly correlated and can be categorized within the broader framework of constructive support (Decristan et al., 2022; La Paro et al., 2004).

Constructive support is one of the aspects of classroom quality most strongly related to student motivational outcomes, and research suggests that students who perceive high levels of constructive support from their teachers report higher levels of various motivational outcomes (Allen et al., 2006; Cornelius-White, 2007). Constructive support strongly predicts student interest, especially in STEM subjects such as mathematics or science (Lazarides et al., 2019; Lazarides & Ittel, 2013; Lerkkanen et al., 2012). Constructive support has also been shown to positively impact student self-efficacy and competency beliefs (Aldridge et al., 2012; Liu et al., 2018; Ma et al., 2021). Meta-analyses on constructive support have revealed positive overall effects on student engagement (Roorda et al., 2011; Tao et al., 2022), as well as positive associations with positive academic emotions, such as enjoyment, and negative associations with negative emotions, such as anxiety (Lei et al., 2018). Students who perceive higher levels of both instructional and social-emotional support from their teachers are more likely to feel stronger motivation, have a higher value for the subject in question, and invest more effort into learning activities (Dietrich et al., 2015; Federici & Skaalvik, 2013; Koca, 2016; A. M. Ryan & Patrick, 2001). Many researchers have explained these effects through the framework of the self-determination theory of motivation. According to this theory, individuals have inherent psychological needs for relatedness, competence, and autonomy, and when these needs are met, motivation and engagement are enhanced (Deci & Ryan, 2002). Through this mechanism, when students perceive their teachers as being actively engaged with, supportive of, and invested in their educational progress, as well as their emotional well-being, those students will feel fulfilled in their psychological needs, leading to higher motivational outcomes (A.J. Martin & Dowson, 2009).

When discussing the effects of interventions on student outcomes, it is essential to consider that these effects may vary between students due to pre-existing individual differences. This phenomenon is a central framework of psychology research called the aptitude-treatment interaction. The aptitude-treatment interaction asserts that treatments, interventions, or other factors will have varying effects on individuals depending on their different aptitudes (Snow & Swanson, 1992). In educational contexts, the term “aptitude” refers to any individual differences and includes cognitive abilities, as well as motivational or affective learner characteristics (Snow, 1992). This paradigm has also been applied to characteristics of instructional quality. In this case, aspects of classroom quality are considered “treatments” and different student characteristics “aptitudes” (Kieft et al., 2008; Rodger et al., 2007). Following this, some studies have found that constructive support may be especially beneficial for students already at risk of lower motivational outcomes (Curby et al., 2009; Decristan et al., 2016; Malecki & Demaray, 2006). This specific phenomenon regarding constructive support has been explained through the academic risk perspective, which proposes that for students who are already at risk of displaying lower levels of certain outcomes, such as motivational outcomes, the presence of supportive relationships and resources in the environment may have a particularly stronger influence (Hamre & Pianta, 2005). In the context of STEM, research has indicated differences in motivational outcomes between male and female students from an early age. Therefore, constructive support might be especially beneficial for females in STEM. Constructive support creates an atmosphere where students feel safe exploring and experiencing failure (Furrer & Skinner, 2003; Hughes & Chen, 2011), which may particularly bolster female students who feel less competent in STEM subjects. Constructive support can also lead to students feeling a deeper sense of belonging in the classroom (Liu et al., 2018). This may also be especially beneficial for females, who tend to perceive themselves as not belonging in STEM classrooms (Dasgupta & Stout, 2014; Good et al., 2012).

Despite sound theoretical reasons suggesting that constructive support might have particularly beneficial effects for females, its relationship to gender differences in motivational outcomes within STEM contexts remains largely uninvestigated. Only a few studies have investigated whether constructive support has differential effects on male and female motivational outcomes. There is some evidence that constructive support may be more connected to motivational outcomes in general for females than for males in both primary and secondary school (McFarland et al., 2016; Rueger et al., 2008). In STEM subjects, in particular, a sparse number of studies have found that different aspects of constructive support were more strongly related to competence beliefs and engagement for females than for males (Fredricks et al., 2018; Vekiri, 2010). Nevertheless, to date, no studies have explicitly explored whether constructive support could potentially act as a moderator between gender and motivational outcomes in STEM subjects, and if constructive support has especially positive effects on the motivational outcomes of females. Whether the overall benefits of constructive support for motivational outcomes differ for male and female students remains to be seen. Indeed, scholars in this field of research have highlighted the need for a deeper investigation into the gender-differential effects of constructive support for student outcomes (Rueger et al., 2008). Investigating these effects is necessary in order to fully understand which classroom characteristics have the potential to reduce gender differences in STEM motivation.

2.3.3 Summary

The following section elaborated on the topic of reducing gender differences in STEM motivational outcomes. While understanding the reasons behind these differences is important, exploring strategies that can alleviate them and increase female participation in STEM is equally crucial. This dissertation takes a multifaceted approach to this goal by exploring two possible factors that could play a role in reducing these gender differences. School-based interventions have been extensively studied as a targeted approach to promote motivational outcomes overall, and various intervention studies have shown positive effects on student motivation (Durlak et al., 2011; Lazowski & Hulleman, 2016). However, a comprehensive overview of the effectiveness of these interventions in relation to gender differences in STEM motivational outcomes is lacking. Another aspect explored in this dissertation is teacher constructive support, which refers to the quality of interactions between teachers and students. Constructive support is strongly and positively linked to student motivational outcomes (Fauth et al., 2014). However, to date, there has been no research on whether there are gender-specific effects of constructive support in STEM subjects. Understanding and evaluating the impact of both school-based interventions and constructive support provides a holistic perspective and guides future research on aspects of the educational experience that could help to reduce gender differences in STEM motivational outcomes.

3 Research Gaps and the Present Research

In light of compelling evidence indicating lower levels of motivational outcomes among females in STEM subjects, educational researchers and policymakers have emphasized the need for further research on this topic. They call for exploring strategies within educational contexts that can foster higher levels of motivational outcomes for female students and, ultimately, assist in reducing the gender gap prevalent in STEM fields (Hammond et al., 2020; Sánchez-Tapia & Alam, 2020). This dissertation seeks to contribute significantly to this field by adopting a multifaceted approach. This approach encompasses both the evaluation of existing research and the investigation of novel methods that have yet to be explored. This approach also recognizes the need to consider both targeted approaches that address students' individual factors, as well as school and classroom contextual factors that may also play a role in motivational differences (M.-T. Wang & Degol, 2013). Targeted approaches such as school-based interventions have been shown to be effective in fostering positive motivational outcomes (Lazowski & Hulleman, 2016) and therefore may also be successful strategies for doing the same in terms of gender differences in those motivational outcomes. However, the overall effectiveness of these interventions in relation to gender differences in STEM motivational outcomes remains unclear. Moreover, inherent elements of daily classroom dynamics profoundly influence student motivation. Among these elements, teacher constructive support stands out as particularly influential in shaping motivational outcomes (Cornelius-White, 2007), yet its relationship with gender differences in STEM motivational outcomes remains underexplored. By integrating these two aspects into one research project, the present dissertation aims to provide a holistic understanding of the factors contributing to differences in gender-specific levels of motivational outcomes. The goal of the present dissertation was to address the specific research gaps highlighted in the previous section regarding the potential of school-based interventions and teacher constructive support for reducing differences in male and female motivational outcomes in STEM subjects. Therefore, the superordinate research question was the following:

Do school-based interventions and teacher constructive support have gender-specific effects on student motivational outcomes in STEM subjects?

Under this overarching research question, more specific research questions were posed, addressing the topics of school-based interventions and teacher constructive support in greater detail.

3.1 Research Question 1

School-based interventions are a common tool used in educational research to promote and foster a broad range of positive student outcomes. They have been frequently used to directly or indirectly influence student motivational outcomes, especially in STEM subjects (Lin-Siegler et al., 2016; Rosenzweig & Wigfield, 2016). However, the results of these intervention studies tend to vary, and gender-specific effects are often overlooked or not reported. The research on this topic lacks a systematic overview that allows for an overall evaluation of the gender-specific effects of these interventions on student motivational outcomes in STEM subjects. Additionally, the strategies used by these interventions, as well as the characteristics of the samples, settings, and empirical process, vary widely across primary studies. There is a need for a comprehensive assessment of these varying characteristics across studies to investigate whether different context factors may also play a role in the gender-specific effects of these school-based interventions. These research gaps lead to the following research questions:

1_a: Do school-based interventions that promote motivational outcomes in students have differential effects for males and females in STEM subjects?

1_b: Are the gender-specific effects of the school-based interventions moderated by factors such as student age, subject, intervention method, or intervention duration?

3.2 Research Question 2

When addressing gender differences in student motivational outcomes in STEM, a substantial body of research has focused on targeted, novel methods such as school-based interventions. However, a notable research gap remains in regard to the effects of various aspects of everyday classroom practices. Students spend a substantial amount of time in the classroom interacting with teachers and peers, and these enduring experiences have also been shown to have a strong relationship with multiple motivational outcomes (Scherer & Nilsen, 2016). Teacher constructive support is one such aspect that is especially closely connected to various student motivational outcomes (Patrick et al., 2007). However, despite these strong connections, limited empirical attention has been given to the relationship of teacher constructive support to gender differences in student motivational outcomes. Gaining insights into this relationship is essential for discerning the everyday student experiences that may relate to differing levels of gender differences in motivational outcomes in STEM. A better understanding of constructive support in relation to gender differences in STEM motivational outcomes would lead to more targeted research on the topic and could also allow teachers to become aware of how they can better support female students in STEM. Therefore, this dissertation aims to fill this research gap by investigating the following research questions:

2_a: Do student perceptions of constructive support moderate the relationship between gender and students' self-concept, self-efficacy, and interest in secondary school mathematics lessons on quadratic equations?

In addition, the following subordinate research question was also investigated:

2_b: What are the direct effects of gender and teacher constructive support on student motivational outcomes in mathematics?

3.3 Empirical Studies

To address the research questions defined above, two empirical studies were conducted and used as the basis of this dissertation. Study I used meta-analysis to examine the gender-specific effects of school-based interventions on student motivational outcomes in STEM subjects, as well as the effects of various contextual factors across studies, thereby addressing both research questions 1_a and 1_b. Study II employed latent interaction structural equation modeling to investigate the moderating effect of perceived teacher constructive support on the relationship between gender and motivational outcomes in mathematics. This study simultaneously examined the direct effects of both gender and perceived constructive support on motivational outcomes as well, thus addressing research questions 2_a and 2_b. An overview of the research goals, methodology, and main results of these two studies can be found in section 4, and the studies can be found in full in the Appendix (Appendix A and B).

4 Summary of Publications

4.1 Study I

Lesperance, K., Hofer, S., Retelsdorf, J., & Holzberger, D. (2022). Reducing gender differences in student motivational-affective factors: A meta-analysis of school-based interventions. *British Journal of Educational Psychology*, *92*(4), 1502-1536.

4.1.1 Research Goals

When considering the extensive body of evidence regarding gender differences in motivational outcomes in STEM subjects (M.-T. Wang & Degol, 2017), it is a clear objective of research on gender and education to investigate strategies that schools and teachers can use to aid in reducing these differences (M.-T. Wang & Degol, 2013). School-based interventions are a widely used method in various aspects of educational research and have also been frequently used to foster motivational outcomes across various ages and subjects (Lazowski & Hulleman, 2016; Wigfield & Wentzel, 2007). Although a large number of studies have evaluated the effect of various school-based interventions on student motivational outcomes overall, many of these studies do not focus on gender-specific effects and those that do often return mixed results. Additionally, these intervention studies differ widely in terms of the interventional strategies and motivational outcomes they address, and it is unclear which interventions are most promising in regard to gender-specific differences in motivational outcomes. To date, there has not been a comprehensive review and analysis of school-based intervention studies to determine if and how the intervention effects differ according to gender. It is unclear which of these interventional methods, if any, might be more effective for especially fostering female motivation in STEM subjects. Without this knowledge, research aiming to reduce gender differences in student motivational outcomes in STEM cannot advance effectively.

Therefore, to address this research gap, the first research goal of Study I was to investigate whether school-based interventions that target student motivational outcomes can aid in reducing gender differences in said outcomes in STEM subjects. To evaluate as many studies as possible, an extensive range of motivational outcomes were considered based on the theoretical definitions of motivational constructs (P. K. Murphy & Alexander, 2000; Pintrich, 2003b). This included various constructs related to motivation, interest, self-beliefs, goals, and academic emotions. The intervention strategies also varied widely, including various intervention methods. These encompassed interventions that specifically targeted female motivational outcomes in STEM and interventions that targeted all student motivational outcomes in STEM. Including this broad range of studies allowed for a comprehensive, systematic evaluation of their varying effects.

Additionally, given the varying samples, methods, and characteristics of studies on school-based interventions, it is difficult to discern from the primary studies alone under what conditions these interventions might be most effective. In the current body of research, it remains unclear whether the interventions are more successful, for example, at certain school levels, in certain subjects, or using certain methods. Therefore, the second research goal of Study I was to evaluate whether certain moderating variables, such as student age, subject, intervention method, or duration, might influence the effectiveness of these interventions.

4.1.2 Methodology

In order to evaluate the gender-specific effects of school-based interventions across multiple primary studies, this study utilized meta-analytic methods. Meta-analysis allows for the aggregation and analysis of effect sizes from multiple studies in one meta-analytic model in order to evaluate the overall effectiveness of the interventions. The first step in a meta-analysis is a systematic literature search for relevant studies, which involves searching electronic research databases, as well as manually searching through related literature and contacting authors in the field for potential studies. This literature search resulted in $N = 8{,}509$ studies after removing duplicates. The second step was to screen the titles and abstracts of each study to decide whether they met the criteria for inclusion in the meta-analysis in accordance with a set of pre-determined inclusion and exclusion criteria. After this screening process, 171 studies were considered relevant for the meta-analysis. The relevant variables were then extracted from these included studies based on the full-text manuscripts. After eliminating outliers, 20 studies with a total of 71 effect sizes were included in the final analyses. The effect sizes of these studies were then calculated using the extracted information on means, standard deviations, and sample sizes. Effect sizes were calculated separately for males and females for each study and were defined as the difference (g) between the standardized mean change score for the treatment and control groups. Meta-analytic models were then implemented to evaluate the overall effect size across studies.

To identify moderating factors that might impact the effectiveness of the interventions, additional variables of interest were extracted from the primary studies, specifically the variables of school subject (mathematics or science), school level (primary, lower secondary, or upper secondary), intervention duration, intervention method (psychosocial or instructional) and intervention target (gender targeted or not). Differences and similarities among types of motivational outcomes were explored by classifying the outcomes into four subcategories. These subcategories were roughly based on the conceptualization of motivational constructs put forth by Murphy and Alexander and consisted of attitudes, self-schemas, motivation, and affect. All moderating variables were then evaluated using random effects meta-regression models. Meta-regression is an extension of the basic meta-analytic model, which allows researchers to investigate whether certain characteristics of the primary studies might explain between-study heterogeneity.

4.1.3 Main Results and Discussion

The meta-analysis showed that while the interventions had significant effects on student motivational outcomes overall for both genders, there were no significantly different effects for males or females. In other words, both genders benefitted equally from the interventions. However, while not statistically significant, results showed that the overall effect size for females ($g = 0.49$) was descriptively larger than that for males ($g = 0.28$). Nevertheless, these results were not statistically significant and therefore cannot be interpreted as evidence of a true effect.

Varied results emerged regarding the moderator variable analyses. There was a significant effect of the intervention method used. For psycho-social interventions, there was a significantly higher effect for females ($g = 0.53$) than males ($g = 0.19$). Interventions specifically targeted towards females had higher effects for females ($g = 0.63$) than interventions not targeted for females ($g = 0.20$). However, there were no significant differences between males and females in targeted or non-targeted interventions. Descriptively, intervention effects were slightly greater on average for students in primary school ($g = 0.48$) and lower secondary school ($g = 0.43$) than for students in upper secondary school ($g = 0.32$). However, these results were not significant. There were also no significant effects on the school subject or intervention duration. No significant differences were found here regarding the type of motivational outcomes. The largest descriptive effect size could be seen for attitudes, followed by self-schemas, motivation, and affect.

These findings demonstrate that while school-based interventions have positive effects for all students, no significant effects were found based on gender in the interventions overall. While descriptive effect sizes suggest that there might be a higher overall effect for females, more research is necessary to determine whether these effects truly exist. However, the results indicate that psycho-social interventions might be particularly effective for enhancing motivational outcomes for female students than for males. Additionally, it seems that interventions explicitly targeted at female students markedly enhance their motivational outcomes in comparison to female students receiving non-targeted interventions. However, the distinction between gender-targeted and non-gender-targeted approaches does not appear to significantly affect outcomes when comparing males and females. While some descriptive trends were evident for the moderator analyses, no significant effects of grade level, school subject, or intervention duration were observed.

4.2 Study II

Lesperance, K., Decristan, J., & Holzberger, D. (2023). The role of teacher constructive support for gender differences in motivational outcomes in secondary school mathematics. *International Journal of Gender, Science and Technology, 15*(3).

4.2.1 Research Goals

The main topic of Study II centered on the moderating role of teacher constructive support between gender and student motivational outcomes in mathematics classrooms. When investigating methods that could play a role in gender differences in student motivational outcomes in STEM subjects, it is also important to identify aspects of students' daily classroom experiences that may be effective in achieving this goal. Constructive support is an aspect that is especially strongly related to student motivational outcomes (Cornelius-White, 2007). Constructive support can be differentiated into instructional support and social-emotional support, and both facets are positively connected to student motivational outcomes (Hamre & Pianta, 2005; Patrick et al., 2007). However, few empirical studies have investigated whether high levels of constructive support especially benefit females in STEM subjects. Therefore, the main goal of Study II was to evaluate whether student perceptions of constructive support had a moderating effect on the relationship between gender and motivational outcomes in mathematics classrooms. The direct effects of gender and teacher constructive support on the student motivational outcomes were also evaluated. Drawing from prior research on gender differences in STEM and the importance of constructive support for motivation, this study aimed to contribute to research on everyday aspects of teaching quality that could also serve to reduce gender differences.

4.2.2 Methodology

To achieve the research goals of Study II, it was necessary to use a dataset containing both student motivational outcomes and student perceptions of constructive support. It was also necessary that the data was collected from students in a STEM subject. The Teaching and Learning International Survey (TALIS) Video Study 2018 is an ideal dataset for this study, as it includes data on student motivational outcomes, as well as student reports of classroom experiences in mathematics classrooms in lower secondary school. Therefore, this study utilized the German subsample of the TALIS data (OECD, 2020b, 2020c). The final sample for this study consisted of $N = 1{,}116$ students (48.5% female) with a mean age of 15 years. The data was obtained from the TALIS Video Study post-test questionnaires, which were administered after students participated in two mathematics lessons on quadratic equations with their regular mathematics teachers. Items from the student post-test questionnaire were used to model the latent outcome variables of self-concept, self-efficacy, and interest in lessons on quadratic equations. Constructive support was assessed through items asking students their perceptions of the amount of constructive support offered by their teachers, both for instructional support and social-emotional support. Student perceptions are considered an appropriate measure of aspects of teaching quality, especially concerning the importance of those aspects for their individual outcomes (Göllner et al., 2021). Student gender was obtained from the pre-test questionnaire.

To evaluate the fit of the items to the latent variables, confirmatory factor analyses were conducted for all latent variables. Measurement invariance was also tested for all variables to establish comparability between male and female variables. To examine the direct effects of gender and both facets of constructive support on motivational outcomes, a structural equation model with gender as the predictor of self-concept, self-efficacy, and interest was fit to the data as a first step. Subsequently, both facets of perceived constructive support were added to the model as predictors without any interaction effects. To examine the moderating effect of constructive support, a latent moderated structural equation model was evaluated, which included gender, perceived instructional support, and perceived social-emotional support as predictors of the three motivational outcomes, as well as the interaction effects between gender and each of the facets of perceived constructive support. Following recommendations from Maslowsky and colleagues (2015), the model with interaction effects was then compared to the model without interaction effects using the likelihood ratio test statistic to evaluate which model best fit the data. For all models, the standard errors were adjusted to account for the nested data structure, which encompassed students within classrooms.

4.2.3 Main Results and Discussion

Confirmatory factor analysis results of all latent variables displayed excellent model fit, and measurement invariance analyses confirmed configural, scalar, and metric invariance. Regarding the direct effects of gender, results showed that girls reported significantly lower self-concept ($\beta = –.15$, $p = .02$) and lower self-efficacy ($\beta = –.33$, $p < .001$) in mathematics than boys. There was no significant effect of gender on mathematics interest. Perceived instructional support significantly predicted student self-concept ($\beta = .36$, $p < .001$), self-efficacy ($\beta = .24$, $p = .01$), and interest ($\beta = .20$, $p = .03$). Perceived social-emotional support significantly predicted interest ($\beta = .20$, $p < .001$), but not self-concept or self-efficacy. The model with the interaction included was not a better fit to the data than the model without interaction. Additionally, the individual interaction coefficients were non-significant. Therefore, in the current data, no interaction effects were found between gender and perceived constructive support.

Study II underscores the presence of gender differences in motivational outcomes related to mathematics. The findings align with prior research, which has also found that female students in secondary school tend to display lower levels than males of competency beliefs, such as self-efficacy and self-concept, in mathematics (OECD, 2013). However, the results also indicate that these negative effects of gender for females do not exist for the motivational outcome of interest in the current sample. While perceived instructional support positively predicted all three motivational outcomes, positive effects of perceived social-emotional support were only seen for interest. This result suggests that various aspects of constructive support may have differential effects on different aspects of motivation. Due to the absence of interaction effects, these findings did not substantiate the idea that constructive support may have an especially strong effect on female motivational outcomes in mathematics. However, constructive support remains pivotal in nurturing secondary school students' motivation in mathematics, irrespective of gender. Further research is needed to explore which aspects of teaching quality are associated with gender differences in STEM subjects, such as mathematics.

5 Discussion

The overall goal of this dissertation was to investigate school-based interventions and teacher constructive support as two possible areas that could be promising avenues for research concerning the reduction of gender differences in STEM motivational outcomes. In Study I, meta-analysis was used to conduct a comprehensive review of the current body of research concerning school-based interventions in order to systematically evaluate the gender-specific effects of these interventions in STEM subjects. The use of meta-analytic methods also allowed for the inspection of various moderating factors that might influence the results of these interventions. In Study II, a large-scale data set was used to evaluate whether there was a moderating effect of teacher constructive support on the relationship between gender and motivational outcomes in mathematics, a core STEM subject. To date, this relationship has not been thoroughly investigated, especially in the context of STEM motivational outcomes. The combination of targeted approaches as well as the inspection of everyday classroom interactions, allowed for a multidimensional investigation of potential strategies, both established and novel, for reducing gender differences in student STEM motivational outcomes. The main findings and their interpretations in the context of research on gender differences in STEM are discussed below in Section 5.1, starting with school-based interventions, followed by constructive support, and a summary and synthesis of the key findings. This is followed by a discussion of the methodological considerations in Section 5.2, including the strengths and limitations of the dissertation. In Section 5.3, directions for further research as well as practical implications of the results are outlined.

5.1 Discussion of the Main Findings

5.1.1 School-Based Interventions

In order to provide a holistic investigation of strategies that might reduce gender differences in STEM motivational outcomes, the current research took two approaches. The first was to investigate the gender-specific effects of school-based interventions on student motivational outcomes in STEM. Although school-based interventions are a commonly used approach for fostering student motivational outcomes overall (Wigfield & Wentzel, 2007), the literature lacks a systematic overview that investigates whether the effects of these interventions differ for male and female students in STEM subjects. Therefore, through meta-analysis, Study I provides a valuable review of current research while extending the implications of these school-based interventions to the field of gender differences.

The first research question of Study I (research question 1_a) was whether school-based interventions that promote motivational outcomes in students have differential effects for males and females in STEM subjects. The results demonstrate that, on an overall level, school-based interventions have a positive effect on motivational outcomes for all students. This corroborates prior meta-analyses on school-based interventions, which also found that on a general student level, interventions have a positive overall effect on student motivational outcomes (Durlak et al., 2011; Hecht et al., 2019; Lazowski & Hulleman, 2016). This is encouraging, as it indicates that these interventions are functioning as designed, and are a worthwhile method to continue pursuing in research on fostering higher levels of student motivational outcomes. However, regarding the gender-specific effects of the school-based interventions, while the observed overall effect size for females was larger than for males, this difference did not reach statistical significance. This indicates that while a descriptive difference was observed, it is not possible to make any definitive claims regarding the gender-specific effects of these interventions. This lack of significance may be due to the relatively small number of primary studies included in the meta-analysis, which could have resulted in the inability to detect smaller differences between the genders. A larger number of primary studies would increase statistical power, potentially allowing for the detection of smaller effects. Nevertheless, these results should be interpreted with caution, and further research is needed to conclusively detect potential gender differences in intervention effects.

Despite the absence of statistically significant findings, the descriptive effects observed in the study provide valuable insights that merit discussion, as they could tentatively inform future research directions. The larger effect size observed for females, although not statistically significant, leaves open the possibility of a difference in the impact of interventions overall on motivational outcomes between genders. It has been suggested that females may reap greater benefits than males from interventions in STEM, as gender-stereotypical societal beliefs put females at a greater need for motivational support in these subjects (Rosenzweig & Wigfield, 2016). As female students tend to start from a lower level of motivational outcomes, their gains from interventions may be greater than males, who may already have more optimal levels of these outcomes. However, this has yet to be confirmed on a meta-analytic level. The descriptive findings from Study I underscore the importance of conducting additional research to further explore this potential gender-specific effect. Given that few research syntheses have been conducted on school-based intervention effects on motivational outcomes in STEM and even fewer on the gender-specific effects of those interventions, these results provide a promising starting point for future research on this topic. By using larger sample sizes, future investigations can shed more light on whether school-based interventions in STEM subjects might have a more significant impact on the motivational outcomes of female students.

The second research question of Study I (research question 1_b) investigated whether the (gender-specific) effects of the school-based interventions were moderated by factors such as student age, subject, intervention method, or intervention duration. The results of this research question revealed interesting differences across studies. One central moderator was the type of intervention method employed in the primary study. To assess the moderating effect of these intervention methods, it was necessary to classify them into two broad categories: psycho-social and instructional approaches. The results indicated that while both approaches had descriptively larger effect sizes for females, there was only a significant gender difference for psycho-social interventions. Furthermore, the effect size of psycho-social interventions for females was notably greater compared to that of males. In other words, psycho-social interventions had a significantly stronger impact on motivational outcomes for females than for males. These psycho-social interventions target the crucial beliefs and underlying mechanisms directly influencing motivational outcomes (Walton & Wilson, 2018). A substantial portion of the gender differences in motivational outcomes in STEM can be attributed to stereotypical beliefs that shape students' identities, ideas, and perceptions of STEM subjects. Interventions that directly address these outcomes may therefore yield particularly beneficial results for females, who face negative stereotypes with respect to motivational outcomes in STEM subjects. Future research on interventions to promote female motivational outcomes in STEM should continue to pursue and evaluate psycho-social methods.

Further moderator analyses showed that, for females, interventions that specifically targeted females yielded significantly larger effect sizes compared to non-gender-targeted interventions. Interventions that target females specifically often challenge traditional gender stereotypes that many students tend to have in STEM in an effort to reduce the negative effects of these stereotypical beliefs on females' motivational outcomes. As female deficits in motivational outcomes in STEM are substantially due to these stereotypical gender beliefs, interventions that directly target these beliefs may therefore have a stronger effect compared to interventions that also target motivational outcomes but from a more general perspective. These findings underscore that when developing interventions to reduce gender differences in STEM, it is essential to address the specific stereotypical beliefs that often lead to lower motivational outcomes for females and consider how those beliefs relate to the motivational outcomes in question.

No significant differences were found for grade level, school subject, or intervention duration. When looking at the effect of grade level, moderator analyses showed larger descriptive effect sizes for primary and lower secondary school students than for upper secondary school students. However, these results were not statistically significant. While not confirmed in the current study, prior research has found that childhood and early adolescence may be more critical periods for using interventions to target motivational outcomes in students (Lazowski & Hulleman, 2016; Wigfield & Wentzel, 2007). This may be due to the overall decline in motivational outcomes that occurs as students get older (Jacobs et al., 2002; Lepper et al., 2005), making earlier years a more effective time for targeting these outcomes. School subject was also not a significant moderator, and there were no significantly different effects between the interventions conducted in mathematics or science. This indicates that the interventions have the same effects, whether conducted in mathematics or science. Although there is some research to suggest that motivational gender differences are larger in mathematics than in science (OECD, 2015; Reilly et al., 2019), it is possible that similar motivational mechanisms in both mathematics and science were addressed in the included interventions, meaning subject-specific effects might not be evident. However, it is also possible that due to the relatively smaller sample size, the specific subsamples of interventions in mathematics and science did not provide enough power to detect subject-specific effects. Lastly, there was no significant effect of intervention duration on the intervention effects. This is consistent with the mixed results from other research syntheses on intervention studies. While certain syntheses have reported shorter interventions as having larger effects (Hattie et al., 1996), others have found longer interventions to have stronger effects (Slavin et al., 2009), and others have even found no effects of intervention duration (Dignath et al., 2008; van Eerde & Klingsieck, 2018). These mixed results may be due to the varying ways in which intervention duration is often quantified in meta-analyses. However, in the case of the current study, the relatively small sample size and measurement of intervention duration as a continuous variable might have made it difficult to detect any potential effects.

5.1.2 Teacher Constructive Support

Following the results concerning the gender-specific effects of school-based interventions, the question arises: What about more regular aspects of students' classroom experience? The interactions students have in the classroom and the support they receive from their teachers also has a strong impact on their motivational outcomes (R. M. Ryan et al., 1994), however, this has not yet been investigated in relationship to gender differences in motivational outcomes in STEM subjects. Study II therefore complements Study I by examining the moderating role teacher constructive support plays in the differing effects of gender on motivational outcomes in mathematics (a core STEM subject). Study II provides a novel investigation of another facet of student experiences in school that could also contribute to addressing differing levels of motivational outcomes between male and female students.

Before investigating the moderating effects of constructive support, it was necessary to first evaluate the direct effects of both gender and constructive support on the motivational outcomes of self-concept, self-efficacy, and interest in mathematics lessons (research question 2_b). Firstly, when looking at the effects of gender on self-concept, self-efficacy, and interest, the results partially aligned with prior research but also revealed some contradictions. Females had significantly lower levels of self-concept and self-efficacy in mathematics lessons, which is a pattern that can be seen in numerous other studies (Goldman & Penner, 2016; Huang, 2013; S. Lindberg et al., 2013; Pajares, 2005). These results, in combination with prior research, support the claim that female students tend to underestimate their own competence in mathematics and have more negative judgments of their own capabilities in mathematics than male students (OECD, 2015). However, no significant effects of gender were found for the variable of interest in mathematics lessons. This is in contrast to other studies that have found females to report lower levels of interest in mathematics (Frenzel et al., 2010; Köller et al., 2001; Preckel et al., 2008). However, a small number of studies have also found minimal or no differences between boys' and girls' mathematics interest (Fredricks et al., 2018; Jacobs et al., 2002). For example, Simpkins and colleagues (2006), in their analysis of 227 elementary school students, found a significant difference between male and female students in mathematics self-concept but not in mathematics interest. Additionally, Ganley and Lubienski (2016) observed that while there was a slight difference between male and female students' mathematics interest, it was notably smaller than their differences in mathematics competency beliefs. The current results, combined with this prior research, suggest that the effects of gender on interest in mathematics may not be as prominent as the effects on competency-related variables such as self-concept and self-efficacy. Interest, while also related to motivation, is a much more affective outcome than self-efficacy or self-concept. An individual's interest in a subject may not necessarily be as affected by gender-stereotypical beliefs. These findings highlight the complexities of the relationship between gender and motivational outcomes and the importance of addressing these nuances when researching gender differences in STEM motivational outcomes.

The direct effects of the two facets of constructive support varied. While both facets of perceived constructive support were positively related to student interest, perceived instructional support was the only facet that had a positive effect on students' self-concept and self-efficacy. This differentiation provides evidence regarding the crucial role of a supportive classroom environment in fostering student interest. Instructional support, including clear feedback, helpful resources, and guidance, alongside social-emotional support, which addresses students' need for emotional well-being and a sense of belonging, seem to collectively fuel student interest. This aligns with prior research, which has also found similar relationships between constructive support and interest (Lazarides & Ittel, 2013; Patrick et al., 2007). However, a notable distinction appears for self-concept and self-efficacy. The current findings suggest that perceived instructional support has a more vital impact for promoting higher levels of these motivational outcomes in students. Both self-concept and self-efficacy are highly related to students' judgments of their own capabilities, both on a specific and general level (Bong & Skaalvik, 2002). Instructional support directly relates to these judgments by allowing an encouraging environment for students to learn from their mistakes and improve their own skills. These opportunities, combined with constructive feedback and direction when needed, could increase feelings of competence (Havik & Westergård, 2020; Pianta et al., 2012). While social-emotional support is certainly also crucial for student well-being overall, the more affective nature of this type of support might not lend itself as much to students feeling confident in their cognitive abilities.

The main research question of Study II (research question 2_a) evaluated whether student perceptions of constructive support moderated the relationship between student gender and the motivational outcomes of self-concept, self-efficacy, and interest in secondary school mathematics classes. Constructive support has been shown to be especially beneficial for specific groups at risk of having lower levels of motivational outcomes (Curby et al., 2009; Malecki & Demaray, 2006). Additionally, various positive mechanisms of constructive support, such as fostering a sense of belonging (Liu et al., 2018) and creating a safe environment to explore (A. M. Ryan & Patrick, 2001), might directly address specific deficits that female students experience in STEM subjects. However, despite the sound theoretical arguments for why this moderating relationship might exist, no significant moderation effects emerged for gender and constructive support for any of the student motivational outcomes. These findings imply that the positive effects of perceived constructive support are consistent across genders, and they do not have an especially beneficial impact for females, as hypothesized. Although constructive support has been shown to have higher effects for certain disadvantaged groups (Curby et al., 2009; Decristan et al., 2022; Malecki & Demaray, 2006), the current study was not able to find support for these same differential effects pertaining to gender differences. It is possible that the beneficial effects of constructive support on motivational outcomes may not be sufficient to account for the specific disadvantages that female students face in STEM. The root causes of lower motivational outcomes among females in STEM subjects, such as gender stereotypical beliefs, may not be addressed by constructive support. Therefore, even though constructive support is beneficial, it raises the motivational outcomes of males and females equally without providing the extra boost necessary to even the levels between genders. It is possible that the inherent nature of constructive support is not targeted enough to combat these stereotypes, and more explicit measures are needed to bolster female students’ motivational outcomes in STEM.

There are methodological aspects of Study II that may have also contributed to the absence of interaction effects between gender and constructive support. The cross-sectional use of the data limited the evaluation of interactions between gender and perceived constructive support over time. Since motivational outcomes develop over the course of a student’s life, a measurement at one time point or even a short period of a few weeks may not be sufficient to detect more complex interactions. Although this study attempted to measure the variables of interest as specific states in a particular education context, future research might consider longitudinal designs to provide more precise insights into these connections. Moreover, two main approaches exist for measuring motivational outcomes. Motivational outcomes measured as states reflect perceptions uniquely associated with specific academic experiences, whereas motivational outcomes measured as traits depict enduring, long-term propensities across various educational contexts (Soland et al., 2019; Wasserman & Wasserman, 2020). There is evidence suggesting that motivational outcomes measured as states can vary in relation to situation-specific factors and interactions, especially in relation to classroom environments (Gaspard & Lauermann, 2021; A. J. Martin et al., 2020; Patall et al., 2018). Given that Study II focused solely on state-related motivational outcomes, the potential for interaction effects in regard to trait motivational outcomes must be acknowledged. Nonetheless, a comprehensive evaluation of such interaction effects would also require a longitudinal study design.

5.1.3 Summary and Synthesis

In this dissertation, the combined findings of the two included studies highlight the intricate dynamics between gender, motivational outcomes in STEM, and approaches for reducing gender differences in these outcomes. A central insight is the need for a more detailed understanding of this relationship and the development of specialized strategies that address these specific dynamics. While existing literature has identified common gender patterns associated with motivational outcomes and intentions to pursue STEM (Eccles & Wang, 2016), the current dissertation reveals a more nuanced understanding of these patterns and suggests that strategies to address them might not always be universally effective. Specifically, Study II underscores that the most pronounced gender differences are found in competency-related beliefs, particularly in self-concept and self-efficacy in mathematics, rather than in interest in mathematics. This observation implies that efforts to boost female students' motivational outcomes in STEM may be more effective by paying particular attention to females' competency beliefs. In turn, Study I demonstrates that not all intervention strategies are equally effective for reducing gender differences in motivational outcomes. In particular, psycho-social interventions, designed to engage directly with motivational mechanisms, appear to have a more profound influence on female students in STEM than males. Further, the findings from Study I regarding the higher effects for female students of interventions that specifically target females, compared to female students who receive non-gender-targeted interventions, emphasizes the importance of not only using psycho-social interventions but also tailoring these interventions to address gender-specific challenges for females in STEM. In synthesis, these findings advocate for a targeted research approach, emphasizing psycho-social interventions that aim to bolster female competency beliefs in STEM. As indicated by this dissertation, such an approach could be pivotal for enhancing specific motivational outcomes where female students in STEM are most in need.

5.2 Methodological Considerations

5.2.1 Strengths

The two studies conducted in the current dissertation have several strengths that are valuable to research on the topic of gender differences in STEM motivational outcomes. One of the main strengths of this dissertation was the complementary approach of considering two different aspects that could contribute to reducing gender differences in STEM motivational outcomes. While substantial research has been conducted on strategies for reducing gender differences in STEM, much of this research tends to focus on singular strategies, such as specific interventions or extracurricular programs (Liben, 2015; Stringer et al., 2020). While worthwhile, these endeavors neglect to consider the multi-faceted nature of gender differences in STEM motivation. Research into these gender differences has called for more comprehensive investigations that consider multiple aspects of educational practice, including classroom interactions and targeted, in-class strategies (dos Santos et al., 2022; Kuchynka et al., 2022). The current dissertation directly addresses this need by evaluating both school-based interventions and constructive support in the classroom as two possible factors for reducing gender differences in STEM motivational outcomes.

One of the prominent strengths of Study I is the use of meta-analysis to evaluate the research questions regarding school-based interventions and gender differences in STEM. Meta-analysis allows for a systematic aggregation of numerous primary studies whose various results might otherwise be overwhelming (Hedges & Pigott, 2001). When evaluating various primary studies conducted on the same topic, it can often be challenging to gain an overview of the differing results. Utilizing meta-analytic methods enables a robust estimate of the overall intervention effects, while also accounting for inconsistencies and variation across studies (Bangert-Drowns & Rudner, 2019). A large range of studies were considered, encompassing a variety of motivational outcomes, intervention strategies, and populations, allowing for the generalizability of the results to various contexts. This generalizability relates to another advantage of meta-analysis, which is the possibility to assess moderating variables that might impact intervention effects across studies (Hedges & Pigott, 2004). Research related to gender differences in STEM has recognized the need for more systematic reviews of intervention studies, not just to examine their overall effect, but also to evaluate moderators of those effects (Master & Meltzoff, 2020). By examining moderating variables such as school level, subject, intervention method, and duration, this study provided a closer look at the factors underlying these interventions that might have had an impact on their success. Understanding these factors is crucial for effectively implementing interventions in real-world settings.

The analysis methods chosen for Study II also reflect a considerable strength of the current dissertation. A primary strength was the use of latent variable modeling. One of the main advantages of using latent variables in analyses is that the measurement error can be adjusted for in the measured variables, allowing for more accurate parameter estimates (Ledgerwood & Shrout, 2011). As a result, the estimations of relationships between variables are more accurate and more reflective of the actual underlying relationships. Additionally, while various approaches have been proposed for modeling interactions between latent variables (Kenny & Judd, 1984; Little et al., 2006; Marsh et al., 2013), the latent moderated structural equation modeling approach used in Study II is one of the most strongly recommended methods (Cortina et al., 2021). Compared to other methods of latent interaction modeling, latent moderated structural equation models produce higher power, more accurate estimates, and smaller standard errors (Cheung et al., 2021; Klein & Moosbrugger, 2000). These strengths make latent moderated structural equation models an especially sound method for investigating complex interaction relationships.

5.2.2 Limitations and Future Research

While this dissertation has many strengths and contributes essential findings to research on gender differences in student STEM motivational outcomes, there are also various limitations to be considered that provide directions for future research on this topic.

Both studies primarily employed self-report data measures to assess student motivational outcomes. In Study I, the reliability of the self-report instruments was evaluated and controlled for in the meta-analysis. Additionally, in Study II, the instruments used demonstrated high levels of reliability and were developed by an expert team. Self-report measures are the most widely used method to assess motivational outcomes and are considered appropriate given the personal, subjective nature of these outcomes (Fulmer & Frijters, 2009). However, the use of self-report instruments inherently involves challenges, including biases such as the social desirability effect, where respondents might inadvertently or consciously present themselves more favorably (Fisher & Katz, 2000). Self-report instruments also introduce the potential for common method variance, where the results could be influenced more by the measurement method rather than the actual constructs being investigated (Donaldson & Grant-Vallone, 2002). Considering these challenges, future research could employ more diverse research designs. For instance, integrating both qualitative and quantitative data through mixed-method approaches could provide a more comprehensive understanding of motivational outcomes and how various strategies affect these motivational outcomes for males and females (Creswell, 1999).

Regarding Study I, one disadvantage of meta-analysis is its limitation to the studies available in the literature at the time of the analysis. In the current study, 35 relevant studies had to be excluded during the coding process for various reasons, resulting in a smaller sample size. A substantial number of studies were excluded due to a lack of a control group, pre-test measures, or other missing study quality criteria. Research syntheses such as meta-analyses require high-quality primary studies to deliver robust results, meaning it is crucial to evaluate the rigor of primary studies included in the analyses and exclude studies that do not meet the established criteria (Siddaway et al., 2019). However, this resulted in a smaller sample size and a narrower range of intervention methods. This issue highlights the need for future intervention research to follow established guidelines for producing high-quality research, including the use of control groups, evaluation of baseline equivalency, reliable measurement instruments, and transparent reporting (What Works Clearing House, 2022). Additionally, several studies investigated relevant interventions without reporting gender-specific results, which made it impossible to calculate distinct effect sizes for males and females. This resulted in the loss of potentially beneficial intervention methods, as well as a more restricted sample size. Future research on school-based interventions and motivational outcomes in STEM should strive to report not only the overall effects of the interventions, but also the differential intervention effects for males and females, even if this is not a central aspect of the research questions. There is a crucial need for intervention research disaggregated by gender to bolster efforts to increase gender equality in STEM subjects (UNESCO, 2016). The relatively smaller pool of studies included in the meta-analysis led to an additional methodological limitation, in that each moderator was evaluated in a separate model. Evaluating all moderators in a single meta-regression model would have allowed for a more complete understanding of the nuanced interactions between these variables and an estimate of the complete variance explained by these moderators overall. However, the limited sample size would not have provided sufficient statistical power to robustly assess these relationships. Finally, the results revealed a great deal of statistical heterogeneity across primary studies that could not be fully explained by the moderators included in this study. This indicates that there are additional context variables influencing these relationships that were not accounted for. Access to a larger pool of high-quality studies that report gender-specific effects will allow future researchers to examine the moderator effects across studies more precisely, as well as further research on additional variables that could play a role in the effects of these interventions on male and female motivational outcomes in STEM.

In Study II, the use of a cross-sectional approach limited the ability to examine the relationship between constructive support and gender over an extended period of time. However, it is possible that the influence of this relationship on motivational outcomes may only become apparent after prolonged time periods. Although Study II attempted to account for this by focusing on context-specific motivational outcomes measured as states, it is possible that the interactions involving gender and constructive support might be less perceptible within such a design. Moving forward, research on this topic should examine the long-term relationship between gender, constructive support, and motivational outcomes in STEM, using longitudinal data to explore these variables over a more substantial period of time. A longitudinal approach would also allow for the examination of reciprocal effects, which was also not possible with the current study design. There is some evidence that over more extended periods of time, motivational outcomes might also influence students' perceptions of their classroom environment (Lazarides & Ittel, 2013). Collecting longitudinal data would allow researchers to examine these potential effects in relation to gender as well. Another limitation of Study II is the generalizability of the data. The findings of Study II are based on a German subsample of the TALIS Video Study 2018 dataset and are therefore specific to the German educational system. Additionally, the majority of students in the sample were from the school track in Germany with the most demanding academic curriculum (Gymnasium). Future studies should explore the interaction between constructive support and gender within more varied samples to determine whether the consistency of these findings holds true for STEM motivational outcomes in more diverse populations. An additional limitation is that Study II only examined the subject of mathematics. Mathematics is one of the core subjects of STEM, and is also relevant for other STEM topics such as physics and engineering (English, 2016). However, motivational outcomes are domain-specific, and gender differences in these outcomes have been shown to vary across different STEM subjects (Green et al., 2007; OECD, 2015). As such, the current results cannot be generalized to other STEM subjects. Future research should also explore these relationships in a broader range of STEM fields to compare the results with the current findings.

5.3 Implications for Research and Practice

Gender differences in STEM motivational outcomes are a pressing concern in education, and there is a need for both research and practical directions to confront this issue. This dissertation examined two primary areas related to gender differences in STEM: the impact of school-based interventions and the role of teacher constructive support. The investigations and findings from this dissertation provide valuable directions for future research beyond the directions already mentioned in terms of overcoming limitations. The current findings also offer insights for educational practice and policy. In the following section, the practical and research implications of the current dissertation are discussed.

5.3.1 Research Implications

Beyond the directions for future research already mentioned in terms of overcoming limitations, further suggestions for continuing research on the topic of gender differences in STEM motivation can be gleaned from the current dissertation. First, the results from Study I reveal the potential efficacy of psycho-social interventions for decreasing gender differences in motivational outcomes in STEM subjects. Going forward, researchers should therefore continue to develop, evaluate, and replicate psycho-social intervention studies in STEM subjects. This will allow for a more thorough and established understanding of the positive effects these psycho-social interventions have on female deficits in STEM motivational outcomes. Additionally, psycho-social interventions encompass a wide variety of strategies, all with the aim of targeting specific psychological processes (Walton, 2014). The strategies used by the psycho-social interventions in Study I were diverse, including role model interventions, utility-value activities, and sense of belonging exercises. However, due to the smaller sample size of the meta-analysis, it was not possible to examine the efficacy of each specific strategy included in this category. In order to fully understand which psycho-social intervention strategies are most effective, further research syntheses with a larger pool of studies are necessary. Future research should use meta-analysis to delve deeper into the specific types of psycho-social interventions, evaluating their individual impacts and identifying optimal contexts for implementation. By pinpointing these specificities, research can move closer to establishing best practices for reducing gender differences in STEM motivation.

Although Study II did not reveal any moderating effects of constructive support on the relationship between gender and motivational outcomes, it is important not to disregard the dynamics of the classroom environment when investigating gender differences in STEM motivational outcomes. Apart from the directions for continued research on gender and constructive support, future research should also consider other dimensions of the classroom environment that might interact with gender and motivational outcomes. For example, some studies have shown that in classrooms with higher levels of cognitively activating activities, such as high-quality discourse or challenging learning tasks, students tend to display higher levels of motivational outcomes, such as interest, self-concept, and motivation (Lepper & Henderlong, 2000; Lipowsky et al., 2009; Seidel et al., 2005). When it comes to gender differences, it is crucial to not only focus on the individual but also to consider the social and contextual environment in which that individual spends time (Taylor, 1996). Recognizing that both context and environment play a role in influencing student motivation is crucial for investigations into gender differences in STEM motivational outcomes, however, there is a lack of research on this topic. This dissertation takes the first step into understanding the relevance of classroom environment for gender differences in motivational outcomes, and further research should continue investigations on this topic

5.3.2 Practical Implications

The current dissertation further solidifies the effectiveness of school-based interventions in STEM education, with benefits for both male and female motivational outcomes. These findings, in combination with prior research demonstrating the effectiveness of these interventions for motivational outcomes (Lazowski & Hulleman, 2016; Wigfield & Wentzel, 2007), present a compelling case for policymakers and politicians to allocate more funding and resources to these initiatives. There is an increasing need to encourage students in STEM studies, and motivational outcomes will play a crucial role in whether or not students will pursue those studies (Eccles & Wang, 2016). Researchers in STEM education have called for the use of more effective pedagogical practices that also address student motivation and engagement (McDonald, 2016; S. Murphy et al., 2019). Educational practice should leverage the substantial body of evidence on school-based interventions in STEM and apply this research to implement effective strategies in curriculums, extracurricular programs, and other school-related contexts. By doing so, educational policymakers and practitioners can not only acknowledge the importance of motivation in STEM learning, but also proactively address a crucial element that can impact the success and enthusiasm of students in these fields.

Along the same lines, this dissertation also reinforces how crucial constructive support in the classroom is for all students and demonstrates how it can positively predict motivational outcomes in STEM subjects such as mathematics. The level of support that teachers provide, both instructionally and on an emotional level, plays a pivotal role in the educational experience and can significantly influence students' motivation. Schools and teachers should prioritize creating environments where these nurturing relationships can thrive. For example, professional development programs could be established to emphasize the significance of constructive support, equipping educators with both the understanding and practical methods to provide ample levels of constructive support to their students. Workshops and training sessions for both in-service and pre-service teachers can be organized to demonstrate how feedback can be given constructively, how to recognize and address individual student needs, and how to foster a classroom environment where every student feels valued. By fostering constructive support in their classrooms, schools, and teachers can help students develop a more profound intrinsic motivation for learning. This ultimately leads to more engaged, confident, and persistent learners who are better equipped to face academic challenges.

On the subject of gender differences, this dissertation offers multiple findings that enable teachers, school leaders, and policymakers to develop and implement tailored strategies that could address gender differences in STEM motivational outcomes. Study I revealed that certain intervention methods, especially those using psycho-social approaches, seem to have an especially positive effect on female students. This presents an opportunity for educational institutions to explore how to integrate these psycho-social intervention strategies into classrooms and curricula. The current dissertation also revealed the specific motivational outcomes where girls might especially need support in regard to STEM subjects. Understanding the specific domains where female students exhibit deficits, notably in competency beliefs but not necessarily in interest, is of utmost importance for educational settings. By recognizing the deficits in girls' self-concept and self-efficacy in STEM subjects, such as mathematics, educators can adapt their instructional methods to include strategies that bolster competency beliefs. By understanding and implementing findings from research studies such as the current one, schools and educators can strategically address and possibly diminish the motivational deficits that girls might experience in STEM. Furthermore, these findings, in combination with the wide body of research on STEM and gender, can inform policy decisions on a larger scale. Educational stakeholders can push for teacher training programs that sensitize pre-service teachers to these issues and train them in evidence-based strategies, ensuring that educators enter the classroom equipped with the tools to specifically address girls' disadvantages in STEM. Schools can also collaborate with organizations that advocate for women in STEM, creating programs that, for example, provide female students with role models or demonstrate the value of STEM in the real world.

6 Conclusion

In the constantly evolving discourse surrounding gender differences within STEM subjects, researchers and policymakers have consistently emphasized the urgency of bolstering motivational outcomes for female students. With a multifaceted approach encompassing both a comprehensive meta-analysis of school-based interventions (Study I) and an investigation of teacher constructive support (Study II), the current research sought to identify factors that could enhance motivational outcomes among female students, contributing to the overarching goal of reducing the gender gap in motivational outcomes in STEM. Notably, the results identified specific intervention methods that could enhance female motivational outcomes in STEM and highlighted areas where female students might require additional support. This dissertation provides valuable implications for researchers, educators, and policymakers working to bridge the gender gap in STEM fields. Through continued research and implementation of evidence-based practices, educational institutions can strive toward a more diverse, equitable, and inclusive STEM landscape, empowering female students and creating a brighter future for STEM education and society.

References

Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research*, *2*(1), 010101. https://doi.org/10.1103/PhysRevSTPER.2.010101

Ahmed, W., & Mudrey, R. R. (2019). The role of motivational factors in predicting STEM career aspirations. *International Journal of School & Educational Psychology*, *7*(3), 201–214. https://doi.org/10.1080/21683603.2017.1401499

Aldridge, J. M., Afari, E., & Fraser, B. J. (2012). Influence of teacher support and personal relevance on academic self-efficacy and enjoyment of mathematics lessons: A structural equation modeling approach. *Alberta Journal of Educational Research*, *58*(4), 614 – 633. https://doi.org/10.11575/ajer.v58i4.55612

Allen, M., Witt, P. L., & Wheeless, L. R. (2006). The role of teacher immediacy as a motivational factor in student learning: Using meta-analysis to test a causal model. *Communication Education*, *55*(1), 21–31. https://doi.org/10.1080/03634520500343368

Almeda, M. V., & Baker, R. S. (2020). Predicting student participation in STEM careers: The role of affect and engagement during middle school. *Journal of Educational Data Mining*, *12*(2), 33–47. https://doi.org/10.5281/zenodo.4008054

Almulla, M. A. (2020). The effectiveness of the project-based learning (PBL) approach as a way to engage students in learning. *SAGE Open*, *10*(3), 2158244020938702. https://doi.org/10.1177/2158244020938702

Anderman, L. H., Andrzejewski, C. E., & Allen, J. (2011). How do teachers support students' motivation and learning in their classrooms? *Teachers College Record*, *113*(5), 969–1003. https://doi.org/10.1177/016146811111300502

Aschbacher, P. R., Ing, M., & Tsai, S. M. (2014). Is science me? Exploring middle school students' STE-M career aspirations. *Journal of Science Education and Technology*, *23*(6), 735–743. https://doi.org/10.1007/s10956-014-9504-x

Ashmore, R. D., & Del Boca, F. K. (1979). Sex stereotypes and implicit personality theory: Toward a cognitive-social psychological conceptualization. *Sex Roles*, *5*(2), 219–248. https://doi.org/10.1007/BF00287932

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (2001). Self-efficacy beliefs as shapers of children's aspirations and career trajectories. *Child Development*, *72*(1), 187–206. https://doi.org/10.1111/1467-8624.00273

Bangert-Drowns, R., & Rudner, L. (2019). Meta-analysis in educational research. *Practical Assessment, Research, and Evaluation*, *2*(1). https://doi.org/10.7275/rw59-1m43

Baram-Tsabari, A., & Yarden, A. (2011). Quantifying the gender gap in science interests. *International Journal of Science and Mathematics Education*, *9*(3), 523–550. https://doi.org/10.1007/s10763-010-9194-7

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral and Brain Sciences*, *11*(2), 169–183. https://doi.org/10.1017/S0140525X00049244

Beyer, S. (2014). Why are women underrepresented in Computer Science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. *Computer Science Education*, *24*(2–3), 153–192. https://doi.org/10.1080/08993408.2014.963363

Bian, L. (2022). Gender stereotypes and education. In D. P. VanderLaan & W. I. Wong (Eds.), *Gender and sexuality development: Contemporary theory and research* (pp. 255–275). Springer International Publishing. https://doi.org/10.1007/978-3-030-84273-4_9

Bloodhart, B., Balgopal, M. M., Casper, A. M. A., McMeeking, L. B. S., & Fischer, E. V. (2020). Outperforming yet undervalued: Undergraduate women in STEM. *PLOS ONE*, *15*(6), e0234685. https://doi.org/10.1371/journal.pone.0234685

Blotnicky, K. A., Franz-Odendaal, T., French, F., & Joy, P. (2018). A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career activities on the likelihood of pursuing a STEM career among middle school students. *International Journal of STEM Education*, *5*(1), 22. https://doi.org/10.1186/s40594-018-0118-3

Bock, R. D., & Kolakowski, D. (1973). Further evidence of sex-linked major-gene influence on human spatial visualizing ability. *American Journal of Human Genetics*, *25*(1), 1–14.

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*(3), 139–153. https://doi.org/10.1207/s15326985ep3403_1

Bong, M., & Skaalvik, E. M. (2002). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*, 1–40. https://doi.org/10.1023/A:1021302408382

Bruder, R., & Prescott, A. (2013). Research evidence on the benefits of IBL. *ZDM*, *45*(6), 811–822. https://doi.org/10.1007/s11858-013-0542-2

Buccheri, G., Gürber, N. A., & Brühwiler, C. (2011). The impact of gender on interest in science topics and the choice of scientific and technical vocations. *International Journal of Science Education*, *33*(1), 159–178. https://doi.org/10.1080/09500693.2010.518643

Burić, I., & Kim, L. E. (2020). Teacher self-efficacy, instructional quality, and student motivational beliefs: An analysis using multilevel structural equation modeling. *Learning and Instruction*, *66*, 101302. https://doi.org/10.1016/j.learninstruc.2019.101302

Burnette, J. L., Billingsley, J., Banks, G. C., Knouse, L. E., Hoyt, C. L., Pollack, J. M., & Simon, S. (2023). A systematic review and meta-analysis of growth mindset interventions: For whom, how, and why might such interventions work? *Psychological Bulletin*, *149*(3-4), 174–205. https://doi.org/10.1037/bul0000368

Cadima, J., Leal, T., & Burchinal, M. (2010). The quality of teacher–student interactions: Associations with first graders' academic and behavioral outcomes. *Journal of School Psychology*, *48*(6), 457–482. https://doi.org/10.1016/j.jsp.2010.09.001

Card, D., & Payne, A. A. (2021). High school choices and the gender gap in STEM. *Economic Inquiry*, *59*(1), 9–28. https://doi.org/10.1111/ecin.12934

Casad, B. J., Oyler, D. L., Sullivan, E. T., McClellan, E. M., Tierney, D. N., Anderson, D. A., Greeley, P. A., Fague, M. A., & Flammang, B. J. (2018). Wise psychological interventions to improve gender and racial equality in STEM. *Group Processes & Intergroup Relations*, *21*(5), 767–787. https://doi.org/10.1177/1368430218767034

Chavatzia, T. (2017). *Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM)* (Vol. 253479). UNESCO Paris, France. http://unesdoc.unesco.org/images/0025/002534/253479e.pdf

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2021). Testing moderation in business and psychological studies with latent moderated structural equations. *Journal of Business and Psychology*, *36*(6), 1009–1033. https://doi.org/10.1007/s10869-020-09717-0

Cimpian, J. R., Kim, T. H., & McDermott, Z. T. (2020). Understanding persistent gender gaps in STEM. *Science*, *368*(6497), 1317–1319. https://doi.org/10.1126/science.aba7377

Cole, J. R., & Cole, S. (1973). *Social stratification in science*. University of Chicago Press.

Conlon, R. A., Barroso, C., & Ganley, C. M. (2023). Young children's career aspirations: Gender differences, STEM ambitions, and expected skill use. *The Career Development Quarterly*, *71*(1), 15–29. https://doi.org/10.1002/cdq.12312

Cook, D. A., & Artino Jr., A. R. (2016). Motivation to learn: An overview of contemporary theories. *Medical Education*, *50*(10), 997–1014. https://doi.org/10.1111/medu.13074

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research*, *77*(1), 113–143. https://doi.org/10.3102/003465430298563

Cortina, J. M., Markell-Goldstein, H. M., Green, J. P., & Chang, Y. (2021). How Are We Testing Interactions in Latent Variable Models? Surging Forward or Fighting Shy? *Organizational Research Methods*, *24*(1), 26–54. https://doi.org/10.1177/1094428119872531

Creswell, J. W. (1999). Mixed-method research: Introduction and application. In G. J. Cizek (Ed.), *Handbook of educational policy* (pp. 455–472). Academic Press. https://doi.org/10.1016/B978-012174698-8/50045-X

Cromley, J. G., Perez, T., & Kaplan, A. (2015). Undergraduate STEM achievement and retention: Cognitive, motivational, and institutional factors and solutions. *Policy Insights from the Behavioral and Brain Sciences*, *3*(1), 4–11. https://doi.org/10.1177/2372732215622648

Crowley, K., Callanan, M. A., Tenenbaum, H. R., & Allen, E. (2001). Parents explain more often to boys than to girls during shared scientific thinking. *Psychological Science*, *12*(3), 258–261. https://doi.org/10.1111/1467-9280.00347

Curby, T. W., LoCasale-Crouch, J., Konold, T. R., Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2009). The relations of observed Pre-K classroom quality profiles to children's achievement and social competence. *Early Education and Development*, *20*(2), 346–372. https://doi.org/10.1080/10409280802581284

Cvencek, D., Brečić, R., Gaćeša, D., & Meltzoff, A. N. (2021). Development of math attitudes and math self-concepts: Gender differences, implicit–explicit dissociations, and relations to math achievement. *Child Development*, *92*(5), e940–e956. https://doi.org/10.1111/cdev.13523

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math–gender stereotypes in elementary school children. *Child Development*, *82*(3), 766–779. https://doi.org/10.1111/j.1467-8624.2010.01529.x

Dasgupta, N., & Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 21–29. https://doi.org/10.1177/2372732214549471

Deci, E., & Ryan, R. (2002). Self-determination theory: A dialectical framework for understanding sociocultural influences on student motivation. In E. Deci & R. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). University Rochester Press.

Decristan, J., Kunter, M., & Fauth, B. (2022). Die Bedeutung individueller Merkmale und konstruktiver Unterstützung der Lehrkraft für die soziale Integration von Schülerinnen und Schülern im Mathematikunterricht der Sekundarstufe. *Zeitschrift für Pädagogische Psychologie*, *36*(1–2), 85–100. https://doi.org/10.1024/1010-0652/a000329

Decristan, J., Kunter, M., Fauth, B., Büttner, G., Hardy, I., & Hertel, S. (2016). What role does instructional quality play for elementary school children's science competence? A focus on students at risk. *Journal for Educational Research Online*, *8*(1), 66–89. https://doi.org/10.25656/01:12032

Dickhäuser, O., & Meyer, W.-U. (2006). Gender differences in young children's math ability attributions. *Psychology Science*, *48*, 3–16.

Dietrich, J., Dicke, A.-L., Kracke, B., & Noack, P. (2015). Teacher support and its influence on students' intrinsic value and effort: Dimensional comparison effects across subjects. *Learning and Instruction*, *39*, 45–54. https://doi.org/10.1016/j.learninstruc.2015.05.007

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively?: A meta-analysis on self-regulation training programmes. *Educational Research Review*, *3*(2), 101–129. https://doi.org/10.1016/j.edurev.2008.02.003

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, *17*(2), 245–260. https://doi.org/10.1023/A:1019637632584

dos Santos, E., Albahari, A., Díaz, S., & Freitas, E. (2022). 'Science and technology as feminine': Raising awareness about and reducing the gender gap in STEM careers. *Journal of Gender Studies*, *31*, 505–518. https://doi.org/10.1080/09589236.2021.1922272

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735

Eccles, J. (2015). Gendered socialization of STEM interests in the family. *International Journal of Gender, Science and Technology*, *7*(2), 116–132. Retrieved from https://genderandset.open.ac.uk/index.php/genderandset/article/view/419

Eccles, J., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage–environment fit on young adolescents' experiences in schools and in families (1993). In J. M. Notterman (Ed.), *The evolution of psychology: Fifty years of the American Psychologist* (pp. 475–501). American Psychological Association. https://doi.org/10.1037/10254-034

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly*, *18*(4), 585–609. https://doi.org/10.1111/j.1471-6402.1994.tb01049.x

Eccles, J. S., & Wang, M.-T. (2016). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development*, *40*(2), 100–106. https://doi.org/10.1177/0165025415616201

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*(3), 830–847. https://doi.org/10.2307/1131221

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, *69*(1), 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127. https://doi.org/10.1037/a0018053

Encinas-Martín, M., & Cherian, M. (2023). *Gender, education and skills: The persistence of gender gaps in education and skills*. OECD Publishing. https://doi.org/10.1787/34680dd5-en

English, L. D. (2016). STEM education K-12: Perspectives on integration. *International Journal of STEM Education*, *3*(1), 3. https://doi.org/10.1186/s40594-016-0036-1

Epstein, C. F. (1970). *Woman's Place; Options and Limits in Professional Careers*. University of California Press.

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education*, *67*, 156–167. https://doi.org/10.1016/j.compedu.2013.02.019

European Commission. (2021). *She Figures 2021: Gender in research and innovation : Statistics and indicators.* Publications Office. https://data.europa.eu/doi/10.2777/06090

European Commission. (2022). *Bridging the gender gap in STEM: Strengthening opportunities for women in research and innovation*. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/774922

European Institute for Gender Equality. (2021). *Concepts and definitions*. https://eige.europa.eu/gender-mainstreaming/glossary

Fatourou, P., Papageorgiou, Y., & Petousi, V. (2019). Women are needed in STEM: European policies and incentives. *Communications of the ACM*, *62*(4), 52–52. https://doi.org/10.1145/3312565

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, *29*, 1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001

Federici, R. A., & Skaalvik, E. M. (2013). Students' perceptions of emotional and instrumental teacher support: Relations with motivational and emotional responses. *International Education Studies*, *7*(1), 21–36. https://doi.org/10.5539/ies.v7n1p21

Fisher, R. J., & Katz, J. E. (2000). Social-desirability bias and the validity of self-reported values. *Psychology & Marketing*, *17*(2), 105–120. https://doi.org/10.1002/(SICI)1520-6793(200002)17:2<105::AID-MAR3>3.0.CO;2-9

Fredricks, J. A., Hofkens, T., Wang, M.-T., Mortenson, E., & Scott, P. (2018). Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *Journal of Research in Science Teaching*, *55*(2), 271–298. https://doi.org/10.1002/tea.21419

Freedman-Doan, C., Wigfield, A., Eccles, J. S., Blumenfeld, P., Arbreton, A., & Harold, R. D. (2000). What am I best at? Grade and gender differences in children's beliefs about ability improvement. *Journal of Applied Developmental Psychology*, *21*(4), 379–402. https://doi.org/10.1016/S0193-3973(00)00046-0

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*(2), 507–537. https://doi.org/10.1111/j.1532-7795.2010.00645.x

Fulmer, S. M., & Frijters, J. C. (2009). A review of self-report and alternative approaches in the measurement of student motivation. *Educational Psychology Review*, *21*(3), 219–246. https://doi.org/10.1007/s10648-009-9107-x

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, *95*, 148–162. https://doi.org/10.1037/0022-0663.95.1.148

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences*, *47*, 182–193. https://doi.org/10.1016/j.lindif.2016.01.002

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B., & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology*, *51*, 1226–1240. https://doi.org/10.1037/dev0000028

Gaspard, H., & Lauermann, F. (2021). Emotionally and motivationally supportive classrooms: A state-trait analysis of lesson- and classroom-specific variation in teacher- and student-reported teacher enthusiasm and student engagement. *Learning and Instruction*, *75*, 101494. https://doi.org/10.1016/j.learninstruc.2021.101494

Ghasemi, E., & Burley, H. (2019). Gender, affect, and math: A cross-national meta-analysis of Trends in International Mathematics and Science Study 2015 outcomes. *Large-Scale Assessments in Education*, *7*(1), 10. https://doi.org/10.1186/s40536-019-0078-1

Goldman, A. D., & Penner, A. M. (2016). Exploring international gender differences in mathematics self-concept. *International Journal of Adolescence and Youth*, *21*(4), 403–418. https://doi.org/10.1080/02673843.2013.847850

Göllner, R., Fauth, B., & Wagner, W. (2021). Student ratings of teaching quality dimensions: Empirical findings and future directions. In Rollet, W., Bijlsma, H., & Röhl, S. (Eds.), *Student feedback on teaching in schools* (pp. 111–122). Springer, Cham.

González-Pérez, S., Mateos de Cabo, R., & Sáinz, M. (2020). Girls in STEM: Is it a female role-model thing? *Frontiers in Psychology*, *11*. https://www.frontiersin.org/articles/10.3389/fpsyg.2020.02204

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*, 700–717. https://doi.org/10.1037/a0026659

Green, J., Martin, A. J., & Marsh, H. W. (2007). Motivation and engagement in English, mathematics and science high school subjects: Towards an understanding of multidimensional domain specificity. *Learning and Individual Differences*, *17*(3), 269–279. https://doi.org/10.1016/j.lindif.2006.12.003

Gumaelius, L., Almqvist, M., Árnadóttir, A., Axelsson, A., Conejero, J. A., García-Sabater, J. P., Klitgaard, L., Kozma, C., Maheut, J., Marin-Garcia, J., Mickos, H., Nilsson, P.-O., Norén, A., Pinho-Lopes, M., Prenzel, M., Ray, J., Roxå, T., & Voss, M. (2016). Outreach initiatives operated by universities for increasing interest in science and technology. *European Journal of Engineering Education*, *41*(6), 589–622. https://doi.org/10.1080/03043797.2015.1121468

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*(3), 153–166. https://doi.org/10.1007/s11199-011-9996-2

Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology*, *51*, 1163–1176. https://doi.org/10.1037/a0039440

Gutman, L. M., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people. A literature review. In *Education Endowment Foundation: London, UK.* [Report]. Education Endowment Foundation. https://educationendowmentfoundation.org.uk/evidence-summaries/evidence-reviews/essential-life-skills/

Hammond, A., Rubiano Matulevich, E., Beegle, K., & Kumaraswamy, S. K. (2020). *The equality equation*. World Bank. https://doi.org/10.1596/34317

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, *76*(5), 949–967. https://doi.org/10.1111/j.1467-8624.2005.00889.x

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, *66*(2), 99–136. https://doi.org/10.3102/00346543066002099

Havik, T., & Westergård, E. (2020). Do teachers matter? Students' perceptions of classroom interactions and student engagement. *Scandinavian Journal of Educational Research*, *64*(4), 488–507. https://doi.org/10.1080/00313831.2019.1577754

Hecht, C. A., Priniski, S. J., & Harackiewicz, J. M. (2019). Understanding long-term effects of motivation interventions in a changing world. *Advances in Motivation and Achievement: A Research Annual*, *20*, 81–98. https://doi.org/10.1108/S0749-742320190000020005

Heckman, J. J., & Kautz, T. (2013). *Fostering and measuring skills: Interventions that improve character and cognition* (Working Paper 19656). National Bureau of Economic Research. https://doi.org/10.3386/w19656

Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, *6*(3), 203–217. https://doi.org/10.1037/1082-989X.6.3.203

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, *9*(4), 426–445. https://doi.org/10.1037/1082-989X.9.4.426

Heilbronner, N. N. (2012). The STEM pathway for women: What has changed? *Gifted Child Quarterly*, *57*(1), 39–55. https://doi.org/10.1177/0016986212460085

Herbert, J., & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of Applied Developmental Psychology*, *26*(3), 276–295. https://doi.org/10.1016/j.appdev.2005.02.007

Heyder, A., Steinmayr, R., & Kessels, U. (2019). Do teachers' beliefs about math aptitude and brilliance explain gender differences in children's math ability self-concept? *Frontiers in Education*, *4*. https://www.frontiersin.org/articles/10.3389/feduc.2019.00034

Hidi, S., & Harackiewicz, J. M. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research*, *70*, 151–179. https://doi.org/10.2307/1170660

Hornstra, L., van der Veen, I., & Peetsma, T. (2016). Domain-specificity of motivation: A longitudinal study in upper primary school. *Learning and Individual Differences*, *51*, 167–178. https://doi.org/10.1016/j.lindif.2016.08.012

Hsieh, P., Pei-H., Acee, T., Chung, W.-H., Hsieh, Y.-P., Kim, H., Thomas, G. D., You, J., Levin, J. R., & Robinson, D. H. (2005). Is educational intervention research on the decline? *Journal of Educational Psychology*, *97*(4), 523–529. https://doi.org/10.1037/0022-0663.97.4.523

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, *28*(1), 1–35. https://doi.org/10.1007/s10212-011-0097-y

Hübner, N., Wille, E., Cambria, J., Oschatz, K., Nagengast, B., & Trautwein, U. (2017). Maximizing gender equality by minimizing course choice options? Effects of obligatory coursework in math on gender differences in STEM. *Journal of Educational Psychology*, *109*(7), 993–1009. https://doi.org/10.1037/edu0000183

Hughes, J. N., & Chen, Q. (2011). Reciprocal effects of student–teacher and student–peer relatedness: Effects on academic self efficacy. *Journal of Applied Developmental Psychology*, *32*(5), 278–287. https://doi.org/10.1016/j.appdev.2010.03.005

Hulleman, C. S., & Barron, K. E. (2016). Motivation interventions in education: Bridging theory, research, and practice. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology, 3rd ed* (pp. 160–171). Routledge/Taylor & Francis Group.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494–495. https://doi.org/10.1126/science.1160364

Hyde, J. S., & Linn, M. C. (2006). Gender similarities in mathematics and science. *Science*, *314*(5799), 599–600. https://doi.org/10.1126/science.1132154

Jacobs, J. E., Davis-Kean, P., Bleeker, M., Eccles, J. S., & Malanchuk, O. (2005). "I can, but I don't want to": The impact of parents, interests, and activities on gender differences in math. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 246–263). Cambridge University Press. https://doi.org/10.1017/CBO9780511614446.013

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

Jansen, M., Schroeders, U., & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences*, *30*, 11–21. https://doi.org/10.1016/j.lindif.2013.12.003

Kadijevich, D. (2000). Gender differences in computer attitude among ninth-grade students. *Journal of Educational Computing Research*, *22*(2), 145–154. https://doi.org/10.2190/K4U2-PWQG-RE8L-UV90

Kenny, D. A., & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*(1), 201–210. https://doi.org/10.1037/0033-2909.96.1.201

Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2008). An aptitude–treatment interaction approach to writing-to-learn. *Learning and Instruction*, *18*(4), 379–390. https://doi.org/10.1016/j.learninstruc.2007.07.004

Kim, C., & Pekrun, R. (2014). Emotions and motivation in learning and performance. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 65–75). Springer, New York, NY. https://doi.org/10.1007/978-1-4614-3185-5_6

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*(4), 457–474. https://doi.org/10.1007/BF02296338

Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras Study. Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik (Ed.), *Klieme, E; Pauli, C; Reusser, K (2009). The Pythagoras Study. Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In: Janik, T. The power of video studies in investigating teaching and learning in the classroom. Münster: Waxmann, 137-160.* (pp. 137–160). Waxmann. https://www.zora.uzh.ch/id/eprint/29452/

Koca, F. (2016). Motivation to learn and teacher-student relationship. *Journal of International Education and Leadership*, *6*(2), 1–20. https://eric.ed.gov/?id=EJ1135209

Kokotsaki, D., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools.*, *19*(3), 267–277. https://doi.org/10.1177/1365480216659733

Köller, O., Baumert, J., & Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education*, *32*(5), 448–470. https://doi.org/10.2307/749801

Kollmayer, M., Schober, B., & Spiel, C. (2018). Gender stereotypes in education: Development, consequences, and interventions. *European Journal of Developmental Psychology*, *15*(4), 361–377. https://doi.org/10.1080/17405629.2016.1193483

Kong, S., Carroll, K., Lundberg, D., Omura, P., & Lepe, B. (2020). Reducing gender bias in STEM. *MIT Science Policy Review*, *1*, 55–63. https://doi.org/10.38105/spr.11kp6lqr0a

Korhonen, J., Tapola, A., Linnanmäki, K., & Aunio, P. (2016). Gendered pathways to educational aspirations: The role of academic self-concept, school burnout, achievement and interest in mathematics and reading. *Learning and Instruction*, *46*, 21–33. https://doi.org/10.1016/j.learninstruc.2016.08.006

Kuchynka, S. L., Eaton, A., & Rivera, L. M. (2022). Understanding and addressing gender-based inequities in STEM: Research synthesis and recommendations for U.S. K-12 education. *Social Issues and Policy Review*, *16*(1), 252–288. https://doi.org/10.1111/sipr.12087

Kurtz-Costes, B., Rowley, S. J., Harris-Britt, A., & Woods, T. A. (2008). Gender stereotypes about mathematics and science and self-perceptions of ability in late childhood and early adolescence. *Merrill-Palmer Quarterly*, *54*(3), 386–409. https://doi.org/10.1353/mpq.0.0001

La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The classroom assessment scoring system: Findings from the prekindergarten year. *The Elementary School Journal*, *104*(5), 409–426. https://doi.org/10.1086/499760

LaForce, M., Zuo, H., Ferris, K., & Noble, E. (2019). Revisiting race and gender differences in STEM: Can inclusive STEM high schools reduce gaps? *European Journal of STEM Education*, *4*(1), 8. https://doi.org/10.20897/ejsteme/5840

Lauermann, F., Tsai, Y.-M., & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy–value theory of achievement-related behaviors. *Developmental Psychology*, *53*(8), 1540–1559. https://doi.org/10.1037/dev0000367

Lazarides, R., Gaspard, H., & Dicke, A.-L. (2019). Dynamics of classroom motivation: Teacher enthusiasm and the development of math interest and teacher support. *Learning and Instruction*, *60*, 126–137. https://doi.org/10.1016/j.learninstruc.2018.01.012

Lazarides, R., & Ittel, A. (2013). Mathematics interest and achievement: What role do perceived parent and teacher support play? A longitudinal analysis. *International Journal of Gender, Science and Technology*, *5*(3), 207–231. Retrieved from https://genderandset.open.ac.uk/index.php/genderandset/article/view/301

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Ledgerwood, A., & Shrout, P. E. (2011). The tradeoff between accuracy and precision in latent variable models of mediation processes. *Journal of Personality and Social Psychology*, *101*(6), 1174–1188. https://doi.org/10.1037/a0024776

Lei, H., Cui, Y., & Chiu, M. M. (2018). The relationship between teacher support and students' academic emotions: A meta-analysis. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02288

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology*, *97*(2), 184–196. https://doi.org/10.1037/0022-0663.97.2.184

Lepper, M. R., & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25 Years of research on intrinsic versus extrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation* (pp. 257–307). Academic Press. https://doi.org/10.1016/B978-012619070-0/50032-5

Lerkkanen, M.-K., Kiuru, N., Pakarinen, E., Viljaranta, J., Poikkeus, A.-M., Rasku-Puttonen, H., Siekkinen, M., & Nurmi, J.-E. (2012). The role of teaching practices in the development of children's interest in reading and mathematics in kindergarten. *Contemporary Educational Psychology*, *37*(4), 266–279. https://doi.org/10.1016/j.cedpsych.2011.03.004

Levine, M., Serio, N., Radaram, B., Chaudhuri, S., & Talbert, W. (2015). Addressing the STEM gender gap by designing and implementing an educational outreach chemistry camp for middle school girls. *Journal of Chemical Education*, *92*(10), 1639–1644. https://doi.org/10.1021/ed500945g

Levine, S. C., & Pantoja, N. (2021). Development of children's math attitudes: Gender differences, key socializers, and intervention approaches. *Developmental Review*, *62*, 100997. https://doi.org/10.1016/j.dr.2021.100997

Liben, L. S. (2015). The STEM gender gap: The case for spatial interventions. *International Journal of Gender, Science and Technology*, *7*(2), 133–150. Retrieved from https://genderandset.open.ac.uk/index.php/genderandset/article/view/418

Lieberman, D. A., Bates, C. H., & So, J. (2009). Young children's learning with digital media. *Computers in the Schools*, *26*(4), 271–283. https://doi.org/10.1080/07380560903360194

Lindberg, S., Linkersdörfer, J., Ehm, J.-H., Hasselhorn, M., & Lonnemann, J. (2013). Gender differences in children's math self-concept in the first years of elementary school. *Journal of Education and Learning*, *2*(3), 1–8. https://doi.org/10.5539/jel.v2n3p1

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, *136*(6), 1123–1135. https://doi.org/10.1037/a0021276

Lin-Siegler, X., Dweck, C. S., & Cohen, G. L. (2016). Instructional interventions that motivate classroom learning. *Journal of Educational Psychology*, *108*(3), 295. https://doi.org/10.1037/edu0000124

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, *19*(6), 527–537. https://doi.org/10.1016/j.learninstruc.2008.11.001

Little, T. D., Bovaird, J. A., & Widaman, K. F. (2006). On the Merits of Orthogonalizing Powered and Product Terms: Implications for Modeling Interactions Among Latent Variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *13*(4), 497–519. https://doi.org/10.1207/s15328007sem1304_1

Liu, R.-D., Zhen, R., Ding, Y., Liu, Y., Wang, J., Jiang, R., & Xu, L. (2018). Teacher support and math engagement: Roles of academic self-efficacy and positive emotions. *Educational Psychology*, *38*(1), 3–16. https://doi.org/10.1080/01443410.2017.1359238

Lubienski, S. T., Robinson, J. P., Crane, C. C., & Ganley, C. M. (2013). Girls' and boys' mathematics achievement, affect, and experiences: Findings from ECLS-K. *Journal for Research in Mathematics Education*, *44*(4), 634–645. https://doi.org/10.5951/jresematheduc.44.4.0634

Lupart, J. L., Cannon, E., & Telfer, J. A. (2004). Gender differences in adolescent academic achievement, interests, values and life-role expectations. *High Ability Studies*, *15*(1), 25–42. https://doi.org/10.1080/1359813042000225320

Ma, L., Luo, H., & Xiao, L. (2021). Perceived teacher support, self-concept, enjoyment and achievement in reading: A multilevel mediation model based on PISA 2018. *Learning and Individual Differences*, *85*, 101947. https://doi.org/10.1016/j.lindif.2020.101947

Malecki, C. K., & Demaray, M. K. (2006). Social support as a buffer in the relationship between socioeconomic status and academic performance. *School Psychology Quarterly*, *21*, 375–395. https://doi.org/10.1037/h0084129

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among U.S. students. *Science Education*, *95*(5), 877–907. https://doi.org/10.1002/sce.20441

Marsh, H. W., Martin, A. J., & Debus, R. (2001). Individual differences in verbal and math self-perceptions: One factor, two factors, or does it depend on the construct? In R. J. Riding & S. G. Rayner (Eds.), *Self perception* (pp. 149–170). Ablex Publishing.

Marsh, H. W., Wen, Z., Hau, K.-T., & Nagengast, B. (2013). Structural equation models of latent interaction and quadratic effects. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course*. IAP Information Age Publishing.

Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research*, *79*(1), 327–365. https://doi.org/10.3102/0034654308325583

Martin, A. J., Mansour, M., & Malmberg, L.-E. (2020). What factors influence students' real-time motivation and engagement? An experience sampling study of high school students using mobile technology. *Educational Psychology*, *40*(9), 1113–1135. https://doi.org/10.1080/01443410.2018.1545997

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, *52*(4), 1119–1134. https://doi.org/10.2307/1129498

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D. M., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills.

Child Development, *79*(3), 732–749. https://doi.org/10.1111/j.1467-8624.2008.01154.x

Maslowsky, J., Jager, J., & Hemken, D. (2015). Estimating and interpreting latent variable interactions: A tutorial for applying the latent moderated structural equations method. *International Journal of Behavioral Development*, *39*(1), 87–96. https://doi.org/10.1177/0165025414552301

Master, A. (2021). Gender stereotypes influence children's STEM motivation. *Child Development Perspectives*, *15*(3), 203–210. https://doi.org/10.1111/cdep.12424

Master, A., & Meltzoff, A. N. (2020). Cultural stereotypes and sense of belonging contribute to gender gaps in STEM. *International Journal of Gender, Science and Technology*, *12*(1), 152–198. Retrieved from https://genderandset.open.ac.uk/index.php/genderandset/article/view/674

Master, A., Meltzoff, A. N., & Cheryan, S. (2021). Gender stereotypes about interests start early and cause gender disparities in computer science and engineering. *Proceedings of the National Academy of Sciences*, *118*(48), e2100030118. https://doi.org/10.1073/pnas.2100030118

Mau, W.-C. (2003). Factors that influence persistence in science and engineering career aspirations. *The Career Development Quarterly*, *51*(3), 234–243. https://doi.org/10.1002/j.2161-0045.2003.tb00604.x

McDonald, C. V. (2016). STEM education: A review of the contribution of the disciplines of science, technology, engineering and mathematics. *Science Education International*, *27*(4), 530–569.

McFarland, L., Murray, E., & Phillipson, S. (2016). Student–teacher relationships and student self-concept: Relations with teacher and student gender. *Australian Journal of Education*, *60*(1), 5–25. https://doi.org/10.1177/0004944115626426

McGuire, L., Hoffman, A. J., Mulvey, K. L., Hartstone-Rose, A., Winterbottom, M., Joy, A., Law, F., Balkwill, F., Burns, K. P., Butler, L., Drews, M., Fields, G., Smith, H., & Rutland, A. (2022). Gender stereotypes and peer selection in STEM domains among children and adolescents. *Sex Roles*, *87*(9), 455–470. https://doi.org/10.1007/s11199-022-01327-9

Meece, J. L., Glienke, B. B., & Burg, S. (2006). Gender and motivation. *Journal of School Psychology*, *44*(5), 351–373. https://doi.org/10.1016/j.jsp.2006.04.004

Michel, Y. A., Steinmayr, R., Frenzel, A. C., & Ziegler, M. (2022). Unpacking domain-specific achievement motivation: The role of contextualising items for test-criterion correlations. *Educational Psychology*, *42*(4), 501–525. https://doi.org/10.1080/01443410.2020.1713303

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of U.S. Draw-A-Scientist studies. *Child Development*, *89*(6), 1943–1955. https://doi.org/10.1111/cdev.13039

Miller, P. H., Slawinski Blessing, J., & Schwartz, S. (2006). Gender differences in high-school students' views about science. *International Journal of Science Education*, *28*(4), 363–381. https://doi.org/10.1080/09500690500277664

Morgenroth, T., Ryan, M. K., & Peters, K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology*, *19*, 465–483. https://doi.org/10.1037/gpr0000059

Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology*, *25*(1), 3–53. https://doi.org/10.1006/ceps.1999.1019

Murphy, S., MacDonald, A., Wang, C. A., & Danaia, L. (2019). Towards an understanding of STEM engagement: A review of the literature on motivation and academic emotions. *Canadian Journal of Science, Mathematics and Technology Education*, *19*(3), 304–320. https://doi.org/10.1007/s42330-019-00054-w

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educational Research and Evaluation*, *21*(4), 343–370. https://doi.org/10.1080/13803611.2015.1057161

O'Dea, R. E., Lagisz, M., Jennions, M. D., & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-06292-0

OECD. (2013). Mathematics self-beliefs and participation in mathematics-related activities. In *PISA 2012 Results: Ready to Learn (Volume III)*. OECD. https://www.oecd.org/pisa/keyfindings/PISA2012-Vol3-Chap4.pdf

OECD. (2015). *The ABC of gender equality in education*. OECD Publishing. https://www.oecd-ilibrary.org/content/publication/9789264229945-en

OECD. (2016). *PISA 2015 results (Volume I): Excellence and equity in education*. OECD. https://doi.org/10.1787/9789264266490-en

OECD. (2020a). Girls' and boys' performance in PISA. In *PISA 2018 Results (Volume II): Where All Students Can Succeed* (pp. 141–155). OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en

OECD. (2020b). *Global Teaching InSights: A Video Study of Teaching*. Organisation for Economic Co-operation and Development. https://www.oecd-ilibrary.org/education/global-teaching-insights_20d6f36b-en

OECD. (2020c). *Global Teaching Insights: A Video Study of Teaching. [Dataset]*. https://www.oecd.org/education/school/global-teaching-insights-technical-documents.htm

Olsson, M., & Martiny, S. E. (2018). Does exposure to counterstereotypical role models influence girls' and women's gender stereotypes and career choices? A review of social psychological research. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02264

Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 294–315). Cambridge University Press.

Parker, P., Nagy, G., Trautwein, U., & Lüdtke, O. (2014). Predicting career aspirations and university majors from academic ability and self-concept: A longitudinal application of the internal–external frame of reference model. In I. Schoon & J. S. Eccles (Eds.), *Gender differences in aspirations and attainment* (1st ed., pp. 224–246). Cambridge University Press. https://doi.org/10.1017/CBO9781139128933.015

Päßler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for women and men? *Journal of Career Assessment*, *20*(4), 479–496. https://doi.org/10.1177/1069072712450009

Patall, E. A., Steingut, R. R., Freeman, J. L., Pituch, K. A., & Vasquez, A. C. (2018). Gender disparities in students' motivational experiences in high school science classrooms. *Science Education*, *102*(5), 951–977. https://doi.org/10.1002/sce.21461

Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, *99*, 83–98. https://doi.org/10.1037/0022-0663.99.1.83

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–105. https://doi.org/10.1207/S15326985EP3702_4

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259–282). Springer.

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: Conceptualizing, measuring, and improving the capacity of classroom interactions. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 365–386). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_17

Pintrich, P. R. (2003a). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, *95*(4), 667–686. https://doi.org/10.1037/0022-0663.95.4.667

Pintrich, P. R. (2003b). Motivation and classroom learning. In *Handbook of psychology: Educational psychology, Vol. 7.* (pp. 103–122). John Wiley & Sons Inc.

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM*, *50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4

Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, *31*, 2–12. https://doi.org/10.1016/j.learninstruc.2013.12.002

Preckel, F., Goetz, T., Pekrun, R., & Kleine, M. (2008). Gender differences in gifted and average-ability students: Comparing girls' and boys' achievement, self-concept, interest, and motivation in mathematics. *Gifted Child Quarterly*, *52*(2), 146–159. https://doi.org/10.1177/0016986208315834

Rathvon, N. (2008). *Effective school interventions, second edition: Evidence-based strategies for improving student outcomes*. Guilford Press.

Reilly, D., Neumann, D. L., & Andrews, G. (2019). Investigating gender differences in mathematics and science: Results from the 2011 Trends in Mathematics and Science Survey. *Research in Science Education*, *49*(1), 25–50. https://doi.org/10.1007/s11165-017-9630-6

Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time. *American Educational Research Journal*, *49*(6), 1048–1073. https://doi.org/10.3102/0002831211435229

Roberts, K. (2014). *Engaging more women and girls in mathematics and STEM fields: The international evidence*. https://doi.org/10.13140/2.1.3947.8402

Robertson, J. (2000). Is attribution training a worthwhile classroom intervention for K–12 students with learning difficulties? *Educational Psychology Review*, *12*(1), 111–134. https://doi.org/10.1023/A:1009089118008

Rodger, S., Murray, H. G., & Cummings, A. L. (2007). Effects of teacher clarity and student anxiety on student outcomes. *Teaching in Higher Education*, *12*(1), 91–104. https://doi.org/10.1080/13562510601102255

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, *81*(4), 493–529. https://doi.org/10.3102/0034654311421793

Rosenzweig, E. Q., & Wigfield, A. (2016). STEM motivation interventions for adolescents: A promising start, but further to go. *Educational Psychologist*, *51*(2), 146–163. https://doi.org/10.1080/00461520.2016.1154792

Rottinghaus, P. J., Lindley, L. D., Green, M. A., & Borgen, F. H. (2002). Educational aspirations: The contribution of personality, self-efficacy, and interests. *Journal of Vocational Behavior*, *61*(1), 1–19. https://doi.org/10.1006/jvbe.2001.1843

Rozek, C. S., Hyde, J. S., Svoboda, R. C., Hulleman, C. S., & Harackiewicz, J. M. (2015). Gender differences in the effects of a utility-value intervention to help parents motivate adolescents in mathematics and science. *Journal of Educational Psychology*, *107*(1), 195–206. https://doi.org/10.1037/a0036981

Rueger, S. Y., Malecki, C. K., & Demaray, M. K. (2008). Relationship between multiple sources of perceived social support and psychological and academic adjustment in early adolescence: Comparisons across gender. *Journal of Youth and Adolescence*, *39*(1), 47. https://doi.org/10.1007/s10964-008-9368-6

Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal*, *38*(2), 437–460. https://doi.org/10.3102/00028312038002437

Ryan, R. M., Stiller, J. D., & Lynch, J. H. (1994). Representations of relationships to teachers, parents, and friends as predictors of academic motivation and self-esteem. *The Journal of Early Adolescence*, *14*(2), 226–249. https://doi.org/10.1177/027243169401400207

Sánchez-Tapia, I., & Alam, A. (2020). Towards an equal future: Reimagining girls' education through STEM. In *UNICEF*. UNICEF. https://eric.ed.gov/?id=ED627513

Scherer, R., & Nilsen, T. (2016). The relations among school climate, instructional quality, and achievement motivation in mathematics. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (Vol. 2, pp. 51–166). Springer International Publishing. https://doi.org/10.1007/978-3-319-41252-8

Schoon, I., & Eccles, J. S. (Eds.). (2014). *Gender differences in aspirations and attainment: A life course perspective*. Cambridge University Press.

Schunk, D. H., & Ertmer, P. A. (2000). Self-regulation and academic learning: Self-efficacy enhancing interventions. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 631–649). Academic Press. https://doi.org/10.1016/B978-012109890-2/50048-2

Seidel, T., Rimmele, R., & Prenzel, M. (2005). Clarity and coherence of lesson goals as a scaffold for student learning. *Learning and Instruction*, *15*(6), 539–556. https://doi.org/10.1016/j.learninstruc.2005.08.004

Shin, J. E. L., Levy, S. R., & London, B. (2016). Effects of role model exposure on STEM and non-STEM student engagement. *Journal of Applied Social Psychology*, *46*(7), 410–427. https://doi.org/10.1111/jasp.12371

Shumow, L., & Schmidt, J. (2013). Academic grades and motivation in high school science classrooms among male and female students: Associations with teachers' characteristics, beliefs, and practices. In R. Haumann & G. Zimmer (Eds.), *Handbook of academic performance: Predictors, learning strategies and influences of gender* (pp. 53–72). Nova Science Publishers, Inc.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, *70*(1), 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Simon, R. A., Aulls, M. W., Dedic, H., Hubbard, K., & Hall, N. C. (2015). Exploring student persistence in STEM programs: A motivational model. *Canadian Journal of Education*, *38*(1), 1–27.

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, *42*(1), 70. https://doi.org/10.1037/0012-1649.42.1.70

Skaalvik, E. M., Federici, R. A., & Klassen, R. M. (2015). Mathematics achievement and self-efficacy: Relations with motivation for mathematics. *International Journal of Educational Research*, *72*, 129–136. https://doi.org/10.1016/j.ijer.2015.06.008

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research*, *79*(4), 1391–1466. https://doi.org/10.3102/0034654309341374

Smith, E., & Farkas, G. (2022). Gender and mathematics achievement: The role of gender stereotypical beliefs of classroom peers. *European Sociological Review*, *39*(2), 161–176 https://doi.org/10.1093/esr/jcac043

Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist*, *27*(1), 5–32. https://doi.org/10.1207/s15326985ep2701_3

Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology*, *43*(1), 583–626. https://doi.org/10.1146/annurev.ps.43.020192.003055

Snyder, K. E., Fong, C. J., Painter, J. K., Pittard, C. M., Barr, S. M., & Patall, E. A. (2019). Interventions for academically underachieving students: A systematic review and meta-analysis. *Educational Research Review*, *28*, 100294. https://doi.org/10.1016/j.edurev.2019.100294

Soland, J., Kuhfeld, M., Wolk, E., & Bi, S. (2019). Examining the state-trait composition of social-emotional learning constructs: Implications for practice, policy, and evaluation. *Journal of Research on Educational Effectiveness*, *12*(3), 550–577. https://doi.org/10.1080/19345747.2019.1615158

Steinke, J. (2017). Adolescent girls' STEM identity formation and media images of STEM professionals: Considering the influence of contextual cues. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00716

Stoet, G., & Geary, D. C. (2018). The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychological Science*, *29*(4), 581–593. https://doi.org/10.1177/0956797617741719

Stringer, K., Mace, K., Clark, T., & Donahue, T. (2020). STEM focused extracurricular programs: Who's in them and do they change STEM identity and motivation? *Research in Science & Technological Education*, *38*(4), 507–522. https://doi.org/10.1080/02635143.2019.1662388

Su, R., & Rounds, J. (2015). All STEM fields are not created equal: People and things interests explain gender disparities across STEM fields. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00189

Tang, M., Pan, W., & Newmeyer, M. D. (2008). Factors influencing high school students' career aspirations. *Professional School Counseling*, *11*(5). https://doi.org/10.1177/2156759X0801100502

Tao, Y., Meng, Y., Gao, Z., & Yang, X. (2022). Perceived teacher support, student engagement, and academic achievement: A meta-analysis. *Educational Psychology*, *42*(4), 401–420. https://doi.org/10.1080/01443410.2022.2033168

Taylor, M. G. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development*, *67*(4), 1555–1571. https://doi.org/10.2307/1131718

Tenenbaum, H. R. (2009). 'You'd be good at that': Gender patterns in parent-child talk about courses. *Social Development*, *18*(2), 447–463. https://doi.org/10.1111/j.1467-9507.2008.00487.x

Thomas, H., & Kail, R. (1991). Sex differences in speed of mental rotation and the X-linked genetic hypothesis. *Intelligence*, *15*(1), 17–32. https://doi.org/10.1016/0160-2896(91)90020-E

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, *92*(1), 144–151. https://doi.org/10.1037/0022-0663.92.1.144

Tollefson, N. (2000). Classroom applications of cognitive theories of motivation. *Educational Psychology Review*, *12*(1), 63–83. https://doi.org/10.1023/A:1009085017100

UNESCO. (2016). *Closing the gender gap in STEM: Drawing more girls and women into science, technology, engineering and mathematics* (pp. 1–4). UNESCO Office Bangkok. https://unesdoc.unesco.org/ark:/48223/pf0000245717

Unrau, N. J., Rueda, R., Son, E., Polanin, J. R., Lundeen, R. J., & Muraszewski, A. K. (2018). Can reading self-efficacy be modified? A meta-analysis of the impact of interventions on reading self-efficacy. *Review of Educational Research*, *88*(2), 167–204. https://doi.org/10.3102/0034654317743199

Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology*, *44*(5), 331–349. https://doi.org/10.1016/j.jsp.2006.04.003

van den Hurk, A., Meelissen, M., & van Langen, A. (2019). Interventions in education to prevent STEM pipeline leakage. *International Journal of Science Education*, *41*(2), 150–164. https://doi.org/10.1080/09500693.2018.1540897

van Eerde, W., & Klingsieck, K. B. (2018). Overcoming procrastination? A meta-analysis of intervention studies. *Educational Research Review*, *25*, 73–85. https://doi.org/10.1016/j.edurev.2018.09.002

van Tuijl, C., & van der Molen, J. H. W. (2016). Study choice and career development in STEM fields: An overview and integration of the research. *International Journal of Technology and Design Education*, *26*(2), 159–183. https://doi.org/10.1007/s10798-015-9308-1

Vekiri, I. (2010). Boys' and girls' ICT beliefs: Do teachers matter? *Computers & Education*, *55*(1), 16–23. https://doi.org/10.1016/j.compedu.2009.11.013

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*, 1174–1204. https://doi.org/10.1037/a0036620

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, *23*, 73–82. https://doi.org/10.1177/0963721413512856

Walton, G. M., & Wilson, T. D. (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological Review*, *125*(5), 617–655. https://doi.org/10.1037/rev0000115

Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review, 33*(4), 304–340. https://doi.org/10.1016/j.dr.2013.08.001

Wang, M.-T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review*, *29*(1), 119–140. https://doi.org/10.1007/s10648-015-9355-x

Wang, M.-T., Degol, J., & Ye, F. (2015). Math achievement is important, but task values are critical, too: Examining the intellectual and motivational factors leading to gender disparities in STEM careers. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00036

Wang, S.-K., & Reeves, T. C. (2007). The effects of a web-based learning environment on student motivation in a high school earth science course. *Educational Technology Research and Development*, *55*(2), 169–192. https://doi.org/10.1007/s11423-006-9016-3

Ward, L. M., & Grower, P. (2020). Media and the development of gender role stereotypes. *Annual Review of Developmental Psychology*, *2*(1), 177–199. https://doi.org/10.1146/annurev-devpsych-051120-010630

Wasserman, T., & Wasserman, L. (2020). Motivation: State, trait, or both. In T. Wasserman & L. Wasserman (Eds.), *Motivation, effort, and the neural network model* (pp. 93–101). Springer International Publishing. https://doi.org/10.1007/978-3-030-58724-6_8

Watt, H. M. G., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., & Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from

Australia, Canada, and the United States. *Developmental Psychology*, *48*(6), 1594–1611. https://doi.org/10.1037/a0027838

What Works Clearing House. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_WWC-HandbookVer5_0-0-508.pdf

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development, Vol. 3, 6th ed* (pp. 933–1002). John Wiley & Sons, Inc.

Wigfield, A., & Wagner, A. L. (2005). Competence, motivation, and identity development during adolescence. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 222–239). Guilford Publications.

Wigfield, A., & Wentzel, K. R. (2007). Introduction to motivation at school: Interventions that work. *Educational Psychologist*, *42*(4), 191–196. https://doi.org/10.1080/00461520701621038

Wille, E., Stoll, G., Gfrörer, T., Cambria, J., Nagengast, B., & Trautwein, U. (2020). It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemporary Educational Psychology*, *61*, 101858. https://doi.org/10.1016/j.cedpsych.2020.101858

World Economic Forum. (2023). *Global Gender Gap Report 2023*. https://www3.weforum.org/docs/WEF_GGGR_2023.pdf

Yeager, D. S., & Dweck, C. S. (2020). What can be learned from growth mindset controversies? *American Psychologist*, *75*(9), 1269–1284. https://doi.org/10.1037/amp0000794

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, *81*, 267–301. https://doi.org/10.3102/0034654311405999

Yeung, A. S., & McInerney, D. M. (2005). Students' school motivation and aspiration over high school years. *Educational Psychology*, *25*(5), 537–554. https://doi.org/10.1080/01443410500046804

Appendix

Appendix A – Study I

Lesperance, K., Hofer, S., Retelsdorf, J., & Holzberger, D. (2022). Reducing gender differences in student motivational-affective factors: A meta-analysis of school-based interventions. *British Journal of Educational Psychology*, *92*(4), 1502-1536.

Received: 16 April 2020 | Accepted: 27 April 2022

DOI: 10.1111/bjep.12512

ARTICLE

the british psychological society
promoting excellence in psychology

Reducing gender differences in student motivational-affective factors: A meta-analysis of school-based interventions

Kaley Lesperance[1] | **Sarah Hofer**[2] | **Jan Retelsdorf**[3] | **Doris Holzberger**[1]

[1]Centre for International Student Assessment, Professorship for Research on Learning and Instruction, TUM School of Social Sciences and Technology, Technical University of Munich, Munich, Germany

[2]Faculty of Education and Educational Psychology, Ludwig Maximilian University of Munich, Munich, Germany

[3]Educational Psychology, Faculty of Education, University of Hamburg, Hamburg, Germany

Correspondence
Kaley Lesperance, Centre for International Student Assessment, Professorship for Research on Learning and Instruction, TUM School of Social Sciences and Technology, Technical University of Munich, Marsstr. 20-22, 80335 Munich, Germany.
Email: kaley.lesperance@tum.de

Funding information
Bundesministerium für Bildung und Forschung, Grant/Award Number: ZIB2022; The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany, Grant/Award Number: ZIB2022

Abstract

Background: Research shows that gender differences tend to exist in student motivational-affective factors in core subjects such as math, science or reading, where one gender is stereotypically disadvantaged.

Aims: This study aimed to investigate strategies that could reduce these gender differences by conducting a meta-analysis on school-based intervention studies that targeted student motivational-affective factors. We therefore evaluated whether interventions had differential effects for male and female students' motivational-affective factors in a given academic subject. We also evaluated potential moderator variables.

Method: After conducting a systematic database search and screening abstracts for inclusion, we synthesized 71 effect sizes from 20 primary studies. All included studies were conducted in science or mathematics-related subjects, which are stereotypically female-disadvantaged.

Results: While the interventions had significant positive effects for both genders, there was no statistically significant difference between the two genders with regard to the intervention effects on motivational-affective factors. However, the descriptive effect size for female students (g = .49) was far greater than for male students (g = .28). Moderator analyses showed no significant effects for grade level, intervention duration, or school subject, but there was a significant influence of intervention method used.

Conclusions: This study demonstrated that school-based interventions have positive effects on motivational-affective factors for both genders. It also provides evidence that interventions in subjects where female students are stereotypically disadvantaged may have greater effects for females than for males. Implications and suggestions for future research are discussed.

KEYWORDS

affect, gender differences, interventions, meta analysis, motivation, students

BACKGROUND

Male and female students often display differences in motivational-affective factors within educational contexts, with either one gender or the other being disadvantaged relative to the domain in question (Wigfield et al., 2002). Many constructs fall under the umbrella of motivational-affective factors. Murphy and Alexander (2000), for example, provide a categorization for motivational constructs. They differentiate between goals, interest, motivation (intrinsic and extrinsic), and self-schema (agency, attribution, self-competence, and self-efficacy). In turn, Pintrich (2003) points to expectancy, value, and affective variables as components of motivation. While in Pintrich's (2003) conceptualization, affective variables are regarded as a subcategory of motivation, variables with an emotional aspect such as enjoyment, anxiety, or boredom also have their own distinction in the literature (Pekrun et al., 2011). All of these factors have been shown to be strongly related to career choices, achievement, and performance outcomes in students (Kim & Pekrun, 2014; Möller et al., 2020; Pintrich, 2003; Wigfield & Eccles, 2000). Research has found that many of these factors strongly predict school performance and academic choices, often above and beyond IQ (Goetz & Hall, 2013; Köller et al., 2001; Parker et al., 2014; Steinmayr & Spinath, 2009). For example, in a review by Rosen et al. (2010), which examined 45 studies on motivation, 27 studies on self-efficacy, and 42 studies on academic self-concept, findings indicated that all of these measures were strongly connected to academic achievement in students from kindergarten to 12th grade. Large-scale studies on affective factors have also demonstrated that pleasant emotions such as enjoyment and pride are positively related to academic achievement, whereas unpleasant emotions such as anxiety are most often negatively related (Pekrun et al., 2002, 2017). Given their strong connection to important academic outcomes, gender differences in motivational-affective factors are concerning.

Gender differences in educational contexts

There is a plethora of evidence demonstrating that gender differences between many student motivational-affective factors exist in various subjects. Results from large-scale international studies such as the Program for International Student Assessment (PISA) provide an overview of student outcomes across many countries. In science, results from PISA 2015 revealed that, on average, across all 72 countries assessed, boys were more likely than girls to report higher intrinsic motivation for science, and greater interest and enjoyment in most science-related topics (OECD, 2016). These findings can be seen in smaller studies as well, with boys displaying more positive science attitudes and a higher likelihood of pursuing a science-related career (Jones et al., 2000; Miller et al., 2006; Weinburgh, 1995). A similar trend exists in mathematics, where boys tend to report higher self-efficacy and more positive attitudes, whereas girls display higher levels of math anxiety and are more likely to perceive math as a low-value subject (Else-Quest et al., 2010; OECD, 2013; Pajares, 2005). On the contrary, in the domain of reading, writing, and language arts, this trend is reversed, with males displaying lower reading and writing self-concepts, more negative attitudes towards reading, and low value beliefs for reading as a subject (Durik et al., 2006; Logan & Johnston, 2009; Marinak & Gambrell, 2010; OECD, 2019; Schleicher, 2019). Given these findings, it comes as no surprise that there continues to be large gaps between males and females in career goals and choices. In the areas of science, technology, engineering, and mathematics (STEM), females are extremely underrepresented in higher education and the labour market (Burke & Mattis, 2007; Dasgupta & Stout, 2014), a pattern that is already evident from career expectations and interest during the school years (Master et al., 2017; Sadler et al., 2012; Shapiro & Williams, 2012). On the other hand, careers in areas such as primary education and healthcare professions display much higher numbers of females than males (Hsu et al., 2010; OECD, 2014, 2019).

The role of gender stereotypes for the development of gender differences

Many research studies have investigated why these gaps between male and female students emerge in different subjects. Gender differences in educational contexts seem to arise as an individual grows and interacts with their environment, and gender stereotypes acquired from the social environment, such as from parents (Casad et al., 2015; Tiedemann, 2000), teachers (Muntoni & Retelsdorf, 2018), and peers (Muntoni et al., 2020) seem to play a large role in the emergence of these differences. There are many different assumed mechanisms of how these gender stereotypes are learned and acquired, such as model learning, reinforcement of gender-typical behaviour, different treatment of boys and girls, or direct expression of gender stereotypical expectations (Gunderson et al., 2012; Heyder et al., 2019). These acquired stereotypes can then have an effect on how individuals process and categorize information, as well as on their choices, behaviours, and beliefs (Martin & Halverson, 1981). As certain school subjects such as math, science, reading, and language arts are often stereotyped towards one gender or another (e.g. math is typically a "boy" subject, reading is typically a "girl" subject), these beliefs are also incorporated into traditional gender stereotypes (Leaper, 2015; Plante et al., 2013; Schmenk, 2004). Eccles' expectancy-value theory provides a promising explanation for the mechanisms by which these stereotypes can affect individual choices and behaviours, stating that whether or not an individual undertakes a task depends on their expectations for success and how valuable they perceive the task to be (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000). Gender stereotypes can shape not only how valuable a task is to an individual, but also their self-concept, attitude, and perceived competence in that task, which in turn affects their expectancy for success (Eccles, 1994). According to this framework, if a girl perceives math as a male-associated subject, she will not only place low value in it but will also perceive herself as less competent in math, therefore expecting less success, and putting less effort into math or not choosing to study math later in life. Through these mechanisms, gender stereotypes then lead to differences in what young girls and boys are interested in and enjoy, their beliefs about their own capabilities, and the choices they make throughout their academic careers (Eccles et al., 1993; Wigfield & Eccles, 1994).

School-based interventions to reduce gender differences?

Ensuring equal opportunities in education means striving for students to learn and develop according to their full potential. Investigating strategies to reduce gender differences in motivational-affective factors is a step forward in the effort to help all students thrive and succeed. With regard to motivational-affective factors of students in general, regardless of gender, an increasing number of researchers have called for studies to develop methods that could positively reinforce or strengthen these factors and thereby positively influence student achievement outcomes (Gutman & Schoon, 2014; Heckman et al., 2006; Lleras, 2008). There is evidence that motivational-affective factors remain malleable throughout an individual's lifespan, and can therefore be built upon and changed through experience and individual development (Heckman & Kautz, 2013). One promising way of achieving this is through school-based interventions. For the purpose of this study, we define "school-based interventions" as any method used in a school-context, which is different from regular instruction, including not only in-class interventions but also novel teaching methods, summer school programmes, or school-organized workshops. Many studies have empirically tested a variety of these interventions to evaluate the effects on student motivational-affective factors. Previous research syntheses have aggregated the effects of some of these interventions. For example, Durlak et al. (2011) performed a meta-analysis on 213 school-based interventions targeting student factors such as attitudes and emotional skills, and found that overall, students who participated in the interventions significantly improved on measures of these outcomes compared to controls. Additionally, in a review by Gutman and Schoon (2014), results showed that factors such as motivation and self-perceptions of children and adolescents were positively affected in all intervention settings. However, these research syntheses did not evaluate the differential effects of the interventions regarding gender and school subject. It is still unclear whether certain intervention methods have stronger effects on males or females, or whether any of these methods are effective in reducing the differences between male and female student motivational-affective factors in a given academic domain. Additionally, these studies focused on a broader range of student factors (such as social skills, conduct problem, social behaviour, self-control, and creativity), not specifically motivational-affective factors, and did not consider possible moderator variables such as student grade level or intervention duration in their analyses.

The methods used by these intervention studies, as well as the targeted motivational-affective factors and student groups, vary widely. For the purposes of this study, intervention studies can be separated into two types: those that target motivational-affective factors in students overall (i.e. no specific gender is targeted, no gender-specific hypotheses) and those that target motivational-affective factors in one gender specifically, generally the gender that is typically disadvantaged in a given subject (i.e. gender-specific hypotheses). Within these two categories, many different interventional methods are used. Some of these methods can be classified as "psycho-social interventions", while others are more related to classroom processes. Psycho-social interventions are designed to directly target students' subjective psychological processes in an attempt to positively alter them (Walton, 2014). These interventions use strategies such as value affirmations, reframing techniques, and mindset changes. One example of psycho-social interventions without gender-specific hypotheses are utility-value interventions, where students are asked to relate the information they learn in class to their everyday lives in an effort to increase interest for the subject (Hulleman et al., 2010). Psycho-social interventions have also been used for gender-specific interventions by targeting the disadvantaged gender in certain subjects to directly challenge gender stereotypes that students hold. One such strategy is exposing students to role models or mentors who occupy non-traditional gender roles (e.g. a female engineer or a male nurse). Exposure to an individual in a non-traditional gender role can challenge gender stereotypes, lessening the effects of these stereotypes on student self-beliefs (Morgenroth et al., 2015). Another strategy is to target gender-specific student value beliefs. According to Eccles (1994), students will be more likely to engage and put forth effort in a subject if they perceive it to be valuable. While students may find little value in subjects that stereotypically do not align with their gender identity, strategies to make the subject material personally relevant to these students can change how important they view it to be, thereby increasing the likelihood that they engage, take interest, and continue studying that subject (Hulleman et al., 2010).

Outside of psycho-social interventions, other interventional methods focus more on the instructional processes in the classroom, using diverse teaching methods to promote higher levels of motivational factors or more positive affect for students in general. Active learning strategies such as cooperative learning or problem-based learning have been shown to increase student engagement, motivation, and self-efficacy in various settings (Bruder & Prescott, 2013; Laal & Ghodsi, 2012). These strategies attempt to engage students socially and cognitively, encouraging them to be active rather than passive learners, thereby increasing their interest, motivation, and enjoyment (Hmelo-Silver, 2004; Slavin, 2011). Additionally, instructional methods that integrate digital media, such as interactive online lessons or digital games, are being evaluated more frequently as possible strategies for positively affecting student motivational-affective factors (Erhel & Jamet, 2013; Lieberman et al., 2009; Wang & Reeves, 2007). Using digital media in instruction can also affect student motivational-affective factors by providing a novel environment for learning and possibilities for adapting to individual learner needs and interests (Annetta, 2008; Christensen, 2002; Uzunboylu & Karagozlu, 2015). While these methods mostly target student motivational-affective factors in general, pre-existing differences between male and female students in certain subjects could lead to differential effects of these interventions as well.

In sum, an increasing number of studies have evaluated the effects of school-based interventions on student motivational-affective factors. However, it remains unclear whether these interventions have different effects for female and male students in a given subject, and which are most effective regarding gender-specific deficits. Additionally, these studies vary widely in the interventional methods and motivational-affective factors they address, as well as the age and grade level of the student population. The duration and implementation of these interventions also differ between studies. It is unclear which variables can influence the intervention effects concerning gender differences. Therefore, one goal of the present meta-analysis is additionally to evaluate potential moderator variables of the intervention effects.

Potential moderator variables

Theoretical moderators

School subject

Motivational-affective factors are often closely linked to academic domains (e.g. mathematics, science, etc.) and can therefore vary across these domains (Marsh et al., 2001; Wigfield et al., 2004). Student characteristics such as self-concept, motivation, interest, self-efficacy, enjoyment, and anxiety have all been shown to be tightly connected to the subject they are measured in (Bong & Clark, 1999; Goetz et al., 2006; Green et al., 2007; Wigfield et al., 2004). Therefore, a student might have high self-concept in mathematics, but low self-concept in reading. The fact that these factors are domain specific is also an important characteristic with regard to gender differences between students. Due to the gender stereotypes associated with certain academic subjects, gender differences also vary with regard to the domain in question (for example, boys are usually disadvantaged in reading, while girls are usually disadvantaged in science), and must therefore be discussed in a domain-specific context. Therefore, we included the school subject as a moderator.

Gender-targeted vs. non-targeted

The studies included in this meta-analysis evaluate interventions that target student motivational-affective factors *in general* and therefore have no gender-specific hypotheses. We also include interventions that target the motivational-affective factors of a *specific gender* and therefore hypothesize that the intervention will have differential effects for males and females. A "non-targeted" intervention aims to positively affect the motivational-affective factors of all students, while a gender-targeted intervention aims to positively affect the motivational-affective factors of a specific gender, usually the stereotypically disadvantaged gender in a given subject (e.g. girls in science or boys in reading). Therefore, we included this as a moderator variable.

Grade level

Intervention effects may also vary as a function of student grade level. While motivational-affective factors retain a malleable quality throughout an individual's life, there is evidence that they are more flexible at earlier ages (Gutman & Schoon, 2014). Children's self-perceptions have been shown to decline with age from first grade to 12th grade (Eccles et al., 1993; Jacobs et al., 2002). Additionally, there is evidence that children endorse traditional stereotypes more as they grow older (Rowley et al., 2007) and gender stereotypical self-beliefs in school have been shown to take effect around grade three (Herbert & Stipek, 2005). Therefore, we also explored grade level as a moderator variable.

Intervention duration

Some previous reviews on intervention studies have found that the duration of the intervention may affect how successful the intervention is. Hattie et al. (1996), found a small effect for intervention duration, with shorter interventions (1 or 2 days) having a greater initial impact, but with longer interventions (4–30 days) being more effective overall. Additionally, in their review of the effects of reading and mathematics programmes on student performance, Slavin and Lake (2009) found that interventions with briefer durations reported somewhat larger effect sizes than those with longer durations. Other meta-analyses on intervention studies have also found that the intervention duration could be a moderator of intervention effects (de Boer et al., 2014; Dignath & Büttner, 2008). Given these findings, we included intervention duration as an additional moderator variable.

Robustness moderators

Study quality

In addition to these theoretical moderators, we included study quality as a methodological moderator to control for the effect of study quality on effect sizes. We used the What Works Clearinghouse Standards for Intervention Studies (What Works Clearinghouse, 2020) as a guide when selecting criteria for study quality. To be accepted for inclusion, studies had to compare an experimental group and a control group, and use either random or quasi-experimental assignment. Studies also had to provide pre- and post-test data of participants. Additionally, we coded certain aspects of the instruments used by the studies to measure the outcome variable, namely whether the instrument was established or self-developed by the researchers, and whether the reliability was high, low, or not reported, according to standard rules of thumb for instrument reliability (Tavakol & Dennick, 2011). In addition to these study quality moderators, we also used pre-test data to assess the baseline equivalency for the control and intervention groups.

Type of motivational-affective factor

We also included the type of motivational-affective factor measured as a moderator in order to be able to investigate differences and commonalities between types of motivational-affective factors. For motivational-affective factors, we considered attitudes, beliefs, expectations, motivation, career aspirations, interest, self-concept, self-efficacy, self-confidence, enjoyment, boredom, engagement, anxiety, and satisfaction as possible outcomes.

The present meta-analysis

The main goal of the present research is to conduct a meta-analysis on studies that tested the effects of interventions on student motivational-affective factors and reported gender-specific results. This allowed us to evaluate the following research questions:

1. Do school-based interventions that promote motivational-affective factors in students have differential effects for the stereotypically disadvantaged gender (e.g. males in reading/language arts and females in STEM) and stereotypically non-disadvantaged gender in a given school subject?
2. Are the effects of school-based interventions moderated by:
   a. whether the intervention is gender-targeted or non-targeted?
   b. the grade level of the students?
   c. the intervention duration?

Given the existing evidence on gender differences in student motivational-affective factors, we expect that school-based interventions that target these factors in students will have differential effects for males and females, in particular, more positive effects on the gender typically affected by effects of negative stereotypes in a given school subject (Research Question 1). We also expected that the intervention target (gender targeted vs, non-targeted), grade level and intervention duration would moderate the differential effects of the intervention on students' motivational-affective outcomes (Research Question 2). We also included type of motivational-affective outcome and intervention method as moderators in order to be able to identify specific effects of the various interventions.

In order to assess the above questions, we are interested in both the absolute intervention effects for male and female students (i.e. the intervention effects for girls or boys respectively) as well as the intervention effects on the difference between girls and boys (i.e. if pre-existing gender differences are significantly reduced by the intervention).

METHOD

Literature search and study selection

We first conducted a literature search in the databases of PsycINFO and Educational Resources Information Center (ERIC). A flowchart of the study selection process can be seen in Figure 1. The keywords used pertained to (1) the population of interest (students, school, etc.), (2) the topic of gender differences, (3) school-based interventions, and (4) motivational-affective student factors of interest. Keywords varied slightly between the two databases based on the subject heading classification system of each database. The full syntax is included in Appendix A. We restricted the search to studies written in English. In order to include grey literature, we did not restrict the search by publication type, therefore including grey literature such as dissertations, conference proceedings, and other literature formats. We did not restrict the search results by year published in order to search all studies on this topic as thoroughly as possible. This first search was conducted on 13 May 2019. The databases searches resulted in 5,650 references (after removing duplicates), with 4,480 results from ERIC and 1,170 results from PsycINFO. The searches from ERIC and PsycINFO were updated on 26 August 2020, resulting in an additional 184 new results from ERIC and 38 new results from PsycINFO. An additional search in the Web of Science database was also conducted on this date, which returned 2,607 results. Additionally, the reference lists of all included studies, as well as existing meta-analyses or reviews on similar topics, was manually screened for possibly relevant articles, which resulted in an additional 28 studies. In order to locate unpublished studies, the authors of all included studies were contacted via email to inquire about any additional unpublished works that might be relevant. This resulted in one additional study. We also sent a call for papers and/or data to the mailing lists and newsletters of the European Association for Research on Learning and Instruction (EARLI), the American Educational Research Association (AERA), the European Educational Research Association (EERA), and the Gender and STEM research network. This call for papers detailed the topic of interest of the meta-analysis, as well as the study inclusion criteria, and invited researchers to send any relevant published or unpublished work. However, we received no responses from this call.

The titles and abstracts of all studies were screened using the following inclusion criteria:

1. **Study Design**. Studies were only included if they compared an experimental group to control groups, with either random or non-random (i.e. clustered) assignment to groups. Studies also were required to include pre-test data.
2. **School Level**. Only studies conducted with students at the primary or secondary school level were included. Studies in higher education or pre-school levels were excluded.

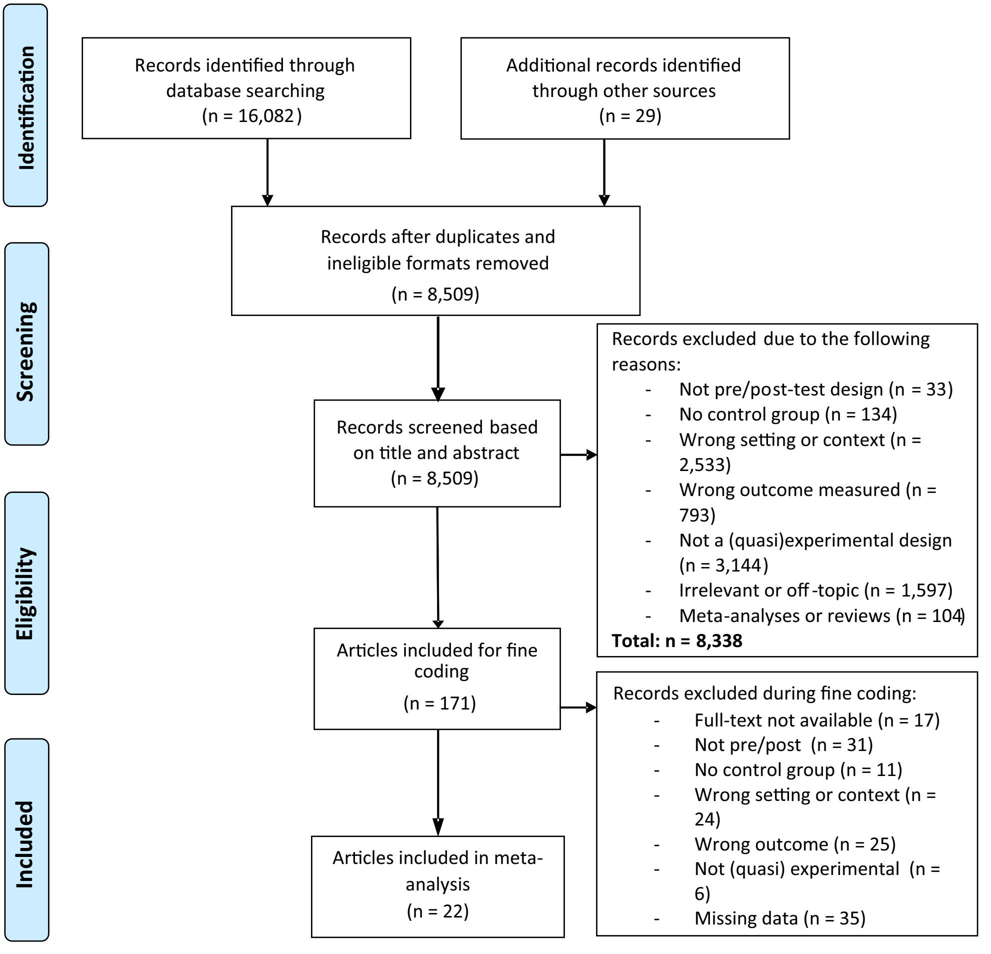


FIGURE 1 Flow chart of the study selection process. Study selection was done following the guidelines of The PRISMA Group (Moher et al., 2009)

3. **School Subject**. Only studies conducted in the core subjects of mathematics, science, reading/(native)language arts or STEM were included. Studies conducted in alternative subjects, such as physical education, arts, or foreign language, were not considered.
4. **Sample Composition**. Only studies that had a sample that represented the average student population of the class were included. Studies with samples consisting exclusively of a specific ethnic or religious group were excluded, if the sample was purposefully selected out of the general school population, for example, if only the African American males in the school were included in the study, or only Latin-American girls were selected to be in the sample (this was not the case if the study takes place in a country where the sample is naturally made up of a specific ethnic group, for example, Saudi Arabia or Mexico). Along the same lines, studies with samples consisting exclusively of gifted or special education students were also excluded. These subgroups are often described as having specific characteristics (based on their background and prior experiences) that are in the focus of corresponding interventions. Accordingly, the results of these intervention studies cannot be transferred to a general student sample.
5. **Intervention Study**. Only studies that evaluated a school-based intervention were included. Interventions conducted solely in the home environment (e.g. with parents or siblings) or outside of a school context (e.g. in church, in sports clubs, etc.) were not included.

6. **Motivational-Affective Factors as Outcome**. Studies were included only if they evaluated the effects of the intervention on one or more motivational-affective student factors.

Using the information from the titles and abstracts of each reference, articles were categorized as "Included" or "Excluded", with a reason given for each exclusion with regard to the criteria. A portion (275) of the articles were double-screened by two separate, trained research assistants. The double coding showed that the inclusion and exclusion process was reliable, with 91% agreement between the coders' overlapping articles. Disagreements were discussed to arrive at a final decision.

Coding of data

In order to code all relevant study variables, full texts were acquired for the studies that were included. The studies were coded using a coding scheme developed by the researchers. The coding scheme was created according to the general publication features (author, date, etc.), as well as methodological variables such as the intervention duration, implementation (instructor of intervention, setting of intervention, etc.), measurement instruments, content-related variables such as the theoretical background, sample characteristics, and school subject, and quantitative data necessary for effect size calculations such as the means, standard deviations, and sample sizes. In order to assess the gender-specific intervention effects, we coded the pre-test and post-test scores for the experimental and control groups for the entire sample, for the female participants only, and for the male participants only. An overview of all variables coded and their purpose can be seen in Table 1. In order to ensure coding reliability, a portion of the included studies were fully double coded by two trained researchers to assess inter-rater agreement. As an indicator for inter-rater reliability, Cohen's Kappa (Cohen, 1968) was calculated for each variable. Values for Cohen's Kappa ranged from $\kappa = .51$ (intervention method), which is considered a substantial agreement, to $\kappa = 1.00$ (grade level), which represents perfect agreement (Landis & Koch, 1977). The agreement for variables without any margin for interpretation, such as means, standard deviations, and sample sizes for effect size calculation, was almost perfect. Any disagreements were discussed between the two coders. During the fine coding process, it was necessary to exclude some studies which originally seemed eligible, as upon further inspection, they were either not a fit for the meta-analysis or did not report the necessary information (the most common case being that the authors had not evaluated gender differences). In the case of missing information, authors of the respective articles were contacted, when possible, to request missing data. In total, we fully coded 22 eligible studies.

Moderator variables

We originally planned to include all moderator variable in one meta-regression model. However, due to the smaller number of studies we included than anticipated, we instead decided to conduct a separate

TABLE 1 Overview of coded variables

Study identification	Descriptive variables	Study quality	Effect size and analyses	Moderator variables
Author(s) Title Year of publication Document type	Intervention method Country	Study design Source of instrument Instrument reliability	Sample sizes Single sex or both sexes Means (pre/post) Standard deviations (pre/post)	Grade level Intervention duration Motivational-affective factor Target of intervention

meta-regression for each moderator in order to assess the effects. For the categorical moderators, we added dummy-coded predictors for each of the different levels.

School subject

We originally coded the school subject for each study as stated in the study (e.g. Biology was coded as "Biology" and not "Science"). However, after completing the coding process, we categorized the different subjects as Science, Mathematics, Reading, and Informatics/Technology (therefore, "Biology" would now be coded as "Science"). Upon completion of the fine coding, we were only left with one study conducted in a "male-disadvantaged" subject (Kerneža & Košir, 2016, conducted in reading literacy). All other studies were conducted in the subjects of science or math, with two exceptions conducted in technology/computer science. We therefore excluded the study by Kerneža and Košir (2016), as one study in a male-disadvantaged subject would not allow us to accurately assess the differential effects. We therefore ended up with only studies that were conducted in the typically female-disadvantaged STEM subjects of science, math, and technology.

Gender targeted vs. non-targeted

An intervention was coded as gender targeted when the authors stated that they expected the intervention to have different effects for male and female students, either in their research questions or hypothesis. An intervention was coded as not gender-targeted when no gender-specific effects were considered in the research question or hypotheses.

Grade level

We originally coded grade level according to the specific grade reported in the study (e.g. Grade 3, Grade 10). If this was not reported, we coded the school level reported in the study (e.g. primary school, high school). Once the coding was complete, we categorized each grade level into Primary (Grades 1–5), Lower Secondary (Grades 6–8) or Upper Secondary (Grades 9–12).

Intervention duration

We used the number of weeks to represent the intervention duration. The shortest duration was 1 week. We centred this variable by subtracting 1 from each value, in order to have a meaningful intercept in the meta-regression model (so the intercept would represent a one-week intervention).

Intervention method

In order to assess which intervention methods were most affective, we classified the interventions into two categories according to our previous description of intervention types. The first category was psycho-social interventions. These interventions directly targeted students' motivational-affective processes by attempting to change or restructure their subjective beliefs and perceptions. This category included interventional methods such as utility-value interventions, role models, and challenging stereotypes. The second category was instructional interventions. These interventions included strategies that focused on the instructional and pedagogical processes in the classroom, with the aim of using diverse teaching methods to change students' motivational-affective factors. This category included interventional methods such a problem-based learning, cooperative learning, or novel curriculum designs. We categorized interventions according to the theoretical frameworks and designs used in the primary studies.

Study quality

We originally planned to use study design (experimental vs. quasi-experimental), statistical reliability of the instrument used to measure outcomes (high vs. low vs. not reported), and the source of the instrument (whether it was an already existing instrument or developed by the researchers of the corresponding study themselves) as variables to evaluate study quality. However, upon completion of the coding, it became evident that all included studies were quasi-experimental, and we therefore could no longer use this variable as a moderator. We coded statistical reliability of the instrument as high, low, or not reported, and the instrument source was coded as an already existing instrument or one that the researchers created themselves. After completing the coding process, only one study was included which did not report the instrument reliability, while all other studies reported high reliability. We also found, upon completing the coding process, that all but one study included used an already existing instrument. We therefore did not include these variables as moderators. We calculated baseline equivalency for the included samples as the standardized mean difference between the control and experimental groups at pre-test, as per the What Works Clearinghouse standards (What Works Clearinghouse, 2020).

Type of motivational-affective factor

While we originally planned to assess the effects of the interventions on each separate motivational-affective factor individually, the final data did not allow enough power for this type of analysis. We therefore classified the motivational-affective factors into four different categories, loosely based on the framework of Murphy and Alexander (2000). The four categories were (1) motivational factors (motivation, interest, value, and engagement), (2) self-schemas (self-efficacy, self-concept, self-confidence, and stereotypes), (3) attitudes, and (4) affective factors (enjoyment, anxiety, and satisfaction). We reversed coded the data for outcomes where a decrease in the mean score represents a positive outcome (e.g. anxiety or stereotypes), so that all effect sizes represented the same directional relationship (i.e. a positive effect size indicates desirable change in motivational-affect factor).

Calculation of effects and general analytic strategies

All analyses were conducted in R Version 4.0.5. (R Core Team, 2019), using the *metafor* package (Viechtbauer, 2010) as well as the *robumeta* package (Fisher & Tipton, 2015). All included studies were quasi-experimental designs with pre–post data from treatment and control groups. Therefore, we calculated the effect sizes for males and females as the difference (g) between the standardized mean change score for the treatment and control groups, including a small-sample bias adjustment. We calculated this using the means, standard deviations, and sample sizes of the treatment and control groups for both males and females in each study. According to Morris (2008), the effect size for pre–post-control design studies is defined as the mean difference between post-test and pre-test scores, divided by the common standard deviation. Following the recommendations made by Morris (2008), we used the pooled pre-test standard deviation to calculate standardized mean change score for pre–post-control studies, as this provides an unbiased estimate of the population effect size. This effect size was calculated as follows:

$$g = c_p\left[\frac{\left(M_{post,T} - M_{pre,T}\right) - \left(M_{post,C} - M_{pre,C}\right)}{SD_{pre}}\right]$$

where the pooled pre-test standard deviation is defined as

$$SD_{pre} = \sqrt{\frac{(n_T - 1)SD^2_{pre,T} + (n_C - 1)SD^2_{pre,C}}{n_T + n_C - 2}}$$

and the bias correction is defined as

$$c_p = 1 - \frac{3}{4(n_T + n_C - 2) - 1}$$

We used random-effects statistical models for this meta-analysis, as this allows the findings to be generalized beyond the included studies (Hedges & Vevea, 1998). We coded at least two effect sizes per study (male vs. female), and if the study measured multiple motivational-affective factors, these were also coded as separate effect sizes. Separate effect sizes that come from the same study are not independent of each other, and therefore violate the independency assumption of classical random-effects meta-analyses (Hedges et al., 2010). We remedied this by using robust variance estimation (RVE) to estimate our model. RVE allows for the inclusion of statistically dependent effect sizes by adjusting the standard errors to account for dependency (Tanner-Smith et al., 2016). We used correlated effect model weights to model the unknown covariance structure (Fisher & Tipton, 2015).

Heterogeneity across the included studies was assessed by calculating the 95% prediction interval, which describes the expected range of true effects by predicting where the true effects are to be expected for 95% of similar future studies (Borenstein et al., 2017). Therefore, this interval can be used to evaluate the variability of intervention effects over different settings (IntHout et al., 2016). To evaluate the first research question, whether the interventions demonstrated a differential effect for male and female students on motivational-affective outcomes in a given school subject, we estimated a simple random effects meta-regression model using RVE, with gender included as a predictor. Due to the lack of studies conducted in a male-disadvantaged subject, we could not evaluate the differential effects of a male-disadvantaged versus female-disadvantaged subject. However, we still included the main subjects of science and mathematics as moderators.

Fourteen of the included studies used a cluster-randomized design, where treatment and control conditions were assigned at the classroom level (i.e. using one pre-existing classroom as the treatment group and another as a control group), but outcomes were reported on the student level. This clustering effect can lead to additional variance, meaning that it was necessary to adjust the variances estimates for studies using this design (Hedges, 2007). As none of the studies reported the intraclass correlation (ICC) necessary for the variance adjustment, we used a conservative estimate of ICC = .20, as recommended by Hedges and Hedberg (2007).

With regard to the second research question, we used separate meta-regression models to investigate the effects of each individual moderator. We originally planned to examine publication bias via selection models. Selection models aim to directly model the selective publication process by considering the probability that certain studies are included in a meta-analysis based on specific characteristics and using weight functions to adjust the overall effect size estimate (Vevea & Hedges, 1995). However, according to McShane et al. (2016), realistic selection models cannot be properly estimated without a large amount of data, and without sufficient data, selection models cannot be relied on to provide accurate estimates. Due to the relatively small final sample size of our meta-analysis, we therefore did not evaluate a selection model. We evaluated publication bias visually by inspecting the funnel plot, as well as statistically using Egger's regression test (Egger et al., 1997) and Kendall's rank correlation test (Begg & Mazumdar, 1994).

In a meta-analysis, outliers and other exceptional cases can affect the interpretability and robustness of results (Viechtbauer & Cheung, 2010). Therefore, we conducted a variety of diagnostic tests on the overall model to examine this. We calculated the: (1) externally standardized residuals, (2) DFFITS values, (3) Cook's distances, (4) covariance ratios, (5) leave-one-out estimates of the amount of heterogeneity, (6) leave-one-out values of the test statistics for heterogeneity, (7) hat values, and (8) weights.

We used the *metafor* package to calculate effect sizes, and the *robumeta* package to estimate our overall model as well as meta-regression models for our moderators, using robust variance estimation to correct for dependent effect sizes. Additional analyses for publication bias, heterogeneity, and outlier detection were all conducted in *metafor*, using a model estimated without correction for dependent effect sizes, as these additional analyses are not possible in *robumeta*. A registered protocol of this study, along with a template of the coding scheme and the R scripts can be found on the Open Science Framework website via the following link: https://osf.io/zb8sc/?view_only=7392757de06e45ffbd3c04c87d569b9c.

RESULTS

Descriptive characteristics

The inclusion criteria were met by 21 independent studies, with 79 relevant effect sizes obtained from a total of 3,458 participants overall. All studies were published or conducted between 2000 and 2019, with the exception of 1 article published in 1981. The sample sizes ranged from 11 to 732 participants. There was a large variation in the intervention techniques. Some focused directly on student motivational-affective factors and attempted to alter these by fostering skills such as goal setting, connecting subject matter to relevant aspects of student lives, or challenging stereotypes through examples and role models. Other interventions incorporated different instructional techniques such as problem-based learning, cooperative learning, or digital games in efforts to increase student motivation, attitudes, and engagement. A descriptive overview of the relevant characteristics from each study can be seen in Table 2.

Outlier analysis

The outlier analysis identified the effect sizes from the Akcay et al. (2010) study as outliers. The effect sizes ranged from $g = 2.30$ to $g = 5.14$, which were much larger than any of the other effect sizes from the other studies. As the presence of outliers can affect the validity and robustness of meta-analytic results, we removed this study from all further analyses.

Overall effect and differences in gender and subject (Research question 1)

After removal of outliers, the final number of studies included in the analyses was 20, with a total of 71 effect sizes. The first research question focused on whether interventions that targeted student motivational-affective factors had differential effects for the stereotypically disadvantaged and non-disadvantaged gender in a certain subject. We first ran a model with just gender as a predictor to examine the overall role of gender on the intervention effects (we ran model with female as the reference category and one model with males as the reference category). The results of this model can be seen in Tables 3 and 4. We found a significant positive effect of the interventions on both male and female student motivational-affective factors; however, there was no statistically significant difference between males and females. The overall descriptive effect for female students, at $g = .49$, was almost double that for male students, which was $g = .28$. The measures of heterogeneity indicated substantial heterogeneity between effects sizes, with $I^2 = 80.33$ ($\tau^2 = .14$), indicating that a large percentage of variance is due to heterogeneity between studies. The 95% prediction interval was .71 to 3.24 for females, and .59 to 2.68 for males, also indicating high heterogeneity. A forest plot for the female and male effect sizes from the various studies can be seen in Figures 2 and 3 respectively.

Due to the lack of studies conducted in male-disadvantaged subjects, we were not able to evaluate if there were differential intervention effects for the stereotypically disadvantaged gender in a given school subject. However, we still included school subject as a moderator variable (with the two categories of either science or mathematics) and gender, as well as the interaction between subject and gender, to evaluate whether there were any significant differences of the intervention effects for the interplay between these two variables. Results of these analyses are displayed in Table 5. A random-effects meta-regression model using robust variance estimation did not reveal any significant difference of the interventions effects moderated by the school subject or its interaction with gender. Descriptively, the estimated effect sizes were highest for females in science $g = .46$, whereas for males in science, the effect size was $g = .30$. For females in math, the estimated overall effect size was $g = .40$, and for males in math, $g = .26$. We also conducted post-hoc subgroup analyses to evaluate the absolute effects on the

TABLE 2 Descriptive overview of included studies

Study	Sample Size (*N*)	Grade Level	Country	Subject	Intervention Method	Gender Targeted	Outcome Measured	Summary of Results
Abed (2016)	87	Lower secondary	Jordan	Science	Drama in science instruction using physical simulations activities (INST)	Yes (only male students served as sample)	Attitudes towards science	The use of drama-based science teaching brought about positive changes in student's attitudes towards science learning and their understanding of scientific concepts compared to traditional science teaching
Al-Balushi and Al-Aamri (2014)	62	Upper secondary	Oman	Science	Project-based learning (INST)	Yes (only female students served as sample)	Attitudes towards science	Students who participated in the project-based learning group developed more positive attitudes towards science than the control group
Alghamdi (2017)	50	Upper secondary	Saudi Arabia	Science	Cooperative learning model by using the Jigsaw instructional strategy (INST)	Yes (only male students served as sample)	Attitudes towards science	Although the use of Jigsaw as a cooperative learning strategy improved the science achievement of students in the experimental group, the strategy did not change the experimental group student's attitudes in a positive direction
Argaw et al. (2016)	81	Upper secondary	Ethiopia	Physics	Problem-based learning in physics (INST)	No	Motivation (to learn physics)	Students taught with a PBL approach showed better outcomes post-test versus control; however, motivation for both groups remains stagnant, no significant differences of intervention in problem solving and in motivation on male and female students

TABLE 2 (Continued)

Study	Sample Size (*N*)	Grade Level	Country	Subject	Intervention Method	Gender Targeted	Outcome Measured	Summary of Results
Batton (2010)	64	Primary	USA	Math	Cooperative learning based on Vygotsky's social learning and Piaget's concept of knowledge (INST)	No	Mathematics anxiety	Decrease in math anxiety is larger for students in cooperative group in comparison to students in non-cooperative group, decrease in math anxiety for female students in cooperative group only, no difference for male students in decrease of math anxiety in cooperative and non-cooperative group
Cady and Terrell (2008)	26	Primary	USA	Computer science	Positive relationship between frequent technology use and self-efficacy and technology acceptance, daily infusion of technology use into science lessons (PSI)	Yes (only female students served as sample)	Attitudes towards computers (importance, enjoyment, self-efficacy)	Students in intervention group develop higher importance and higher self-efficacy of technology use versus control, no significant differences in computer enjoyment between groups
Chen et al. (2019)	68	Primary	Taiwan	Science	Modified Argument-driven Inquiry (MADI) teaching approach (INST)	Yes	Engagement in learning science	The engagement in learning science of the experimental group was statistically significantly higher than those in the control group. However, no statistically significant changes between boys and girls of the experimental group were found

(Continues)

TABLE 2 (Continued)

Study	Sample Size (*N*)	Grade Level	Country	Subject	Intervention Method	Gender Targeted	Outcome Measured	Summary of Results
Chiu (2011)	247	Lower secondary	Taiwan	Science	Social cognitive approach with lessons focused on women in sciences and men in humanities (PSI)	No	Attitudes towards learning science (interest, confidence, value)	Initial gender gap favouring boys regarding interest and confidence decreased in experimental group only, initial gender gap favouring boys regarding value decreased and changed from favouring boys to favouring girls in experimental group
Fabian et al. (2018)	52	Primary	Scotland	Math	Mobile learning (based on constructivist learning, collaborative learning, situated learning and others) (INST)	No	Attitudes towards mathematics (enjoyment, self-confidence, value of math)	Students in the experimental group had significantly higher gains in performance in comparison to the control group. However, no significant difference was found in student's attitudes towards mathematics between the groups. No significant gender differences were found either
Falco et al. (2008)	228	Lower secondary	USA	Math	Improving students' self-efficacy beliefs towards learning mathematics through fostering the skills time management, goal setting, study habits, and help-seeking. (PSI)	Yes	Attitudes towards mathematics (self-confidence, value, enjoyment, motivation)	Students in experimental group develop more positive attitudes towards mathematics compared to control group, gains for female students compared to male students were significantly higher

TABLE 2 (Continued)

Study	Sample Size (*N*)	Grade Level	Country	Subject	Intervention Method	Gender Targeted	Outcome Measured	Summary of Results
Falco and Summers (2019)	88	Upper secondary	USA	STEM	Career group intervention based on self-efficacy theory (PSI)	Yes (only female students served as sample)	STEM self-efficacy and career decision-making self-efficacy	Significant differences were found between the experimental and the control group at post-test and follow-up test for both variables, which means the intervention had a positive impact on participant's career decision self-efficacy and STEM self-efficacy
Fennema et al. (1981)	Not reported	Upper secondary	USA	Math	Videotapes aimed to increase knowledge about sex-related differences in mathematics and improvement of attitudes (PSI)	Yes	Mathematics value, anxiety and stereotypes	Students in the experimental group reported that they would increase the amount of mathematics courses they planned to take. The enrolment data also confirmed that the mathematics enrolment of students in the experimental group increased compared to the control group, indicating that the intervention was effective in increasing women's participation in mathematics
Mavridis et al. (2017)	79	Lower secondary	Greece	Math	Online flexible educational game targeting student attitudes as a supplementary teaching method during mathematics lessons (PSI)	No	Attitudes towards mathematics	Students taught with an educational game approach develop more positive attitudes towards mathematics and higher achievement scores versus control, educational game approach works equally well for male and female student's attitudes
Orabuchi (2013)	155	Lower secondary	USA	Math	Online visual and interactive technological tool based on self-efficacy theory (PSI)	No	Mathematics anxiety, attitudes towards mathematics	No significant differences between groups regarding mathematics performance, mathematics anxiety and attitudes, no gender differences

(Continues)

TABLE 2 (Continued)

Study	Sample Size (*N*)	Grade Level	Country	Subject	Intervention Method	Gender Targeted	Outcome Measured	Summary of Results
Shin et al. (2019)	416	Primary and lower secondary	Korea	Science	Expectancy-value-theory, utility value intervention to positively influence STEM motivation (PSI)	No	Attitudes towards science	Students taught with a utility value intervention develop more positive attitudes towards mathematics versus control, science utility value intervention works equally well for male and female students, no significant differences of intervention on male and female students
Soyibo and Hudson (2000)	77	Upper secondary	Jamaica	Biology	Computer-assisted instruction (CAI) in addition to lecture and discussion (INST)	Yes (only female students served as sample)	Attitudes towards biology and the computer/CAI	The experimental group had significantly better post-test attitudes towards biology and computers/CAI than the control group. The experimental group also significantly outscored the control group in their understanding of reproduction in plants and animals
Starkey (2013)	168	Lower secondary	Kenya	Math	Digital game as a supplementary teaching method during mathematics lessons (PSI)	Yes	Attitudes towards mathematics and situational motivation	Significant increase in mathematics achievement, motivation and attitudes towards mathematics for students who played the digital game versus the control group, digital game approach works significantly better for male students regarding their motivation increase

TABLE 2 (Continued)

Study	Sample Size (*N*)	Grade Level	Country	Subject	Intervention Method	Gender Targeted	Outcome Measured	Summary of Results
Sung et al. (2015)	111	Primary	Taiwan	Math	Technology-based 3D virtual manipulatives and step-by-step progressive teaching for teaching the surface area of composite solids (INST)	No	Attitudes towards mathematics	There was no significant difference in student's attitudes towards mathematics between experimental and control groups, nor between male and female students. The experimental group exhibited better performance compared to the control group
Van der Meij et al. (2015)	61	Lower secondary	Netherlands	Science	Animated pedagogical agents in virtual environments (INST)	Yes	Task relevance (value) and self-efficacy	Students taught with APA develop higher self-efficacy beliefs and knowledge versus control group, significant increase of self-efficacy beliefs for female students in experimental conditions and decrease of self-efficacy beliefs for female students in control conditions, opposite pattern for male students
Zhao et al. (2018)	77	Lower secondary	China	Math	Systematic intervention on collective representations, situational cues and personal characteristics based on the Identity Threat Model (PSI)	Yes (only female students served as sample)	Math-gender stereotypes	The level of math-gender stereotypes in the experimental group was significantly lower than that of the control group after the intervention and at a follow-up test, which indicates that the intervention was effective in reducing math-gender stereotypes among adolescent girls

Note: In the "Intervention Method" column, we indicated whether we classified the intervention as instructional (INST) or psycho-social (PSI).

TABLE 3 Model with gender as predictor

			95% CI			
Predictor	**Estimate**	***SE***	**Lower**	**Upper**	***df***	***p*-value**
Intercept	.487	.103	.266	.708	14.4	.0003***
Gender (Male)	−.208	.150	−.529	.108	17.6	.1841

Note: Female is the reference category (intercept). ***$p < .01$.

TABLE 4 Model with gender as predictor

			95% CI			
Predictor	**Estimate**	***SE***	**Lower**	**Upper**	***df***	***p*-value**
Intercept	.279	.126	.005	.554	12.0	.0468**
Gender (Female)	.208	.150	−.108	.524	17.6	.1841

Note: Male is the reference category (intercept). **$p < .05$.

intervention for with females in science, females in math, males in science, and males in math, in order to see the effectiveness of the interventions for each group separately. These results can be seen in Tables A1–A4 in Appendix B.

Differences in effect sizes depending on moderator variables (Research question 2)

In order to investigate our second research question, we conducted various moderator analyses to evaluate whether any potential moderating variables were responsible for the variance among studies. Results for all moderator analyses can be seen in Table 6. For the categorical moderators (i.e. grade level, intervention target, etc.), we chose the categorical level that displayed the descriptively strongest association with the outcome variable as the reference category.

There were no significant differences between the grade levels with regard to the intervention effects. The results descriptively showed the biggest effect size estimate for primary school levels ($g = .48$), followed closely by lower secondary school levels ($g = .43$), and then upper secondary school levels ($g = .32$). We also evaluated the effect of the intervention method (psycho-social vs. instructional) and its interaction with gender on the intervention effects. Results displayed largest effect size estimates for females when psycho-social interventions were used, $g = .53$, whereas the effect size for males when psycho-social interventions were used was significantly lower ($g = .19$). Effects from interventions using instructional interventions, with $g = .42$ for females, and $g = .41$ for males, did not differ from the reference group (effects of psycho-social interventions for females).

For the moderator of “intervention target”, we also examined the effects of this variable in combination with gender. In reference to the intercepts, which represented gender-targeted interventions and females, the only significant difference found was for non-gender-targeted interventions and females. The effect sizes for gender-targeted interventions for females was $g = .63$ ($p = .002$) while the effect size for females who received a non-gender-targeted intervention was at $g = .20$ ($p = .015$). The effect size for males who received a gender-targeted intervention was $g = .32$, while the effect size for males who received a non-gender-targeted intervention was $g = .24$. Neither of these conditions was significantly different from the reference category (gender targeted and female).

The duration of the intervention in weeks had no significant moderating effect ($g = .01$). Upon completion of this moderator analysis, we ran an additional post-hoc moderator analysis with duration as a categorical variable. We used the categories of less than or more than 4 weeks, based on a

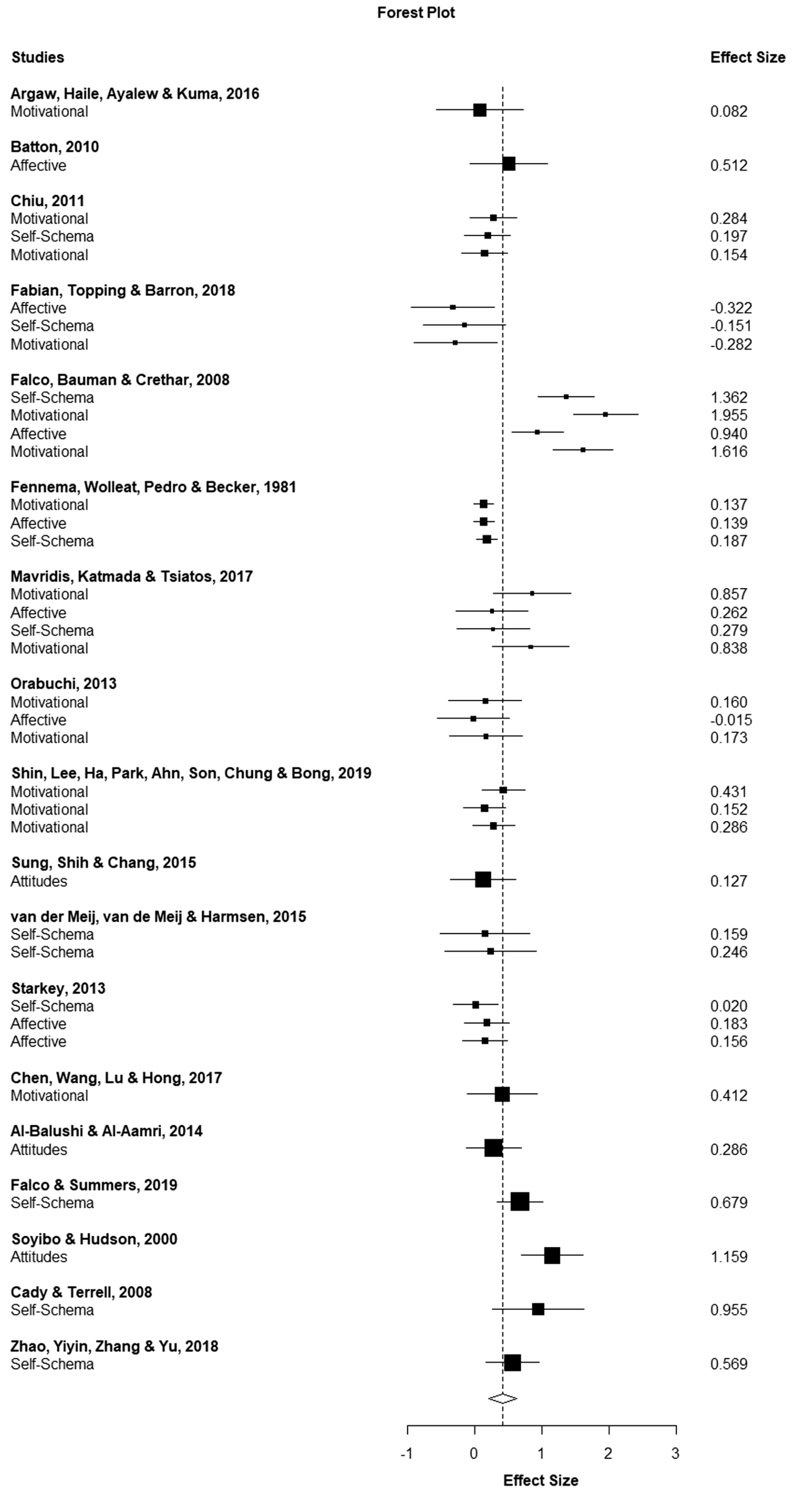


FIGURE 2 Forest plot of female effect sizes. Effect sizes adjusted for dependency using RVE

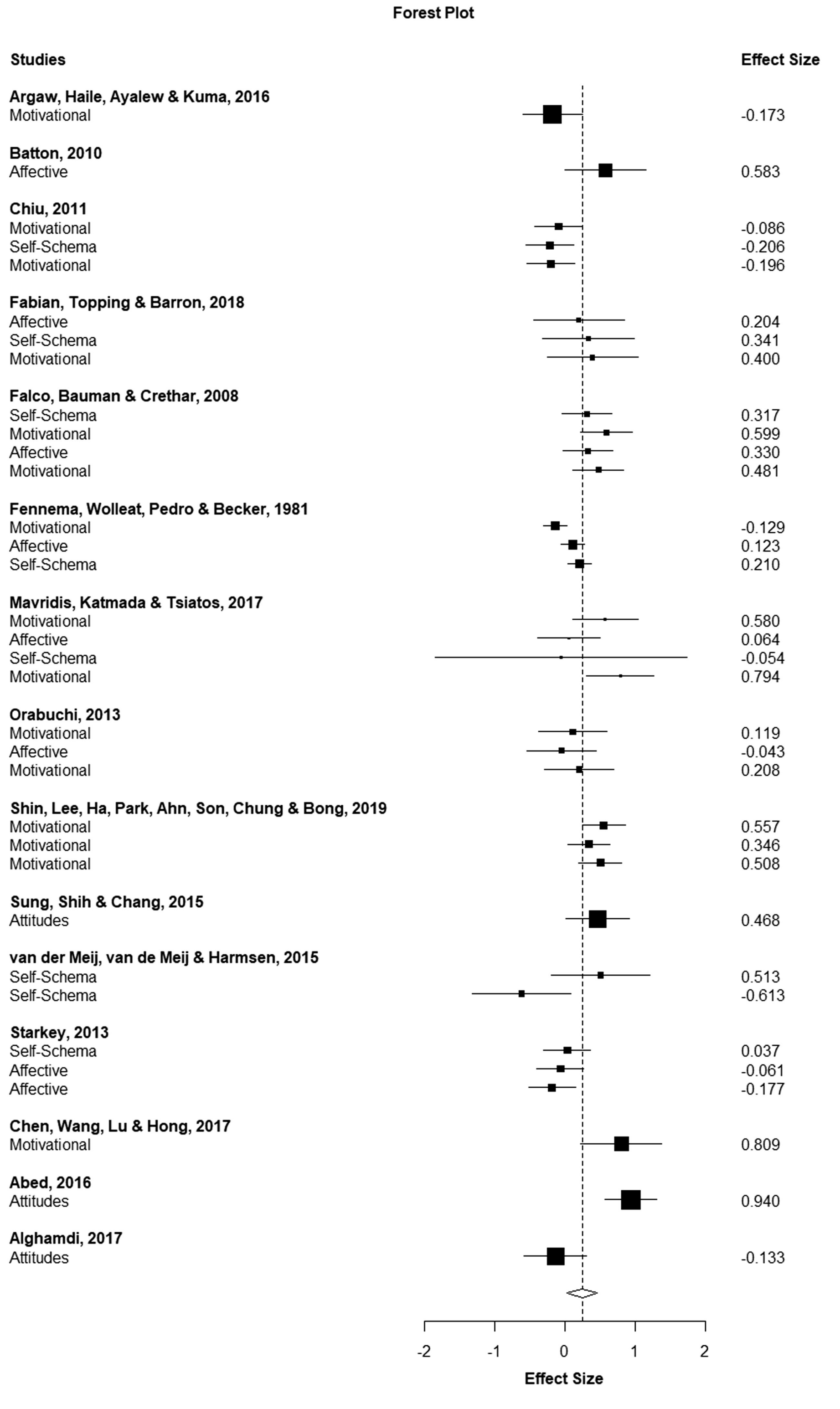


FIGURE 3 Forest plot of male effect sizes. Effect sizes adjusted for dependency using RVE

TABLE 5 Meta-regression with interaction of subject and gender

Moderator Level	*j*	*k*	*df*	Estimate	*SE*	*p*-value	Reference
Subject							
Female & Science	7	12	5.07	.46	.19	.064	REF
Female & Math	9	23	11.51	−.06	.25	.817	
Male & Science	7	12	7.66	−.16	.29	.598	
Male & Math[a]	8	22	15.30	.02	.32	.954	

Note: j represents the number of studies, and *k* represents the number of effect sizes. REF indicates which level of the variable was used as the reference category (intercept).

[a]Indicates that this combination was an interaction term in the meta-regression.

meta-analysis by Hattie et al. (1996), which found that interventions under 30 days had a positive correlation with effect sizes. However, this additional analysis also resulted in non-significant results. There were also no significant differences with regard to type of motivational affective outcome measured. Descriptively, the biggest effect sizes could be seen for interventions where attitude was the outcome measured (g = .52), followed by self-schema outcomes (g = .47), motivational outcomes (g = .34), and affective outcomes (g = .24).

Publication bias & robustness checks

Egger's regression test was not significant (p = .960), indicating that there was no evidence for risk of publication bias. The rank correlation test for funnel plot asymmetry was also insignificant (p = .198), also indicating no evidence for publication bias. The funnel plot can be seen in Figure 4.

With regard to study quality, we found that for 18 of the 71 effect sizes, the baseline equivalency standard recommended by the What Works Clearinghouse (an effect size of more than .25) was not met. A large percentage of these problematic baseline equivalencies were within studies, for only one gender but not the other. Therefore, due to our goal to compare the effects of the interventions by gender, we did not exclude these effect sizes from our analyses. However, we conducted a sensitivity analysis to evaluate the effect of baseline equivalency, which can be seen in Table A5 in Appendix C.

DISCUSSION

In the present paper, we aimed to investigate whether interventions that targeted motivational-affective factors in students had differential effects for the stereotypically disadvantaged and stereotypically non-disadvantaged gender in a given school subject. While other meta-analyses and reviews have evaluated the effects of school-based interventions on these factors in students (Durlak et al., 2011; Gutman & Schoon, 2014), the present study takes a closer look at gender-specific effects of these interventions. We also examined additional variables that might moderate the effects of these interventions.

Gender-specific intervention effects

On a general level, the results of this meta-analysis demonstrate that interventions have the potential to promote motivational-affective factors for both male and female students. This is in line with prior research, which shows that school-based interventions can effectively promote or foster these factors in students. Large-scale meta-analyses have demonstrated the positive effect of interventions on student's attitudes, emotional skills, and motivation (Durlak et al., 2011; Lazowski & Hulleman, 2016; Taylor et al., 2017). The findings that these interventions have, on average, a positive effect for both male and

TABLE 6 Overview and results of moderator analyses

Moderator level	*j*	*k*	%	*df*	Estimate	*SE*	*p*-value	Reference category
Grade level								
Primary	5	13	18%	3.96	.48	.14	.029**	REF
Lower secondary	8	42	59%	7.92	−.05	.20	.812	
Upper secondary	7	16	23%	8.17	−.16	.22	.499	
Intervention method								
Psycho-social & female	10	26	37%	7.83	.53	.12	.002***	REF
Instructional & female	8	11	15%	12.02	−.11	.23	.644	
Psycho-social & male	8	23	33%	7.70	−.34	.11	.010**	
Instructional & male[a]	8	11	15%	16.31	.33	.30	.281	
Intervention target								
Gender targeted & female	10	18	25%	7.87	.63	.13	.002***	REF
Non gender targeted & female	8	19	27%	12.37	−.43	.15	.015**	
Gender targeted & male	7	15	21%	8.90	−.31	.24	.220	
Non gender targeted & male[a]	8	19	27%	13.80	.35	.26	.208	
Intervention duration								
Intercept (1 week)	19	67		5.74	.27	.15	.119	
Duration				2.96	.01	.01	.296	
Outcome type								
Attitudes	5	6	8%	3.99	.52	.23	.092	REF
Affective	7	16	23%	7.97	−.28	.27	.322	
Motivational	9	30	42%	9.59	−.18	.27	.506	
Self-schema	9	19	27%	9.68	−.07	.26	.805	

Note: j represents the number of studies, *k* represents the number of effect sizes, % represents what percent of the effect sizes were at each of the various moderator categories, and REF indicates which level of the variable was used as the reference category (intercept). ***$p < .01$, **$p < .05$.

[a]Indicates that this combination was an interaction term in the meta-regression.

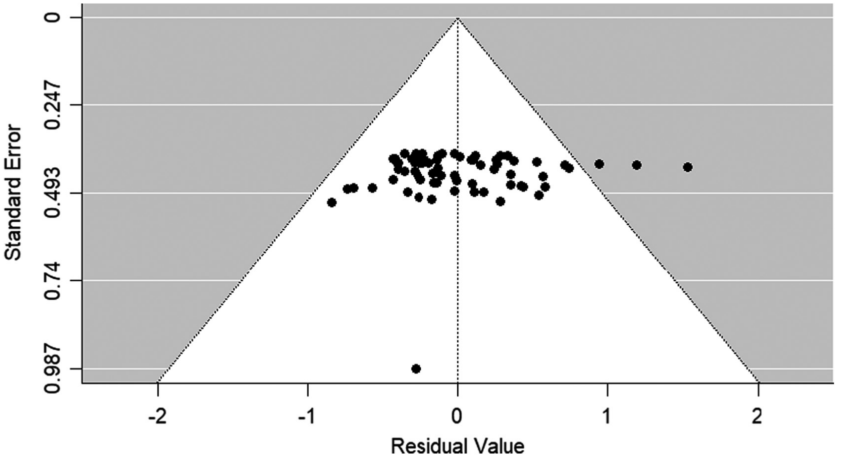


FIGURE 4 Funnel plot

female students is encouraging, as it demonstrates the efficacy of school-based interventions across genders. Descriptively, we found that the overall average effect size for females was larger than that for males; however, this difference in effect sizes was not statistically significant. While these results provide preliminary evidence that there may be a difference in how effective these interventions are based on student gender, more research is needed to determine if these differences in effects are significant, and if they hold true in a larger sample. One possible explanation for this may be that interventions function more effectively when levels of these motivational-affective factors are lower to start out with, which would explain the possibly larger average effect for females in subjects that are typically "female disadvantaged" where females have been shown to have lower levels of motivational-affective factors. Male students tend to have higher levels of these factors on average in subjects such as science or math, and therefore, while these interventions are still significantly and positively affecting them, they may not be as "in need" of support as females in these subjects.

The relevance of intervention characteristics

Descriptively, the results show that the intervention effects were slightly greater on average for students in primary school (Grades 1–5) and lower secondary school (Grades 6–8) than for students in upper secondary school (Grades 9–12). Although these differences were not statistically significant, the findings are in line with other studies, which have also found intervention effects to be stronger for students in childhood and early adolescence. For example, Lazowski and Hulleman (2016) found almost identical effects in a meta-analysis on motivation interventions in education, with the largest effect sizes for students in grades 6–8 ($d = .57$), followed closely by students in grades 1–5 ($d = .52$) and lastly, by grades 9–12 ($d = .42$). The literature suggests that these lower secondary and primary-school-aged groups tend to reap the most benefits from these interventions when it comes to motivational-affective factors and this stage of development might be the prime opportunity to target these factors in students (Juvonen, 2007; Wentzel & Wigfield, 2007).

This study was novel in that it included various types of interventions in order to gain a wide view of the current state of research, whereas most prior meta-analyses on school-based interventions have focused on one specific intervention strategy, such as digital media (Hillmayr et al., 2020) or problem-based learning (Batdi, 2014). We differentiated between psycho-social intervention approaches and instructional approaches. We found that both approaches had a descriptively larger effect sizes for females than for males; however, these gender differences were significant only for psycho-social approaches. The gender difference in effect sizes was particularly prominent for psycho-social interventions. Contrary to instructional interventions, psycho-social interventions are specifically designed to directly target motivational-affective factors in students (Walton & Wilson, 2018) and therefore could explain why the effect size was larger for females in stereotypically male subjects (with lower starting values). Although we did not test the overall differences between psycho-social and instructional interventions (independent from gender), strategies which aim directly for student motivation, attitude, emotions, or self-beliefs may be more effective at promoting these factors than other types of interventions.

Along the same line, our results showed larger descriptive effects for interventions that target females in mathematics or science than interventions that were not gender targeted. However, as our sample only included studies that were conducted in subjects where females were disadvantaged, the interventions likewise were targeted towards females only. Our results showed that in these female-targeted interventions, males had lower effect sizes than females, although not significantly different. All of this evidence points to the need for more interventions that directly target these factors in students, as well as interventions that consider which gender might be disadvantaged in a certain subject.

The finding that intervention duration did not significantly moderate intervention effects fits with prior research, as there seem to be mixed results on if and how the length of an intervention moderates its effects. While some studies have found small effects for intervention duration (Hattie et al., 1996; Slavin & Lake, 2009), other meta-analyses on interventions have also not found any significant effects of intervention duration (van Eerde & Klingsieck, 2018). However, given the relatively small sample size of our meta-analysis, and that we treated intervention duration as a continuous variable, it is possible that we did not have enough power to detect any effects that might be present. More research is needed to determine if the length of an intervention is overall a moderating variable of the intervention effects, and if there is an effect of length, whether or not it differs depending on student gender.

There were no significant differences between the type of motivational-affective outcome measured. Descriptively, the biggest effect size could be seen for when attitudes were the measured outcome. This may be due to the broader definition of attitudes, as compared to the other categories of outcomes measured. Measures of attitude usually encompass many various aspects such as value, self-efficacy, beliefs, and relevance (Thurstone, 1970), and therefore capture a wider range of student factors.

Limitations & future research

While this study provides a crucial first look at gender-specific effects of school-based interventions, there are certain limitations that must be considered when interpreting the results. Out of 171 potentially eligible articles, we excluded 35 due to missing data or information. This was due to two main issues. First, while many studies were eligible for inclusion in this meta-analysis based on the characteristics of the interventions themselves, many of them could not be included because they did not evaluate gender-specific effects, and therefore, the data necessary for calculating effect sizes for the different genders were not available. This highlights the need for future research to evaluate and report not only the overall effects of interventions but also the gender-specific effects, especially in subjects where there is typically a gender gap in these factors. Second, we excluded a number of studies because they did not include a control group or did not include any pre-test measures, and therefore did not fulfil our strict criteria for study design. This illustrates the greater need for high-quality intervention research in education, with stricter adherence to standards for effective intervention studies, as robust meta-analytic results rely on well-powered and well-designed primary studies.

These issues with missing information from primary studies resulted in a relatively smaller sample size, leading to a number of methodological limitations and changes to our original study design. Firstly, while we originally planned to include all moderators into one model, our final sample size did not allow us to robustly test this model. We therefore evaluated each moderator individually, which did not allow us to control for any possible confounding interactions between moderators. It is possible that relationships between moderators could have had confounding effects on the results of our analyses. More high-quality intervention studies on gender differences in interventions would allow future meta-analytic studies to control for these possible interactions, and gain a closer look at how these moderators might relate to each other.

Due to the limited sample, we also were not able to examine interventions conducted in subjects where male students are typically disadvantaged. This highlights the need for more intervention research in subjects where male students are stereotypically disadvantaged. A surge of studies have focused on increasing female students interest, performance, and participation in STEM subjects in recent years (Kanny et al., 2014). While this is undoubtedly an important topic, it is equally as important to continue research in subject where male students are disadvantaged as well. Reading, writing, and language arts is a subject where boys have consistently displayed lower levels of interest, self-efficacy, and motivation (OECD, 2019; Retelsdorf et al., 2015). Increasing intervention research in these subjects is also a crucial step in closing the gender gap between all students.

Another limitation that must be considered is the quality of the included studies. We had certain quality requirements for studies to be included, namely that they use a pre-/post-test design and include a control group. We also assessed study quality through multiple additional criteria, which were study design (experimental vs. quasi-experimental), instrument reliability (high vs. low), instrument source (established vs. self-developed), and equivalence at baseline. While almost all studies used highly reliable, established instruments to measure their outcomes, 18 of the 71 included effect sizes did not meet the recommended standards of the What Works Clearinghouse for baseline equivalence. Due to our relatively small sample size and the dependency of our data (males vs. females), we chose not to exclude these effect sizes in order to include as many primary studies as possible in an area with an already limited number of intervention studies. However, this is a point that must be taken into account when interpreting our results. Additionally, we chose to include studies that used non-random assignment when selecting treatment and control groups. In educational research, it is often difficult to use completely random assignment when conducting studies in schools or classrooms, as separating already existing groups of students is usually impractical and, at times, unethical (Gopalan et al., 2020). Due to this, many studies conducted in schools will often use pre-existing classes as a treatment and control group, making this method of group assignment quite common throughout the literature. In order to avoid excluding any potentially relevant interventions, we chose to include these studies. However, using pre-existing classes for group assignment could be a confounding factor in these studies. Even though we statistically adjusted for these clustered assignments, future research should aim to examine studies that only use random assignment in order to eliminate any potentially confounding effects.

Lastly, we had a great degree of statistical heterogeneity among primary studies. Although the moderator analyses revealed several interesting descriptive trends regarding various characteristics of school-based interventions, none of them significantly influenced the intervention effects, and we were not able to completely explain the amount of heterogeneity. There are numerous contextual characteristics that were not taken into account in this study that may have played a role in the varying effect sizes. For example, a number of other factors have been shown to be connected to both gender stereotypes, as well as motivational-affective factors, such as high-achieving vs. low-achieving students, socio-economic status and cultural background (Dietrich et al., 2013; Guo et al., 2015; Rowley et al., 2007; Rozek et al., 2015). These features were not examined in the current study due to the lack of sufficient information from primary studies and low statistical power, but they are important variables for future researchers to consider including when examining the gender-specific effects of school-based interventions.

Implications and conclusion

This meta-analysis provides a comprehensive overview of the interventions on student motivational-affective factors and their effectiveness, and sheds light on the need to investigate what strategies are most effective in promoting and strengthening these factors in student in order to gain a deeper empirical understanding. This provides an important step forward in continuing to investigate ways in which schools and teachers can combat gaps that arise between male and female students in motivation, interest, self-efficacy, enjoyment, and engagement. School-based interventions are clearly a promising method for promoting the motivational-affective factors of school-aged children and adolescents. We defined school-based interventions as any method used in a school context, which is different from regular instruction. This includes not only in-class interventions but also novel teaching methods, summer school programmes, or school-organized workshops. A variety of possibilities are there for implementing these types of interventions into a school curriculum. The results of this study show that these interventions seem to have positive effects, and that researchers should continue to pursue investigations into these interventions in order to gain a better understanding of what interventional strategies are most effective for promoting student motivational-affective factors overall. Future studies should build on this current work, using strong theoretical frameworks with regard to the development of student motivation and affect to design and test school-based interventions.

This study also provides a comprehensive overview of high-quality intervention studies to date that have evaluated the gender-specific effects of school-based interventions. Our findings provide possible evidence that gender may play a role in the effectiveness of a given intervention. This is something that is evident in a number of primary studies. For example, the study by Falco et al. (2008) designed an intervention to improve student self-efficacy beliefs through fostering various skills such as time management and goal-setting, based on social-cognitive and expectancy-value theories. This primary study found that while all students in the experimental group developed more positive attitudes compared to the control, the gains for female students were significantly higher. Researchers should continue to investigate how these interventions affect male and female students differently, and conduct more studies on promising strategies to build a strong evidence base for practitioners and policy-makers, who can then make evidence-based recommendations for best classroom practices to combat gender differences between students.

This meta-analysis also identified contextual variables of these interventions such as intervention method, student grade level, and school subject that might play a role in how effective these interventions are for students. Future research should investigate this more deeply to determine how salient these effects may be. Continued research on this topic will help to create educational settings that are more inclusive and assist all students equally in achieving their full potential, regardless of gender.

ACKNOWLEDGEMENTS

We would like to thank the student research assistants who helped with the data collection and preparation for this study. The project on which this report is based was funded by the German Federal Ministry of Education and Research and the Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Federal States in the Federal Republic of Germany [ZIB2022].

CONFLICT OF INTEREST

All authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Kaley Lesperance: Conceptualization; Data curation; Investigation; Methodology; Project administration; Resources; Writing – original draft; Writing – review & editing. **Sarah Hofer:** Conceptualization; Investigation; Methodology; Project administration; Supervision; Writing – original draft; Writing – review & editing. **Jan Retelsdorf:** Conceptualization; Writing – review & editing. **Doris Holzberger:** Conceptualization; Methodology; Project administration; Supervision; Visualization; Writing – review & editing.

DATA AVAILABILITY STATEMENT

The data supporting this meta-analysis will be from previously reported studies and datasets, which have been cited. The processed data will be made openly available upon publishing of the study.

ORCID

Kaley Lesperance https://orcid.org/0000-0002-0669-2744
Sarah Hofer https://orcid.org/0000-0001-7267-9356
Jan Retelsdorf https://orcid.org/0000-0002-9892-9613
Doris Holzberger https://orcid.org/0000-0001-6535-289X

REFERENCES

References marked with an asterisk (*) indicate studies included in the meta-analysis.

*Abed, O. H. (2016). Drama-based science teaching and its effect on students' understanding of scientific concepts and their attitudes towards science learning. *International Education Studies*, *9*(10), 163–173. https://doi.org/10.5539/ies.v9n10p163

*Akcay, H., Yager, R. E., Iskander, S. M., & Turgut, H. (2010). Change in student beliefs about attitudes toward science in grades 6-9. In *Asia-pacific forum on science learning and teaching* (Vol. *11*, No. 1, pp. 1–18). The Education University of Hong Kong, Department of Science and Environmental Studies.

*Al-Balushi, S. M., & Al-Aamri, S. S. (2014). The effect of environmental science projects on students' environmental knowledge and science attitudes. *International Research in Geographical and Environmental Education*, *23*(3), 213–227. https://doi.org/10.1080/10382046.2014.927167

*Alghamdi, A. (2017). *Impact of jigsaw on the achievement and attitudes of saudi arabian male high school science students* [Doctoral dissertation, University of Akron].

Annetta, L. A. (2008). Video games in education: Why they should be used and how they are being used. *Theory Into Practice*, *47*, 229–239. https://doi.org/10.1080/00405840802153940

*Argaw, A. S., Haile, B. B., Ayalew, B. T., & Kuma, S. G. (2016). The effect of problem based learning (PBL) instruction on students' motivation and problem solving skills of physics. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(3), 857–871. https://doi.org/10.12973/eurasia.2017.00647a

Batdi, V. (2014). The effects of a problem based learning approach on students' attitude levels: A meta-analysis. *Educational Research and Reviews*, *9*(9), 272–276. https://doi.org/10.5897/ERR2014.1771

*Batton, M. (2010). *The effect of cooperative groups on math anxiety* [Doctoral dissertation, Walden University]. Walden Dissertations and Doctoral Studies.

Begg, C. B., & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics*, *50*(4), 1088. https://doi.org/10.2307/2533446

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*, 139–153. https://doi.org/10.1207/s15326985ep3403_1

Borenstein, M., Higgins, J. P., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Bruder, R., & Prescott, A. (2013). Research evidence on the benefits of IBL. *ZDM Mathematics Education*, *45*, 811–822. https://doi.org/10.1007/s11858-013-0542-2

Burke, R., & Mattis, M. (2007). *Women and minorities in science, technology, engineering and mathematics: Upping the numbers*. Edward Elgar Publishing. https://doi.org/10.4337/9781847206879

*Cady, D., & Terrell, S. R. (2008). The effect of the integration of computing technology in a science curriculum on female students' self-efficacy attitudes. *Journal of Educational Technology Systems*, *36*(3), 277–286. https://doi.org/10.2190/ET.36.3.d

Casad, B. J., Hale, P., & Wachs, F. L. (2015). Parent-child math anxiety and math-gender stereotypes predict adolescents' math education outcomes. *Frontiers in Psychology*, *6*, 1597. https://doi.org/10.3389/fpsyg.2015.01597

*Chen, H. T., Wang, H. H., Lu, Y. Y., & Hong, Z. R. (2019). Bridging the gender gap of children's engagement in learning science and argumentation through a modified argument-driven inquiry. *International Journal of Science and Mathematics Education*, *17*(4), 635–655. https://doi.org/10.1007/s10763-018-9896-9

*Chiu, M. S. (2011). Effects of a women-in-sciences/men-in-humanities intervention on Taiwanese adolescents' attitudes towards learning science. *Asia-Pacific Education Researcher (De La Salle University Manila)*, *20*(2), 322–335.

Christensen, R. (2002). Effects of technology integration education on the attitudes of teachers and students. *Journal of Research on Technology in Education*, *34*, 411–433. https://doi.org/10.1080/15391523.2002.10782359

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220. https://doi.org//10.1037/h0026256

Dasgupta, N., & Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 21–29. https://doi.org/10.1177/2372732214549471

de Boer, H., Donker, A. S., & van der Werf, M. P. C. (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research*, *84*, 509–545. https://doi.org/10.3102/0034654314540006

Dietrich, J., Schnabel, K., Ortner, T. M., Eagly, A., Retamero-Garcia, R., Kröger, L., & Holst, E. (2013). Internalized gender stereotypes vary across socioeconomic indicators. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2273466

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*, 231–264. https://doi.org/10.1007/s11409-008-9029-x

Durik, A. M., Vida, M., & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology*, *98*, 382–393. https://doi.org/10.1037/0022-0663.98.2.382

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Eccles, J. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al model of achievement-related choices. *Psychology of Women Quarterly*, *18*, 585–609. https://doi.org/10.1111/j.1471-6402.1994.tb01049.x

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*, 830–847. https://doi.org/10.2307/1131221

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

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

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education*, *67*, 156–167. https://doi.org/10.1016/j.compedu.2013.02.019

*Fabian, K., Topping, K. J., & Barron, I. G. (2018). Using mobile technologies for mathematics: Effects on student attitudes and achievement. *Educational Technology Research and Development*, *66*(5), 1119–1139. https://doi.org/10.1007/s11423-018-9580-3

*Falco, L. D., Crethar, H., & Bauman, S. (2008). Skill-builders: Improving middle school students' self-beliefs for learning mathematics. *Professional School Counseling*, *11*(4), 229–235. https://doi.org/10.5330/PSC.n.2010-11.229

*Falco, L. D., & Summers, J. J. (2019). Improving career decision self-efficacy and STEM self-efficacy in high school girls: Evaluation of an intervention. *Journal of Career Development*, *46*(1), 62–76. https://doi.org/10.1177/0894845317721651

*Fennema, E., Pedro, J. D., Wolleat, P. L., & Becker, A. D. (1981). Increasing women's participation in mathematics: An intervention study. *Journal for Research in Mathematics Education*, *12*(1), 3–14. https://doi.org/10.2307/748654

Fisher, Z., & Tipton, E. (2015). Robumeta: An R-package for robust variance estimation in meta-analysis. *arXiv* [Preprint]. https://doi.org/10.48550/arXiv.1503.02220

Goetz, T., Frenzel, A. C., Pekrun, R., & Hall, N. C. (2006). The domain specificity of academic emotional experiences. *The Journal of Experimental Education*, *75*, 5–29. https://doi.org/10.3200/JEXE.75.1.5-29

Goetz, T., & Hall, N. C. (2013). Emotion and achievement in the classroom. In J. Hattie, & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 192–195). Routledge.

Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, *44*(1), 218–243. https://doi.org/10.3102/0091732X20903302

Green, J., Martin, A. J., & Marsh, H. W. (2007). Motivation and engagement in English, mathematics and science high school subjects: Towards an understanding of multidimensional domain specificity. *Learning and Individual Differences*, *17*, 269–279. https://doi.org/10.1016/j.lindif.2006.12.003

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, *66*, 153–166. https://doi.org/10.1007/s11199-011-9996-2

Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology*, *51*(8), 1163–1176. https://doi.org/10.1037/a0039440

Gutman, L. M., & Schoon, I. (2014). *The impact of non-cognitive skills on outcomes for young people* (pp. 1–5). Education Endowment Foundation.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, *66*(2), 99–136. https://doi.org/10.3102/00346543066002099

Heckman, J., & Kautz, T. (2013). *Fostering and measuring skills: Interventions that improve character and cognition (No. w19656)*. National Bureau of Economic Research. https://doi.org/10.3386/w19656

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*, 411–482. https://doi.org/10.1086/504455

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*, 341–370. https://doi.org/10.3102/1076998606298043

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, *29*(1), 60–87. https://doi.org/10.3102/0162373707299706

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. https://doi.org/10.1002/jrsm.5

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods*, *3*(4), 486–504. https://doi.org/10.1037/1082-989X.3.4.486

Herbert, J., & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of Applied Developmental Psychology*, *26*, 276–295. https://doi.org/10.1016/j.appdev.2005.02.007

Heyder, A., Steinmayr, R., & Kessels, U. (2019). Do teachers' beliefs about math aptitude and brilliance explain gender differences in children's math ability self-concept? *Frontiers in Education*, *4*, 34. https://doi.org/10.3389/feduc.2019.00034

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, *153*, 103897. https://doi.org/10.1016/j.compedu.2020.103897

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, *16*, 235–266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Hsu, H.-Y., Chen, S.-H., Yu, H.-Y., & Lou, J.-H. (2010). Job stress, achievement motivation and occupational burnout among male nurses. *Journal of Advanced Nursing*, *66*, 1592–1601. https://doi.org/10.1111/j.1365-2648.2010.05323.x

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*, *102*, 880–895. https://doi.org/10.1037/a0019506

IntHout, J., Ioannidis, J. P., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *British Medical Journal Open*, *6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

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

Jones, M. G., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science Education*, *84*, 180–192.

Juvonen, J. (2007). Reforming middle schools: Focus on continuity, social connectedness, and engagement. *Educational Psychologist*, *42*(4), 197–208. https://doi.org/10.1080/00461520701621046

Kanny, M. A., Sax, L. J., & Riggers-Piehl, T. A. (2014). Investigating forty years of STEM research: How explanations for the gender gap have evolved over time. *Journal of Women and Minorities in Science and Engineering*, *20*(2), 127–148. https://doi.org/10.1615/JWomenMinorScienEng.2014007246

*Kerneža, M., & Košir, K. (2016). Comics as a literary-didactic method and their use for reducing gender differences in reading literacy at the primary level of education. *Center for Educational Policy Studies Journal*, *6*(2), 125–149. https://doi.org/10.26529/cepsj.91

Kim, C., & Pekrun, R. (2014). Emotions and motivation in learning and performance. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 65–75). Springer. https://doi.org/10.1007/978-1-4614-3185-5_6

Köller, O., Baumert, J., & Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal of Research in Mathematics Education*, *32*, 448–470. https://doi.org/10.2307/749801

Laal, M., & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia - Social and Behavioral Sciences*, *31*, 486–490. https://doi.org/10.1016/j.sbspro.2011.12.091

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Leaper, C. (2015). Gender and social-cognitive development. In R. M. Lerner, L. S. Liben, & U. Muller (Eds.), *Handbook of child psychology and developmental science* (pp. 1–48). American Cancer Society. https://doi.org/10.1002/9781118963418.childpsy219

Lieberman, D. A., Bates, C. H., & So, J. (2009). Young children's learning with digital media. *Computers in the Schools*, *26*, 271–283. https://doi.org/10.1080/07380560903360194

Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social Science Research*, *37*, 888–902. https://doi.org/10.1016/j.ssresearch.2008.03.004

Logan, S., & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading*, *32*, 199–214. https://doi.org/10.1111/j.1467-9817.2008.01389.x

Marinak, B. A., & Gambrell, L. B. (2010). Reading motivation: Exploring the elementary gender gap. *Literacy Research and Instruction*, *49*, 129–141. https://doi.org/10.1080/19388070902803795

Marsh, H. W., Martin, A. J., & Debus, R. (2001). Individual differences in verbal and math self-perceptions: One factor, two factors, or does it depend on the construct. *International Perspectives on Individual Differences*, *2*, 149–170.

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, *52*, 1119–1134. https://doi.org/10.2307/1129498

Master, A., Cheryan, S., Moscatelli, A., & Meltzoff, A. N. (2017). Programming experience promotes higher STEM motivation among first-grade girls. *Journal of Experimental Child Psychology*, *160*, 92–106. https://doi.org/10.1016/j.jecp.2017.03.013

Mavridis, A., Katmada, A., & Tsiatsos, T. (2017). Impact of online flexible games on students' attitude towards mathematics. *Educational Technology Research and Development*, *65*(6), 1451–1470. https://doi.org/10.1007/s11423-017-9522-5

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science*, *11*(5), 730–749. https://doi.org/10.1177/1745691616662243

Miller, P. H., Slawinski Blessing, J., & Schwartz, S. (2006). Gender differences in high-school students' views about science. *International Journal of Science Education*, *28*(4), 363–381. https://doi.org/10.1080/09500690500277664

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*, 264–269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90*(3), 376–419. https://doi.org/10.3102/0034654320919354

Morgenroth, T., Ryan, M. K., & Peters, K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology*, *19*, 465–483. https://doi.org/10.1037/gpr0000059

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*, 364–386. https://doi.org/10.1177/1094428106291059

Muntoni, F., & Retelsdorf, J. (2018). Gender-specific teacher expectations in reading—The role of teachers' gender stereotypes. *Contemporary Educational Psychology*, *54*, 212–220. https://doi.org/10.1016/j.cedpsych.2018.06.012

Muntoni, F., Wagner, J., & Retelsdorf, J. (2020). Beware of stereotypes: Are classmates' stereotypes associated with students' reading outcomes? *Child Development*, *92*(1), 189–204. https://doi.org/10.1111/cdev.13359

Murphy, K. P., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology*, *25*, 3–53. https://doi.org/10.1006/ceps.1999.1019

OECD. (2013). Mathematics self-beliefs and participation in mathematics-related activities. In *OECD, PISA 2012 results: Ready to learn* (Vol. *III*, pp. 87–112). OECD. https://doi.org/10.1787/9789264201170-8-en

OECD. (2014). *Indicator D5: Who are the teachers (Education at a Glance 2014: OECD Indicators)*. OECD Publishing. https://doi.org/10.1787/eag-2018-en

OECD. (2016). *PISA 2015 results (Volume I): Excellence and equity in education* (Vol. *1*). OECD. https://doi.org/10.1787/9789264266490-en

OECD. (2019). *PISA 2018 results (Volume II): Where all students can succeed*. OECD. https://doi.org/10.1787/b5fd1b8f-en

*Orabuchi, N. (2013). *Effects of online visual and interactive technological tool (OVITT) on early adolescent students' mathematics performance, math anxiety and attitudes toward math* [Doctoral dissertation, Texas Woman's University]. ProQuest Dissertations Publishing.

Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. In A. M. Gallagher, & J. C. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 294–315). Cambridge University Press. https://doi.org/10.18848/2327-7920/CGP/v23i03/1-9

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, *34*, 29–48. https://doi.org/10.1080/01443410.2013.797339

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology*, *36*, 36–48. https://doi.org/10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*, 91–105. https://doi.org/10.1207/S15326985EP3702_4

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*, 1653–1670. https://doi.org/10.1111/cdev.12704

Pintrich, P. R. (2003). Motivation and classroom learning. In W. M. Reynolds, & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (pp. 103–122). Wiley. https://doi.org/10.1002/0471264385.wei0706

Plante, I., de la Sablonnière, R., Aronson, J. M., & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology*, *38*, 225–235. https://doi.org/10.1016/j.cedpsych.2013.03.004

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology*, *107*(1), 186–194. https://doi.org/10.1037/a0037107

Rosen, J. A., Glennie, E. J., Dalton, B. W., Lennon, J. M., & Bozick, R. N. (2010). *Noncognitive skills in the classroom: New perspectives on educational research*. RTI International.

Rowley, S. J., Kurtz-Costes, B., Mistry, R., & Feagans, L. (2007). Social status as a predictor of race and gender stereotypes in late childhood and early adolescence. *Social Development*, *16*, 150–168. https://doi.org/10.1111/j.1467-9507.2007.00376.x

Rozek, C. S., Hyde, J. S., Svoboda, R. C., Hulleman, C. S., & Harackiewicz, J. M. (2015). Gender differences in the effects of a utility-value intervention to help parents motivate adolescents in mathematics and science. *Journal of Educational Psychology*, *107*(1), 195–206. https://doi.org/10.1037/a0036981

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

Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. Organisation for Economic Co-operation and Development.

Schmenk, B. (2004). Language learning: A feminine domain? The role of stereotyping in constructing gendered learner identities. *TESOL Quarterly*, *38*, 514–524. https://doi.org/10.2307/3588352

Shapiro, J. R., & Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles*, *66*, 175–183. https://doi.org/10.1007/s11199-011-0051-0

*Shin, D. D., Lee, M., Ha, J. E., Park, J. H., Ahn, H. S., Son, E., Chung, Y., & Bong, M. (2019). Science for all: Boosting the science motivation of elementary school students with utility value intervention. *Learning and Instruction*, *60*, 104–116. https://doi.org/10.1016/j.learninstruc.2018.12.003

Slavin, R. E. (2011). Instruction based on cooperative learning. In R. E. Mayer, & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 344–360). Routledge. https://doi.org/10.4324/9780203839089.ch17

Slavin, R. E., & Lake, C. (2009). *Effective educational programs: Meta-findings from the best evidence encyclopedia*. Annual Meetings of the Society for Research on Educational Effectiveness, 8.

*Soyibo, K., & Hudson, A. (2000). Effects of computer-assisted instruction (CAI) on 11th graders' attitudes to biology and CAI and understanding of reproduction in plants and animals. *Research in Science & Technological Education*, *18*(2), 191–199. https://doi.org/10.1080/713694977

*Starkey, P. (2013). *The effects of digital games on middle school students' mathematical achievement* [Doctoral dissertation, Lehigh University]. ProQuest Dissertations Publishing.

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*(1), 80–90. https://doi.org/10.1016/j.lindif.2008.05.004

*Sung, Y. T., Shih, P. C., & Chang, K. E. (2015). The effects of 3D-representation instruction on composite-solid surface-area learning for elementary school students. *Instructional Science*, *43*(1), 115–145. https://doi.org/10.1007/s11251-014-9331-8

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology*, *2*(1), 85–112. https://doi.org/10.1007/s40865-016-0026-5

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, *88*(4), 1156–1171. https://doi.org/10.1111/cdev.12864

Thurstone, L. L. (1970). Attitudes can be measured. In G. F. Summers (Ed.), *Attitude measurement* (pp. 127–141). Rand McNally.

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, *92*(1), 144–151. https://doi.org/10.1037/0022-0663.92.1.144

Uzunboylu, H., & Karagozlu, D. (2015). Flipped classroom: A review of recent literature. *World Journal on Educational Technology*, *7*, 142–147. https://doi.org/10.18844/wjet.v7i2.46

*van der Meij, H., van der Meij, J., & Harmsen, R. (2015). Animated pedagogical agents effects on enhancing student motivation and learning in a science inquiry learning environment. *Educational Technology Research and Development*, *63*(3), 381–403. https://doi.org/10.1007/s11423-015-9378-5

van Eerde, W., & Klingsieck, K. B. (2018). Overcoming procrastination? A meta-analysis of intervention studies. *Educational Research Review*, *25*, 73–85. https://doi.org/10.1016/j.edurev.2018.09.002

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, *60*(3), 419–435. https://doi.org/10.1007/BF02294384

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, *23*, 73–82. https://doi.org/10.1177/0963721413512856

Walton, G. M., & Wilson, T. D. (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological Review*, *125*(5), 617. https://doi.org/10.1037/rev0000115

Wang, S.-K., & Reeves, T. C. (2007). The effects of a web-based learning environment on student motivation in a high school earth science course. *Educational Technology Research and Development*, *55*, 169–192. https://doi.org/10.1007/s11423-006-0638-2

Weinburgh, M. (1995). Gender differences in student attitudes toward science: A meta-analysis of the literature from 1970 to 1991. *Journal of Research in Science Teaching*, *32*, 387–398. https://doi.org/10.1002/tea.3660320407

Wentzel, K. R., & Wigfield, A. (2007). Motivational interventions that work: Themes and remaining issues. *Educational Psychologist*, *42*(4), 261–271. https://doi.org/10.1080/00461520701621103

What Works Clearinghouse. (2020). *Standards handbook* (Version 4.1). https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

Wigfield, A., Battle, A., Keller, L. B., & Eccles, J. S. (2002). Sex differences in motivation, self-concept, career aspiration, and career choice: Implications for cognitive development. In A. V. McGillicuddy, & R. De Lisi (Eds.), *Biology, society, and behavior: The development of sex differences in cognition* (Vol. *21*, pp. 93–124). Ablex.

Wigfield, A., & Eccles, J. S. (1994). Children's competence beliefs, achievement values, and general self-esteem: Change across elementary and middle School. *Journal of Early Adolescence*, *14*, 107–138. https://doi.org/10.1177/027243169401400203

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81. https://doi.org/10.1006/ceps.1999.1015

Wigfield, A., Guthrie, J. T., Tonks, S., & Perencevich, K. C. (2004). Children's motivation for reading: Domain specificity and instructional influences. *Journal of Educational Research*, *97*, 299–310. https://doi.org/10.3200/JOER.97.6.299-310

*Zhao, F., Zhang, Y., Alterman, V., Zhang, B., & Yu, G. (2018). Can math-gender stereotypes be reduced? A theory-based intervention program with adolescent girls. *Current Psychology*, *37*(3), 612–624. https://doi.org/10.1007/s12144-016-9543-y

How to cite this article: Lesperance, K., Hofer, S., Retelsdorf, J., & Holzberger, D. (2022). Reducing gender differences in student motivational-affective factors: A meta-analysis of school-based interventions. *British Journal of Educational Psychology*, *00*, 1–35. https://doi.org/10.1111/bjep.12512

APPENDIX A

Search syntax for ERIC

(SU (students) AND SU (education* OR school OR classroom OR parents OR teachers) AND SU (female OR male OR gender differences OR sex stereotypes OR sex fairness OR gender bias OR gender issues OR sex role) AND AB ("intervention*" OR "mentor*" OR "role model*" OR "training*" OR "program*" OR "instruction*" OR "strategy*" OR "support*" OR "outreach*" OR "teaching*" OR "experiment*" OR "control group*") AND SU (interest OR self concept OR self esteem OR self efficacy OR motivation OR attribution theory OR stereotype OR career choice OR attitude OR beliefs OR values OR learner engagement OR participation OR satisfaction)) NOT SU (higher education OR universities OR college*)

Search syntax for PsycINFO

(SU (students) AND SU (education* OR school OR classroom OR parents OR teachers) AND SU ("human females" OR "human males" OR sex differences OR gender gap OR gender equality OR sex role) AND AB ("intervention*" OR "mentor*" OR "role model*" OR "training" OR "program" OR "instruction" OR "strategy*" OR "support" OR "outreach" OR "teaching" OR "experiment*" OR "control group*") AND SU (belonging OR interest OR self concept OR self-confidence OR self-esteem OR self-efficacy OR motivation OR attribution OR occupational choice OR attitude OR occupational preference OR values OR student engagement OR cognitive appraisal OR participation OR expectations)) NOT SU (higher education OR college*)

Search syntax for Web of Science

((((TS=(student*) AND TS=("education*" OR "school*" OR "classroom*" OR "teacher*" OR "parent*") AND TS=("gender gap*" OR "gender difference*" OR "gender stereotype*" OR "gender equality*" OR "gender bias*" OR "gender specific*" OR "sex difference*" OR "sex stereotype*" OR "sex role*") AND TS=("intervention*" OR "mentor*" OR "role model*" OR "training*" OR "program*" OR "instruction*" OR "strategy*" OR "support*" OR "outreach*" OR "teaching*" OR "experiment*" OR "control group*") AND TS=("interest*" OR "belonging*" OR "self concept*" OR "self efficacy*" OR "self confidence*" OR "self esteem*" OR "motivation*" OR "attribution*" OR "stereotype*" OR "career choice*" OR "attitude*" OR "belief*" OR "value*" OR "engagement*" OR "participation*" OR "expectation*")))))

APPENDIX B

Subgroup analyses

TABLE A1 Meta-analytic subgroup model for females in science

			95% CI			
Predictor	**Estimate**	***SE***	**Lower**	**Upper**	***df***	***p*-value**
Intercept	.389	.135	.052	.725	5.64	.0303**

Note: Model estimated using robust variance estimation to correct for dependent effect sizes. **$p < .05$.

TABLE A2 Meta-analytic subgroup model for males in science

			95% CI			
Predictor	**Estimate**	***SE***	**Lower**	**Upper**	***df***	***p*-value**
Intercept	.247	.187	−.211	.705	5.91	.234

Note: Model estimated using robust variance estimation to correct for dependent effect sizes.

TABLE A3 Meta-analytic subgroup model for females in math

			95% CI			
Predictor	**Estimate**	***SE***	**Lower**	**Upper**	***df***	***p*-value**
Intercept	.379	.163	.001	.756	7.85	.049*

Note: Model estimated using robust variance estimation to correct for dependent effect sizes. *$p < .10$.

TABLE A4 Meta-analytic subgroup model for males in math

			95% CI			
Predictor	**Estimate**	***SE***	**Lower**	**Upper**	***df***	***p*-value**
Intercept	.227	.092	−.004	.457	5.51	.053*

Note: Model estimated using robust variance estimation to correct for dependent effect sizes. *$p < .10$.

APPENDIX C

TABLE A5 Sensitivity analysis for baseline equivalency

			95% CI			
Predictor	**Estimate**	***SE***	**Lower**	**Upper**	***df***	***p*-value**
Intercept	−.127	.069	−.278	.0.024	11.3	.0909*
Unequal baseline	.526	.080	.337	.716	6.9	.0003***

Note: Model estimated using robust variance estimation to correct for dependent effect sizes. * $p < .10$, ***$p < .01$.

Appendix B – Study II

Lesperance, K., Decristan, J., & Holzberger, D. (2023). The role of teacher constructive support for gender differences in motivational outcomes in secondary school mathematics. *International Journal of Gender, Science and Technology, 15*(3).

International Journal of Gender, Science and Technology
http://genderandset.open.ac.uk

Selected papers presented at the 6th Network Gender & STEM Conference, 21-23 July 2022 in München, Germany

In association with

NETWORK GENDER & STEM
educational and occupational pathways and participation

The Role of Teacher Constructive Support for Gender Differences in Motivational Outcomes in Secondary School Mathematics

Kaley Lesperance[1], Jasmin Decristan[2] & Doris Holzberger[1]

[1]Technical University of Munich, TUM School of Social Sciences and Technology, Centre for International Student Assessment (ZIB), [2]University of Wuppertal, Institute of Educational Research, School of Education

ABSTRACT
In secondary school mathematics, females often display lower levels of motivational outcomes than males, which can lead to gender gaps in future study or career choice. To reduce these gaps, it is crucial to evaluate which aspects of classroom teaching quality might be involved. Teacher constructive support is one especially promising aspect, as it strongly relates to student motivational outcomes. The present pre-registered study investigates how student gender is related to self-concept, self-efficacy, and interest in secondary school mathematics lessons on the quadratic equation, and examines the moderating role of teacher constructive support for this relationship. Using questionnaire data from the Teaching and Learning International Survey Video Study Germany (n = 1,116 secondary school students), we applied latent moderated structural equation models to examine the direct effects as well as the interaction of student gender and constructive support on student motivational outcomes in mathematics lessons. Female gender was negatively associated with self-concept and self-efficacy, but not with interest. Various facets of constructive support were positively associated with motivational outcomes, but no interaction effects with gender were found. These findings are discussed in regards to constructive support and the persisting gender gap in mathematics. Directions for future research are suggested.

KEYWORDS
Teacher support; gender; motivation; mathematics; secondary school

The Role of Teacher Constructive Support for Gender Differences in Motivational Outcomes in Secondary School Mathematics

INTRODUCTION

Male and female students tend to differ in mathematics in regards to their motivational outcomes, with female students often displaying lower levels of these outcomes, especially in secondary school (Else-Quest et al., 2010). Research shows that these outcomes can in turn affect students' achievement, as well as future academic and career choices (Pintrich, 2003; Wigfield & Eccles, 2000). In order to understand and reduce these gender gaps, it is important to investigate what aspects of students' daily experiences in school might alleviate the effects of gender on their motivational outcomes. Students spend a large portion of their lives in the classroom, and various aspects of the classroom environment have strong effects on a plethora of student outcomes, including motivational outcomes such as self-concept, self-efficacy, and interest (Scherer & Nilsen, 2016). Teacher constructive support, which is defined as positive teacher-student relationships regarding both instructional and emotional matters, is one such aspect of the classroom environment that is linked to these motivational factors (Fauth et al., 2014; Ryan & Patrick, 2001). Students who perceive teachers as more supportive and caring tend to have higher self-concept, self-efficacy, and interest in subjects such as mathematics (Yu & Singh, 2018). However, what is still unclear is whether there are differential effects of perceived constructive support on these student outcomes in regards to student gender. The goal of this study is to investigate whether perceived teacher constructive support plays a moderating role between gender and student motivational outcomes in the specific learning context of quadratic equations in secondary school mathematic classrooms. By doing so, this study aims to contribute to research which identifies classroom factors that could help to reduce gender differences in motivational outcomes in mathematics, which to date remains relatively underinvestigated.

Theoretical background

Gender differences in mathematics self-concept, self-efficacy, and interest

Gender differences in mathematics are traditionally discussed in terms of achievement outcomes such as grades or exam scores. However, research has shown that, on average, male and female students barely differ in these outcomes, if at all (Else-Quest et al., 2010; Hyde et al., 2008; OECD, 2013; Wang et al., 2013). Where research does find substantial gender differences is in regards to student motivational outcomes (Kollmayer et al., 2018; Wigfield et al., 2002). When discussing motivational outcomes in educational contexts, many different constructs are often classified under this term. Generally, motivational outcomes can be defined as student self-beliefs, values, and goals that relate to their choices, persistence, and achievement in academic settings (Wigfield et al., 2012). Gender differences in these motivational outcomes are concerning, as they have been shown to be strongly connected to further academic choices such as postsecondary study track and career choices. These differences can often be seen in mathematics classrooms (Köller et al., 2001; Möller et al., 2020; Parker et al., 2014).

Eccles' expectancy-value motivation model proposes a widely supported explanation for how these motivational outcomes might affect student choices and behaviors. This theory posits that achievement motivation is determined by a combination of expectancy components and values for tasks in particular domains (Eccles & Wigfield, 2020; Wigfield, 1994). Expectancies for success, also commonly referred to as competency beliefs, are closely related to constructs such as self-concept and self-efficacy. These constructs pertain to how well students expect to perform in certain domains, and their perceptions of their own competence. Meanwhile, task value constructs generally refer to how useful, important, or interesting students perceive certain domains to be (Wigfield & Eccles, 2000). Following this expectancy-value model, we chose to focus our study on the motivational outcomes of self-concept, self-efficacy, and interest in order to encompass both expectancy-related (self-concept and self-efficacy) as well as value-related (interest) constructs.

Self-efficacy and self-concept both involve self-beliefs and therefore are strongly related. Self-efficacy refers to an individual's belief in their own abilities to perform certain tasks and produce desired outcomes, and the degree of confidence in one's capabilities to utilize their skills, knowledge, and resources to accomplish goals (Schunk & Pajares, 2002). Self-concept, on the other hand, is a much broader belief or perception one holds about themselves, including self-perceptions about their own abilities, characteristics, values, and affect (Bong & Clark, 1999). Both self-efficacy and self-concept are often used as indicators for the expectancy component, but slightly differ from expectancies for success, which are defined as beliefs regarding the consequences that a specific behavior will produce with respect to a task outcome (Eccles & Wigfield, 2020). Interest can be defined as the long-term orientation of a person toward an object, activity, or field of knowledge, which involves both positive emotions and positive value attributions (Schiefele, 1991). Interest is strongly related to the intrinsic value that an individual holds for a specific task or activity (Eccles, 2005).

All of these motivational outcomes are domain-specific, meaning that they are linked to the academic subject in question and can vary across different subjects (Brunner, 2008; Hornstra et al., 2016). For example, a student may have high self-efficacy in reading, but low self-efficacy in mathematics. Studies have attempted to understand how gender differences arise between students in these motivational outcomes. Differences seem to stem from multiple social, environmental, and systemic factors that influence an individuals' behaviors, interests, and self-beliefs, especially when it comes to gender-stereotyped activities (Wang & Degol, 2013). As children develop and interact with their social environment, gender stereotypes are acquired from significant people in their lives, such as parents, teachers, and peers (Smith & Farkas, 2022; Tiedemann, 2000), as well as from society and culture (Kollmayer et al., 2018). These learned gender stereotypes can have an impact on children's identity, choices, behaviors, and beliefs (Martin & Halverson, 1981).

Numerous theoretical paradigms have been proposed to explain the underlying mechanisms behind these observed effects. One prominent explanation pertains to the concept of stereotype threat. Stereotype threat theory posits that when students are aware of negative stereotypes concerning their social or demographic group (e.g., gender) in certain domains, they may feel concerned about fulfilling that stereotype, leading to decreased motivation, more negative self-beliefs, and increased anxiety in that domain (Fogliati & Bussey, 2013; Thoman et al., 2013). Certain themes and subjects may be more affected by gender stereotypes than others, and gender differences in motivational outcomes can usually be seen more often in subjects that are stereotypically favored towards one gender or the other.

Mathematics is one such subject that is consistently seen as a stereotypically "male" subject (Makarova et al., 2019). In mathematics, female students tend to display lower levels of motivational outcomes when compared to male students (Pajares, 2005). These differences have been shown to increase and become more pronounced as students get older (Bharadwaj et al., 2016; Contini et al., 2017), which can be attributed to an increasing endorsement of traditional gender stereotypes (Rowley et al., 2007) as well as an overall decrease in students' motivational outcomes after the transition to secondary school (Frenzel et al., 2010; Plenty & Heubeck, 2013). Results from the Program for International Student Assessment (PISA) 2012 showed that, on average, across 15-year old students in 72 countries, female students reported lower levels of math self-efficacy and self-concept, and more negative math attitudes as compared to male students (OECD, 2013). More recent results from PISA 2018 also demonstrated that, on average, only one percent of female students reported that they aspired to pursue a mathematics- or science-related career, compared to eight percent of male students (Schleicher, 2019). This is especially interesting in light of PISA 2018 mathematics performance results, which showed that male and female students barely differed in regards to math achievement and ability. Although there is currently no research explicitly exploring gender differences in specific sub-disciplines of math such as algebra or calculus, gender differences in self-concept, self-efficacy, and interest have been found in numerous studies spanning a range of mathematical content (Barth & Masters, 2020; Frenzel et al., 2010; Goldman & Penner, 2016; Preckel et al., 2008; Wang, 2012). It can therefore be inferred that these gender differences are likely to be present in various mathematical sub-disciplines.

In light of the clear evidence that female students consistently display lower levels of motivational outcomes than male students in mathematics, and the established importance of these outcomes for future academic and career choices, it is crucial to identify factors that can positively promote motivational outcomes for female students in math. While various aspects of one's environment such as family, peers, and culture do have a large impact on academic outcomes, the school environment and students' experience in the classroom play arguably one of the largest roles in shaping motivation and educational beliefs (Tiedemann, 2000; Wigfield & Harold, 1992). As teachers and their lessons are the main focus of most time spent in the classroom, teaching quality has a considerable impact on both achievement-related and motivational outcomes of students (Burić & Kim, 2020; Yang & Kaiser, 2022).

Therefore, an important step in reducing gender differences is identifying what aspects of teaching quality could serve to positively promote motivational outcomes, especially for female students who are typically at risk of having lower levels of these outcomes in mathematics.

In considering the measurement of motivational outcomes, it is important to differentiate between these outcomes as traits and states. While motivational outcomes as traits reflect long-term tendencies in perceptions of abilities across various academic situations, motivational outcomes can also be measured as states, which capture perceptions specific to particular learning experiences or domains (Wasserman & Wasserman, 2020). Indeed, various motivational outcomes have been treated as both trait and state in the literature (Hausen et al., 2022; Soland et al., 2019). Evidence has also shown that these motivational outcomes, when measured as states, can be influenced by situation-specific factors and contextual interactions, especially in the context of classroom instruction (Gaspard & Lauermann, 2021). This study therefore focuses on assessing students' motivational outcomes as states in relation to two focal lessons on quadratic equations. By examining these motivational outcomes as states, we aimed to explore how perceived constructive support might influence students' immediate motivational outcomes within the targeted domain of quadratic equations, providing insights into the potential impact of support factors in a specific academic context.

Teacher constructive support and student self-concept, self-efficacy, and interest
Teacher constructive support is one such aspect of teaching quality that has been proven to be extremely important for student motivational factors (Cornelius-White, 2007). Constructive support is a rather broad construct in the literature and has also been referred to as supportive climate (Klieme et al., 2009) or positive climate (Burić & Kim, 2020). Constructive support can be defined as the quality of social interactions between teachers and students in the classroom, and to what degree they are characterized by interest, respect, support, and productive feedback (Fauth et al., 2014; Praetorius et al., 2014). Teachers who provide constructive support treat students with courtesy and warmth when correcting errors and giving feedback, allow for differentiation and adapt to individual needs, and strive to foster positive student relationships (Praetorius et al., 2018). Given its broad nature, constructive support can be further separated into two facets, namely instructional support and social-emotional support. Whereas instructional support refers to teachers who care about student learning and want to help them learn, social-emotional support refers to teachers who care about students on a personal level and provide emotional support (Patrick et al., 2007). Some examples of instructional support are aiding students with content or instruction-related problems, or providing individualized support and consistent feedback. Examples of social-emotional support include empathizing with student struggles and establishing safe emotional dynamics in the classroom (Hamre & Pianta, 2005). While empirically distinguishable, both facets are strongly correlated and can be grouped under the overarching measure of constructive support (Decristan et al., 2022; Wentzel, 1997).

Given that both facets of constructive support are mostly interpersonal and emotional by nature, it is not surprising that perceived constructive support has been shown to be the aspect of teaching quality most strongly related to students' motivation and enjoyment of a subject, as well as most influential for their self-beliefs (Allen et al., 2006; Cornelius-White, 2007). According to the self-determination theory of motivation, individuals have three basic psychological needs that relate to their motivation: relatedness, competence, and autonomy (Reeve et al., 2004). When students perceive teachers as involved, encouraging, and interested in both their learning and emotional well-being, this helps to fulfill their need for relatedness, competence, and autonomy (Martin & Dowson, 2009). Students' sense of relatedness is fostered when they feel an atmosphere of connection, acceptance, and belonging within the classroom environment (Ryan & Powelson, 1991). When teachers provide feedback, guidance, and support for students' progress and well-being, it also enhances students' beliefs in their own abilities and can increase their feelings of competence (Niemiec & Ryan, 2009). Finally, by showing interest in students' learning progress, individual perspectives, and ideas, teachers encourage students to exercise greater self-direction and be active in their own learning processes, addressing their need for autonomy (Ruzek et al., 2016). These mechanisms can then lead students to experience higher levels of motivational outcomes and increased engagement (Reeve, 2012; Skinner & Belmont, 1993). Indeed, multiple studies have shown positive effects of perceived constructive support on a plethora of student academic outcomes. Students who perceive their teachers as more supportive have been shown to display more interest in the subject in question (Fauth et al., 2014; Lazarides & Ittel, 2013). Perceived constructive support has also been shown to positively relate to student self-efficacy (Fast et al., 2010; Sakiz et al., 2012) and self-concept (Demaray et al., 2009; McFarland et al., 2016). How supported students feel by their teachers, both instructionally and emotionally, has also been shown to affect their engagement, self-esteem, and intrinsic motivation (Patrick et al., 2007; Ryan et al., 1994; Wang & Eccles, 2012).

When measuring constructive support, or any other aspects of teaching quality, researchers tend to rely on either external observers, teacher perceptions, or student perceptions. While each method has both advantages and disadvantages, when aiming to investigate the effects of constructive support on student outcomes, student perceptions may be considered the more appropriate method, as whether or not instruction is perceived as supportive is something students are best able to judge for themselves (Göllner et al., 2021). Therefore, in this study, we focus on student perceptions of constructive support.

While it is quite established in the literature that constructive support is important for student motivational-affective outcomes (Sabol & Pianta, 2012), it remains unclear whether these effects are the same for both genders, or whether they differ for male and female students. Researchers in this field have called for more investigation into gender differences in the relationship between student outcomes and constructive support (Rueger et al., 2008).

Teacher constructive support as a moderator between gender and student motivational outcomes

When discussing the effects of constructive support on student motivational outcomes, it is important to consider that students have individual differences and pre-existing characteristics. This perspective is in line with a central paradigm of psychology research, known as the aptitude-treatment-interaction, which states that aspects of a treatment will influence individuals differently depending on their pre-existing cognitive and motivational-affective characteristics (Snow & Swanson, 1992). This paradigm has also been applied to educational research, with aspects of teaching quality considered "treatments" and student characteristics considered "aptitudes" (Kieft et al., 2008). As gender differences have been shown to exist in student motivational outcomes from an early age, this implies that males and females have different preconditions in terms of motivation, which could in turn lead to differential effects of teaching quality.

These differing effects are especially important to investigate when considering certain groups that are in jeopardy of low motivation in mathematics. Specifically in regards to constructive support, there is some evidence to suggest that it may be more important for some students than for others (Curby et al., 2009; Decristan et al., 2016). For example, Malecki and Demaray (2006) found that perceived teacher support had a stronger relation to student academic outcomes for students from lower socio-economic backgrounds than for those from higher socio-economic backgrounds. Hamre and Piante (2005) explained this effect as the academic risk perspective, which posits that relational assets in the environment may have a greater influence on student outcomes for students who are already at risk of having lower levels of those outcomes. Applied to the mathematics classroom, research shows that female students are typically at risk of displaying lower levels of motivational outcomes such as self-concept, self-efficacy, and interest. Due to these lower levels of motivational outcomes, female students should be especially supported in mathematics classrooms.

When viewed through the lens of mathematics classrooms, there are various positive mechanisms of constructive support that may be particularly relevant for female students. Female students tend to experience feelings of lower competence and confidence in mathematics, and may thus benefit more from perceived constructive support. Both instructional and social-emotional support from teachers can help create an atmosphere where students feel more inclined to explore and engage (Birch & Ladd, 1997; Furrer & Skinner, 2003). As studies have shown that for girls, warm and caring climates are more important to their motivation and engagement in mathematics than for boys (Fredricks et al., 2018; Rueger et al., 2008), the positive effects of supportive classroom environments may be especially pronounced for female students. Research has also shown that in mathematics, female students tend to feel less of a sense of belonging, which is another important aspect related to their competency beliefs and interest (Dasgupta & Stout, 2014; Good et al., 2012). As supportive teachers who provide feedback and make students feel respected can also lead to a deeper sense of belonging in the classroom (Liu et al., 2018), this may also be a mechanism which is especially important for female students in mathematics. Lastly, students who feel supported by their teachers report feeling more self-assured and less afraid of making mistakes or asking for help (Hughes & Chen, 2011). By providing constructive support, teachers create a safe space for students to experience failure, and students are also less likely to attribute mistakes to their lack of ability (Ryan & Patrick, 2001). Given that female students are more likely to attribute their failures in math to their own ability and display lower competency beliefs in mathematics (Dickhäuser & Meyer, 2006; Herbert & Stipek, 2005), this is also a relationship that could be exceptionally relevant for female students' motivational outcomes. Despite theoretical explanations for why constructive support may be particularly relevant for female students, there remains a sparse number of studies that have empirically investigated this relationship. A small number of studies have provided some preliminary evidence that support from teachers in general may have differential effects on male and female students' motivational outcomes. For example, McFarland et al. (2016) found that perceived closeness in student-teacher relationships significantly predicted general self-concept for girls, but not for boys, in primary school. There is also evidence that these differential effects may be seen in stereotypically "gendered" subjects. Vekiri (2010), for example, found a stronger association for girls than for boys between perceived teacher support and competence beliefs in middle school information technology classrooms, which is a stereotypically "male" subject. Additionally, Hochweber and Vieluf (2018) found that higher levels of teacher support were related to smaller gender differences in reading enjoyment (a stereotypically female subject) for ninth grade students. In one of the few studies that has examined these relationships in mathematics, Fredericks and colleagues (2018) found that teacher social-emotional support was more strongly related to girls' behavioral engagement in mathematics, and teacher instructional support was more strongly related to girls' emotional engagement than boys.

The present study

Research consistently points to gender differences in students' motivation in mathematics. Although several approaches have been implemented to positively affect particularly female students' mathematical motivation, there has been little research related to regular classroom instruction. This study thus aims to examine the role of student perceptions of constructive support for motivation in mathematics in general, and the connection between students' gender and motivational outcomes in particular in order to understand whether constructive support reduces gender differences.

The research goal of this study is to investigate the effects of gender and both facets of perceived constructive support (i.e., instructional and social-emotional) on student motivational outcomes, as well as whether the two facets of constructive support moderate the relationship between gender and student motivational outcomes. Therefore, we evaluated the following research questions (RQ) and hypotheses (H):

RQ1: Do female students have significantly lower levels of self-concept, self-efficacy, and interest than male students in secondary school mathematics lessons on quadratic equations?

> H1: We hypothesize that male and female students significantly differ in their self-concept, self-efficacy, and interest in mathematics lessons on quadratic equations. We specifically hypothesize that female students will display lower levels of self-concept, self-efficacy, and interest on average than male students.

RQ2: Are student perceptions of constructive support related to secondary school students' self-concept, self-efficacy, and interest in secondary school mathematics lessons on quadratic equations?

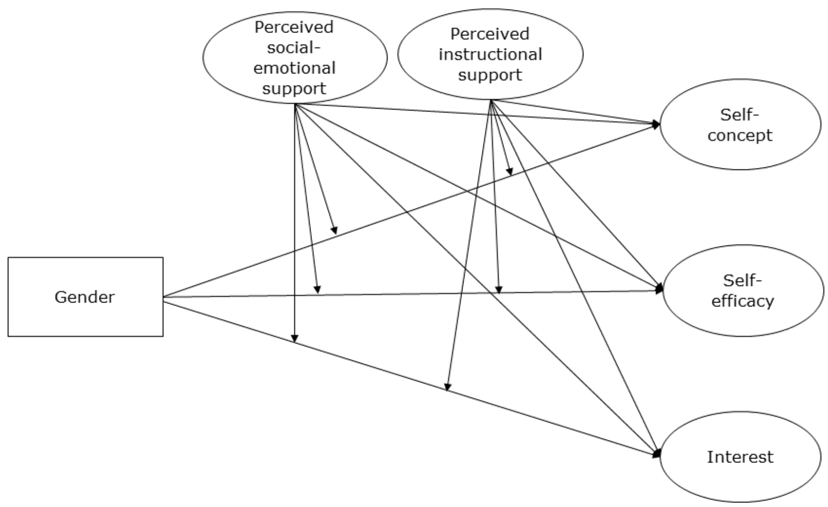
> H2: We hypothesize that student perceptions of constructive support (i.e., instructional and social-emotional) are significantly and positively related to student self-concept, self-efficacy, and interest in mathematics lessons on quadratic equations.

RQ3: Do student perceptions of constructive support moderate the relationship between gender and students' self-concept, self-efficacy, and interest in secondary school mathematics lessons on quadratic equations?

> H3: We hypothesize that student perceptions of constructive support (i.e., instructional and social-emotional), moderate the relationship between gender and student self-concept, self-efficacy, and interest in mathematics lessons on quadratic equations in secondary school. Specifically, we hypothesize that higher levels of perceived constructive support will have a stronger, positive effect on the relationship between gender and student self-concept, self-efficacy, and interest for female students than for male students.

Figure 1 shows the assumed relationship between gender and the two facets of perceived constructive support with students' motivational outcomes as examined in RQ3. This study is pre-registered, and all hypothesis and planned analyses were uploaded to the Open Science Framework (OSF) platform prior to conducting the data analyses. The pre-registration can be viewed via the following link: https://osf.io/q9bej/?view_only=499c86b19d964d018ed8520bee753430

Figure 1
Interplay between gender and perceived constructive support and its effects on motivational outcomes



Note. The oval shapes represent latent variables and the rectangle shapes represent manifest variables.

METHODS

Sample and procedure

This study is a secondary analysis of data from the German sample of the Teaching and Learning International Survey (TALIS) Video Study Germany, an international field study in secondary school mathematics education (OECD, 2020a, 2020b). The full data contains coded videos of lessons, as well as student and teacher questionnaires and student achievement tests. The German sample is made up of 50 classes from grades 8 to 10 from 38 schools throughout Germany. However, for this study, one class from the overall sample was excluded, as data for motivational outcomes and perceived teacher support was missing for all students in this class. The sample included in this study therefore consisted of 49 classes with a total of 1,116 students (48.5% female, n = 554). The mean age of the students was 15.0 years (SD = 0.80), with 6 classes from grade 8 (12.25%), 37 classes from grade 9 (75.5%), and 6 classes from grade 10 (12.25%). The portion of students with a migration background (defined as mother, father, or student not born in Germany) was 15.8% (n = 165). From the 49 teachers, 22 (44.9%) were female.

Students were taught by their usual mathematics teachers. In order to ensure that student outcomes could be accurately compared across classes, all students received the same two lessons on quadratic equations. Quadratic equations are an integral concept in algebra, which is one of the main components of mathematics education across numerous countries (National Council of Teachers of Mathematics

(NCTM), 2000; OECD, 2019) and therefore an appropriate topic for comparison.

Data was collected over a timespan of about 8 weeks. Prior to starting with the focus unit on quadratic equations, the pre-test questionnaire was administered to both students and teachers. Two separate lessons were then randomly selected to be recorded throughout the unit, one from the first half of the unit and one from the second half. Once the unit was finished, students and teachers completed a post-test questionnaire.

Measures

All measures used in this study, except for student gender, were taken from the student post-test questionnaire. Student gender was obtained from the pre-test questionnaire. For the post-test measures, students were instructed to answer all items in relation to the focal lessons on quadratic equations ("While answering the following questions, please always think about your learning during the unit on the topic of quadratic equations"). All scales were developed or adapted for the TALIS study by an international project team (Mihaly et al., 2021).

Self-concept

Student self-concept in quadratic equations was assessed via six items on a four-point Likert Scale (1 = *Strongly disagree* to 4 = *Strongly agree*). Items were adapted from the self-concept scale used in PISA (Mihaly et al., 2021). Sample items included "Learning about quadratic equations was easy for me" and "When I was taught the topic of quadratic equations, I could understand the concepts very well". The scale showed good reliability (ɑ = .88).

Self-efficacy

Student self-efficacy in quadratic equations was assessed using five items based on the Self-Efficacy for Learning and Performance component of the Motivated Strategies for Learning Questionnaire (Mihaly et al., 2021). Items were answered on a four-point Likert scale (1 = *Not at all true of me* to 4 = *Extremely true of me*). Items included phrases such as "I expected to do well in quadratic equations" or "I believed I would receive an excellent grade for the topic of quadratic equations". The reliability of the scale was good (ɑ = .89).

Interest

Student personal interest in quadratic equations was assessed via three items on a four-point Likert scale (1 = *Strongly disagree* to 4 = *Strongly agree*; see Mihaly et al., 2021). Items asked students to answer based on their current math lessons they had just participated in (e.g., "I was interested in the topic of quadratic equations" or "After my mathematics class on the topic of quadratic equations I was often already curious about the next mathematics class"). The reliability of the scale was good (ɑ = .82).

Student perceptions of constructive support

The items for both facets of constructive support were answered on a four-point Likert scale (1 = *Strongly disagree* to 4 = *Strongly agree*). The facet of instructional support was assessed via three items (Mihaly et al., 2021), for example, "Our mathematics teacher helped us with our learning" or "Our mathematics teacher continued teaching until we understood". Reliability was good, with α = .86. The facet of social-emotional support was assessed via five items (Mihaly et al., 2021), for example, "I got along well with my mathematics teacher" or "My mathematics teacher really listened to what I had to say". This scale also showed good reliability, with α = .90.

Student gender
Student gender was collected in the pre-test questionnaire. Students were asked to indicate whether they were male or female. This variable was then dichotomously coded in the data set, with males as 0 and females as 1.

Analyses and missing data
All inferential analyses were conducted in Mplus Version 8.3 (Muthén & Muthén, 1998–2017). We first conducted confirmatory factor analyses (CFA) to evaluate a measurement model in order to assess the fit of the observed items to the latent variables of self-concept, self-efficacy, interest, student perception of instructional support, and student perception of social-emotional support. A CFA with all latent variables included was used as the final measurement model. Goodness-of-fit was assessed using the following fit indices: Comparative Fit Index (CFI), Tucker-Lewis index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Residual (SRMR). We considered (a) CFI and TLI > .95 and .90, (b) RMSEA < .06 and .08, and (c) SRMR < .08 as indicators of excellent and adequate model fit, respectively (Hu & Bentler, 1999).

Before moving on to the structural equation modelling, we also tested for measurement invariance across gender as a prerequisite for investigating mean differences. Specifically, we assessed configural, metric, and scalar invariance for two latent variable models: one with the three motivational outcomes of self-concept, self-efficacy, and interest, as well as one model with the two perceived constructive support facets of instructional support and social-emotional support. We used the cutoff values recommended by Chen (2007) when evaluating the measurement invariance and considered a reduction of ΔCFI ≥ .010 and ΔRMSEA ≥ .015 or ΔSRMR ≥ .010 as indicative of non-invariance.

In order to evaluate RQ1, we used a structural equation model (SEM) with gender as a predictor of student self-concept, self-efficacy, and interest. Gender was always included with males as the reference category (intercept) in order to assess the effect of "being female" on the outcome variables. We included the latent variables of student self-concept, self-efficacy, and interest all together in one model. To investigate RQ2, we simultaneously added both facets of perceived constructive support (i.e., instructional and social-emotional) to the model as additional predictors of self-concept, self-efficacy, and interest. We used the same fit indices cutoffs described above to evaluate the model fit of the SEMs.

For RQ3, we used a latent moderated structural equation modelling approach (Klein & Moosbrugger, 2000). We followed the approach for latent moderated SEMs recommended by Maslowsky and colleagues (2015), which recommends comparing models with and without interaction terms in order to test for moderation. Therefore, we used the model with gender and both facets of perceived constructive support as predictors as the model without interaction effects (Model 0). We then added two latent interaction terms between gender and each facet of perceived constructive support to create a model with interaction effects (Model 1)[1]. Traditional model fit indices are not applicable to latent moderated SEMs. We therefore used the log-likelihood ratio test as recommend by Maslowsky and colleagues (2015) to compare the fit of Model 1 relative to Model 0. The likelihood ratio test statistic (LRTS) was calculated via the following equation:

$$\text{LRTS} = -2[(\text{log-likelihood for Model 0}) - (\text{log-likelihood for Model 1})]$$

The LRTS can be evaluated using the chi-square distribution. The degrees of freedom were calculated as the difference between the number of free parameters in the model with interaction and the number of free parameters in the model without interaction. A significant result from the log-likelihood ratio test indicates that Model 0 constitutes a significant loss in fit as compared to Model 1 (the models with interaction), and therefore suggests that Model 1 is a better fit to the data. Additionally, we compared the Bayesian information criterions (BIC) of the two models. A smaller BIC value is suggestive of a better fit to the data (Lin et al., 2017).

In order to account for the nested structure of the data (i.e., students nested within classes), all analyses were conducted using a robust sandwich estimator (TYPE = COMPLEX in Mplus) to adjust standard errors of parameter estimates and correct for the non-independence of observations (Asparouhov, 2005; Muthén & Satorra, 1995). To handle the non-normality of the data, maximum likelihood estimation with robust standard errors (MLR) was used for all models. For the post-test questionnaire, 91 students (8.2%) did not participate, resulting in missing data. We handled this missing data using the full information maximum likelihood approach (Schafer & Graham, 2002). Significance was evaluated using a *p*-value cutoff of < .05. As Mplus only provides two-tailed p-values, we recalculated the p-values for the directional hypotheses to be one-tailed. The outcome variables of self-concept, self-efficacy, and interest were allowed to correlate in all models. All relevant syntaxes and corresponding materials are public and can be found on the OSF webpage for this study via the following link:
https://osf.io/q9bej/?view_only=499c86b19d964d018ed8520bee753430

RESULTS

Descriptive statistics

Descriptive statistics and intercorrelations among latent and observed (i.e., gender) variables are reported for the total sample, males, and females in Table 1. Overall, the correlation patterns suggest significant relationships between all variables of interest, except for gender and student perceptions of constructive support.

Table 1
Descriptive statistics and latent correlations for all variables

Measure	2	3	4	5	6	*M*	*SD*
Total sample							
1. Self-concept	.80*	.60*	.41*	.34*	-.09*	2.59	0.65
2. Self-efficacy	-	.51*	.31*	.27*	-.16*	2.35	0.72
3. Interest		-	.42*	.43*	.08*	2.17	0.73
4. Instructional support			-	.81*	< .01	2.87	0.77
5. Social-emotional support				-	< .01	2.99	0.74
6. Gender					-	-	-
Measure	2	3	4	5		*M*	*SD*
Males							
1. Self-concept	.83*	.60*	.37*	.33*		2.64	0.67
2. Self-efficacy	-	.52*	.28*	.24		2.46	0.72
3. Interest		-	.40*	.43*		2.12	0.74
4. Instructional support			-	.81*		2.88	0.80
5. Social-emotional support				-		2.99	0.79
Measure	2	3	4	5		*M*	*SD*
Females							
1. Self-concept	.78*	.63*	.38*	.31*		2.54	0.62
2. Self-efficacy	-	.57*	.29*	.25*		2.25	0.70
3. Interest		-	.41*	.42*		2.21	0.70
4. Instructional support			-	.79*		2.86	0.74
5. Social-emotional support				-		2.99	0.70

Note. Gender: 0 = male, 1 = female. Range of values for all scales was from 1–4. *p < .05.

Confirmatory factor analysis and measurement invariance

The results of the CFA for the final measurement model with all latent variables (self-efficacy, self-concept, interest, perceived instructional support and perceived social-emotional support) revealed excellent model fit for all scales: CFI = .97, TLI = .97, RMSEA = .04 [.037, .045], and SRMR = .04. All items loaded strongly and significantly onto the respective latent factors.

As can be seen in Table 2, the results of the measurement invariance confirmed configural, metric, and scalar invariance across gender based on the cutoff values recommended by Chen (2007) for both the three motivational outcome variables grouped together, as well as the two facets of perceived constructive support grouped together. In other words, it can be assumed that both male and female students interpreted the perceptions of constructive support and motivational outcome measures in the same manner.

Table 2
Measurement invariance across gender

	df	χ^2	CFI	ΔCFI	RMSEA	ΔRMSEA	SRMR
Motivational outcomes							
Configural	90	339.863	.971		.052		.039
Metric	79	349.121	.972	.001	.050	.002	.040
Scalar	68	352.005	.973	.001	.047	.003	.040
Constructive support							
Configural	50	95.454	.990		.056		.021
Metric	44	102.184	.990	< .001	.053	.003	.024
Scalar	38	105.455	.990	< .001	.048	.005	.024

Main effects of gender on motivational outcomes
For RQ1, we fit a SEM with gender as a predictor of student self-concept, self-efficacy, and interest in quadratic equations. The model showed a good fit to the data, with CFI = .97, TLI = .96, RMSEA = .05 [.045, .058], and SRMR = .04. Model coefficients showed that gender (in this case, being female) significantly and negatively predicted self-concept (b = -.15, p = .02) and self-efficacy (b = -.33, p < .001) in quadratic equations. There were no significant effects of gender on interest in quadratic equations.

Main effects of teacher constructive support on motivational outcomes
To evaluate RQ2, we added both facets of perceived constructive support (instructional and social-emotional support) to the model as predictors of self-concept, self-efficacy, and interest in quadratic equations (Model 0). The results indicated good model fit (CFI = .97, TLI = .96, RMSEA = .04 [.036, .049], and SRMR = .04). Perceived instructional support had a significant positive effect on self-concept (b = .36, p < .001), self-efficacy (b = .24, p = .01) and interest (b = .20, p = .03) in quadratic equations. Perceived social-emotional support had a significant positive effect on interest in quadratic equations (b = .31, p < .001), but not on self-concept or self-efficacy in quadratic equations.

Interaction effects of gender and perceived constructive support
To answer RQ3, we added latent interactions (gender x perceived instructional support and gender x perceived social-emotional support) to Model 0, resulting in Model 1. When comparing the Model 1 to Model 0, the log-likelihood test was not significant (df = 6, χ^2 = 0.65), indicating that the model with interaction did not represent a better fit to the data. There was also no reduction in BIC in the interaction model as compared to the non-interaction model, additionally indicating that the interaction model was not a better fit to the data. The specific estimates for the interaction within the model were non-significant. Path coefficients for the model without interaction (Model 0) and the latent interaction model (Model 1) can be seen in Table 3.

Table 3
Path coefficients of the main effect model (Model 0) and the latent interaction model (Model 1) for gender and perceived instructional support

	Model 0			Model 1		
Outcome and predictor	Est. (*SE*)	Std. est.	*p*	Est. (*SE*)	Std. est.	*p*
Self-concept						
1. Gender	-0.15 (0.07)	-0.14	.016	-0.15 (0.07)	-0.14	.016
2. Instructional support	0.36 (0.10)	0.34	<.001	0.36 (0.16)	0.34	.011
3. Social-emotional support	0.05 (0.08)	0.05	.271	0.05 (0.15)	0.05	.361
4. Gender × Instructional support				0.01 (0.20)	0.01	.492
5. Gender × Social-emotional support				-0.01 (0.20)	-0.01	.491
R^2	.15			.15		
Self-efficacy						
1. Gender	-0.33 (0.07)	-0.31	<.001	-0.33 (0.07)	-0.31	<.001
2. Instructional support	0.24 (0.10)	0.23	<.001	0.25 (0.15)	0.24	.045
3. Social-emotional support	0.06 (0.11)	0.06	.287	0.04 (0.15)	0.04	.388
4. Gender × Instructional support				-0.01 (0.18)	-0.01	.476
5. Gender × Social-emotional support				0.04 (0.17)	0.04	.405
R^2	.10			.10		
Interest						
1. Gender	0.16 (0.07)	0.15	.992	0.17 (0.07)	0.15	.992
2. Instructional support	0.20 (0.10)	0.18	.028	0.17 (0.13)	0.15	.101
3. Social-emotional support	0.31 (0.11)	0.28	<.001	0.33 (0.13)	0.29	.008
4. Gender × Instructional support				0.06 (0.16)	0.05	.350
5. Gender × Social-emotional support				-0.02 (0.15)	-0.02	.439
R^2	.19			.20		
Goodness-of-fit						
AIC	46698.235			46709.586		
BIC	47082.475			47123.009		

Note. Est. = unstandardized parameter estimate; Std. est. = standardized estimate; AIC = Akaike information criterion; BIC = Bayesian information criterion.

DISCUSSION

Despite more attention in recent years to gender equality in education, female students continue to display lower levels of motivational outcomes than males in STEM subjects, including mathematics. The question of how to reduce these gender differences is not new in educational research and has been a topic of discourse for many years. Numerous interventions have been conducted that attempt to promote motivational outcomes in gender stereotypical subjects and therefore reduce differences between male and female students (for a meta-analysis, see Lesperance et al., 2022). While these interventions are a promising avenue for reducing gender differences, they do not refer to daily classroom learning and instruction. Although it has been shown that constructive support is positively connected to various motivational outcomes (e.g., summarized by Cornelius-White, 2007), there are still very few studies examining whether constructive support can especially bolster these outcomes for female students in mathematics. The present research therefore aimed to evaluate the main effects of student gender and perceived constructive support on the student motivational outcomes of self-concept, self-efficacy, and interest in quadratic equations in secondary school mathematics classrooms, as well as to investigate whether perceived constructive support moderated the relationship between gender and self-concept, self-efficacy, and interest.

Although our analyses focused on the specific mathematical topic of quadratic equations, a thorough understanding of quadratic equation concepts is also necessary for more advanced mathematical topics such as geometry and calculus (López et al., 2016) and can therefore be considered as a core concept of mathematics educations. Although there has been no systematic analysis of the specific differences in quadratic equations to date, we would not assume that there are particularly extreme differences in the area of quadratic equations from other areas of mathematics, making it appropriate to relate to the literature on general mathematics education.

The impact of gender on student motivational outcomes

Our hypothesis regarding the effect of gender on self-concept, self-efficacy, and interest in secondary school mathematics lessons on quadratic equations (H1) was partially supported. Consistent with previous findings on gender differences in math self-concept (Marsh & Yeung, 1998; Nagy et al., 2010; Watt, 2004) and self-efficacy (Huang, 2013; Pajares, 2005; Zander et al., 2020), female students reported lower levels of these outcomes than male students. However, we did not find any significant effect of gender on interest in quadratic equations. These results are in contrast to other studies that have found girls to report lower levels of interest in mathematics, especially in secondary school (Frenzel et al., 2010; Köller et al., 2001; Watt, 2004). However, there is also some evidence to suggest that the gender gap in math interest may not be as prominent as the gap in competency-related outcomes. For example, Ganley and Lubienski (2016) found that the differences between male and female students' math interest were substantially smaller than differences in confidence in their math abilities (although differences in interest were present). Furthermore, some studies have found no gender differences in math interest, even while still finding gender differences in math competency beliefs (Fredricks et al., 2018; Jacobs et al., 2002; Simpkins et al., 2006). Combined with the results of the current study, this suggests that while gender differences have been observed in both math competency-related outcomes (e.g., self-concept, self-efficacy) and math interest, the gender differences in competency-related outcomes seem to be stronger and more consistent. Therefore, it is possible that the persistent gender gap in mathematics-related areas is less a result of females not being interested in math, but rather more due to beliefs concerning their own capabilities in these areas.

Drawing on prior research, these lower levels of competency-beliefs in quadratic equations found in female students in our study may be due to learned gender stereotypes that students have incorporated into their own identities. Indeed, studies have shown that students who endorse traditional gender stereotypes in STEM-related subjects tend to display more gender stereotypical self-beliefs about their competencies in those subjects (e.g., female students do not perceive themselves as competent compared to male students) (Casad et al., 2015; Correll, 2001; Koul et al., 2016). Stereotype threat could provide one explanation for this mechanism. When female students are cognizant of the stereotype that females are not seen as competent or capable in STEM subjects such as math or science, this awareness can affect their self-assessment of their own abilities (Inzlicht & Schmader, 2012; Pennington et al., 2016; Shapiro & Williams, 2012).

The positive impact of perceived constructive support on motivational outcomes

Our findings partially supported our second hypothesis (H2). We found that while both perceived instructional support and perceived social-emotional support had a significant, positive effect on student interest in quadratic equations, only perceived instructional support had a significant, positive effect on self-concept and self-efficacy in quadratic equations. These results provide valuable insights into the unique effects of these two facets of perceived constructive support on specific student motivational outcomes.

The positive effect of both facets of perceived constructive support on student interest in quadratic equations corroborates prior research that has demonstrated the significance of a supportive learning environment for student interest. The significant effects found in our study suggest that when students perceive that their teachers provide clear feedback, helpful resources, and guidance, as well as foster a sense of belonging, empathy, and emotional well-being, they are more likely to develop a genuine interest in a given subject (Lazarides et al., 2019; Lazarides & Ittel, 2013; Prewett et al., 2019).

However, our results showed an interesting distinction in regards to self-concept and self-efficacy in quadratic equations. We found that for these two motivational outcomes, only the facet of perceived instructional support had a significant and positive effect. This implies that while perceived instructional support plays a crucial role in shaping students’ perception of their own competence and efficacy in quadratic equations, perceived social-emotional support may not. When placing these results in the context of prior research, it is important to acknowledge that many studies on the relationship between perceived constructive support and student motivational outcomes do not distinguish teacher support into two facets as we have done in our study, but rather use an overall measure, with aspects of both facets creating a single factor or just one facet as a measure. Studies combing both facets into a single factor also find significant effects on competency-related beliefs (Ahmed et al., 2008; Lapointe et al., 2005; Yu & Singh, 2018). However, it is difficult to determine what aspects of perceived constructive support are driving these effects when they are combined in one model. A sparse number of studies looking at the sole effect of perceived instructional support also found positive effects for self-efficacy and self-concept (Liu et al., 2018; Ma et al., 2021; Yildirim, 2012). While some of the few studies investigating solely perceived social-emotional support have found that it is not related to student competency outcomes (Ruzek et al., 2016), there is sparse evidence for relations between perceived social-emotional support and self-efficacy (Yang et al., 2021).

The differential effects of perceived instructional support and perceived social-emotional support on student self-concept, self-efficacy, and interest in quadratic equations underscore the multifaceted nature of perceived constructive support in mathematics classrooms. Of the three outcomes evaluated, interest has been shown to be the most affective in nature, is very closely tied to positive emotions experienced while performing a given task, and relies heavily on interactions and experience (Frenzel et al., 2010). Therefore, it may be possible that the personal and emotional implications of perceived social-emotional support are especially important for student interest. In contrast, it seems as though it is instructional support that plays the more critical role in fostering students' confidence and beliefs about their abilities in quadratic equations. This is in line with the nature of self-concept and self-efficacy, which both involve more cognitive appraisals of ones' capabilities and are therefore heavily tied to learning and instruction (Bong & Skaalvik, 2003). Additionally, these results imply that even when students perceive their teachers as warm, caring, and personally interested in their emotional well-being, this has little effect on their evaluations of their own competency expectations. This underscores the nuanced nature of the associations between expectancy-related outcomes and contextual factors such as constructive support compared to the associations observed between value-related outcomes and these same contextual factors.

The interaction of perceived constructive support and gender

Our findings did not support our hypothesis regarding the interaction of gender and perceived constructive support (H3). There were no significant effects of the interaction between gender and perceived constructive support for any of the student motivational outcomes, indicating that while perceived constructive support has positive effects for students with regard to motivational outcomes, our study did not support the hypothesis that these effects might be stronger for female students in mathematics than for boys. One possible explanation for the lack of interaction effects could be due to the design of the data. The data used was cross-sectional and only assessed student motivation in the classroom of interest at one time point. Although the instruments used attempted to measure the motivational constructs as specifically as possible in a state context (e.g., "I was interested in the topic of quadratic equations", "I expected to do well in quadratic equations"), it is possible that detecting interaction effects in this design is not feasible. Motivational outcomes are continuously shaped over the entire course of a student's life and therefore may need to be investigated over the long-term.

Another possible explanation for the lack of interaction effects may be that constructive support is equally beneficial for male and female students. Through the framework of the academic risk perspective, it is plausible that females would especially benefit from high levels of perceived constructive support due to the risk of them having lower levels of motivational outcomes in mathematics. However, our results suggest that the effect of perceived constructive support on student motivational outcomes does not differ in regards to gender. While it is positive to confirm that perceived constructive support has beneficial effects for all students, this might signify that even with highly supportive teachers, female students still display lower levels of motivational outcomes in mathematics, which may be due to factors outside of the classroom, and therefore may require more targeted approaches to combat these deficits.

Limitations and future research

Although this study takes a crucial first look at the role of perceived constructive support for gender differences in student motivational outcomes, there are a few limitations that should be considered when interpreting the results. First, the variables used were measured as states. In order to investigate the long-term effects of perceived constructive support on student motivational outcomes, it may be crucial to examine these relationships as traits using longitudinal data. While the current work took a first step in evaluating if any interactions exist between these variables, assessing them over longer periods of time would allow for a more detailed view, as well as the possibility to evaluate how perceived constructive support is related to changes in motivational outcomes over time. Students already start to display gender differences in motivational outcomes in primary school (Eccles et al., 1993). Therefore, these differences tend to already exist by the time they enter secondary school and may require a longer period of time to evaluate what factors influence their development. Additionally, the current data did not allow us to take reciprocal effects into account, however, results from other studies have suggested that some motivational outcomes, for example, interest, may also have an effect on how students perceive their classroom environment over longer periods of time (Lazarides & Ittel, 2012). Future research should therefore consider using longitudinal designs when examining these relationships.

Secondly, all scales included in the study were self-reported by the students. Self-report data is susceptible to common method bias (Podsakoff et al., 2012). It is also possible that social desirability influenced the self-reported information. We chose to only use student data in this study because we were interested in the individual experiences of each student and how that related to their motivational outcomes. However, future studies could consider drawing information about constructive support from additional sources such as teacher reports or third-party observations to reduce the possibility of biases from purely self-reported data.

It is also important to mention that the secondary school system in Germany consists of different school types, each with a distinct curriculum. The students in the current data set were mainly (82%) from the *Gymnasium* school type, which can be considered as the most academically rigorous type. The high percentage of students from this school type in the data did not allow us to assess any differences between school types, however, it is plausible that gender differences may present as more or less pronounced in different types of schools. Indeed, gender differences have been shown to vary in regards to ability level (Preckel et al., 2008). Additionally, schools with a stronger vocational focus tend to have a different study body composition in regards to socioeconomic status.

Socioeconomic status can also have an effect on gender differences in academic contexts (Cascella & Pampaka, 2020). Therefore, future research should also strive to include a more diverse sample of schools when continuing studies in the German secondary school system.

Moreover, future research could further explore additional contextual variables that have also been associated with variations in gender differences in motivational outcomes such as gifted versus average-ability students (Preckel et al., 2008; Rudasill et al., 2009; Zhou et al., 2017) and teacher gender (Duffy et al., 2001; Gong et al., 2018; Martin & Marsh, 2005). These factors have been previously associated with gender differences in motivational outcomes, however, due to the focus and scope of our present research questions, they were not explicitly examined. Expanding the investigation to include these variables would provide a deeper understanding of the relationships between these various factors.

Implications and conclusions

This study is one of the first to examine the potential of perceived constructive support as a moderator of gender differences in student motivational outcomes in secondary school mathematics. Concurrent with prior research, our results showed that female students display lower levels of both self-concept and self-efficacy in mathematics. However, contradictory to prior research, we did not find the same negative effect for female math interest. While more research is needed to examine the concrete gender differences in these outcomes, our results hint that when trying to encourage female students in mathematics, it might be more pertinent to focus on their competency beliefs. This is valuable as many educational initiatives that focus on females in STEM subjects tend to address a wide range of outcomes, and knowledge of which motivational constructs are most affected by gender is crucial going forward. These results also highlight the importance of developing a more differentiated view of gender differences in motivational outcomes in mathematics.

This study also illustrates that perceived constructive support continues to be an important predictor of student motivational outcomes in mathematics, and that there are nuanced relationships between specific facets of perceived constructive support and specific motivational outcomes. These findings emphasize the need for educational practitioners to support students in ways that not only promote interest but also enhance students' self-perceptions and feelings of competence in their mathematical abilities. Perceived constructive support should continue to be studied by researchers in more diverse contexts and populations to deeply understand the magnitude and variation of these effects. Additionally, these results provide evidence for teacher training and continuing teacher education that training teachers to be supportive, caring actors in students' lives can make an important impact on student motivation, and subsequently, future educational outcomes.

While we did not find the hypothesized interaction effects between gender and perceived constructive support, this line of research is, to our knowledge, one of the first studies to approach this topic. Future studies should continue to investigate the interplay of these variables and their effects in various contexts and with different populations. These results provide a starting point for that research. Additionally, this study highlights the relative scarcity of studies that investigate how various aspects of teaching quality and classroom environment may influence gender differences in mathematics. Continued research in this area is crucial for understanding what educators can do to combat gender differences in mathematics and create an environment where all students can reach their full potential, regardless of gender.

AUTHOR NOTE

Kaley Lesperance
https://orcid.org/0000-0002-0669-2744

Jasmin Decristan
https://orcid.org/0000-0003-4664-1367

Doris Holzberger
https://orcid.org/0000-0001-6535-289X

We have no conflicts of interest to disclose. Please direct all correspondence to Kaley Lesperance: kaley.lesperance@tum.de.

ENDNOTES

[1] Due to the high correlation between perceived instructional support and perceived social-emotional support, we also ran two separate latent moderation SEMs in order to ensure that the lack of interaction effects was not due to the high correlation between these two facets. However, we also did not find any interaction effects for the separate models with just perceived instructional support or just perceived social-emotional support.

REFERENCES

Ahmed, W., Minnaert, A., van der Werf, G., & Kuyper, H. (2008). Perceived social support and early adolescents' achievement: The mediational roles of motivational beliefs and emotions. *Journal of Youth and Adolescence*, *39*(1), 36–46. https://doi.org/10.1007/s10964-008-9367-7

Allen, M., Witt, P. L., & Wheeless, L. R. (2006). The role of teacher immediacy as a motivational factor in student learning: Using meta-analysis to test a causal model. *Communication Education*, *55*(1), 21–31. https://doi.org/10.1080/03634520500343368

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *12*(3), 411–434. https://doi.org/10.1207/s15328007sem1203_4

Barth, J., & Masters, S. (2020). Effects of classroom quality, gender stereotypes, and efficacy on math and science interest over school transitions. *International Journal of Gender, Science and Technology*, *12*(1), 4–31. https://genderandset.open.ac.uk/index.php/genderandset/article/view/658

Bharadwaj, P., De Giorgi, G., Hansen, D., & Neilson, C. A. (2016). The gender gap in mathematics: Evidence from Chile. *Economic Development and Cultural Change*, *65*(1), 141–166. https://doi.org/10.1086/687983

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, *35*(1), 61–79. https://doi.org/10.1016/S0022-4405(96)00029-5

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist*, *34*(3), 139–153. https://doi.org/10.1207/s15326985ep3403_1

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40. https://doi.org/10.1023/A:1021302408382

Brunner, M. (2008). No g in education? *Learning and Individual Differences*, *18*(2), 152–165. https://doi.org/10.1016/j.lindif.2007.08.005

Burić, I., & Kim, L. E. (2020). Teacher self-efficacy, instructional quality, and student motivational beliefs: An analysis using multilevel structural equation modeling. *Learning and Instruction*, *66*, Article 101302. https://doi.org/10.1016/j.learninstruc.2019.101302

Casad, B. J., Hale, P., & Wachs, F. L. (2015). Parent-child math anxiety and math-gender stereotypes predict adolescents' math education outcomes. *Frontiers in Psychology*, *6*, Article 1597. https://doi.org/10.3389/fpsyg.2015.01597

Cascella, C., & Pampaka, M. (2020). Attitudes towards gender roles in family: A Rasch-based validation study. *Journal of Applied Measurement*, *21*(2), 1-24. https://research.manchester.ac.uk/en/publications/attitudes-towards-gender-roles-in-family-a-rasch-based-validation

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Contini, D., Tommaso, M. L. D., & Mendolia, S. (2017). The gender gap in mathematics achievement: Evidence from Italian data. *Economics of Education Review*, *58*, 32–42. https://doi.org/10.1016/j.econedurev.2017.03.001

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research*, *77*(1), 113–143. https://doi.org/10.3102/003465430298563

Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal of Sociology*, *106*(6), 1691–1730. https://doi.org/10.1086/321299

Curby, T. W., LoCasale-Crouch, J., Konold, T. R., Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2009). The relations of observed pre-K classroom quality profiles to children's achievement and social competence. *Early Education and Development*, *20*(2), 346–372. https://doi.org/10.1080/10409280802581284

Dasgupta, N., & Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights from the Behavioral and Brain Sciences*, *1*(1), 21–29. https://doi.org/10.1177/2372732214549471

Decristan, J., Kunter, M., & Fauth, B. (2022). Die Bedeutung individueller Merkmale und konstruktiver Unterstützung der Lehrkraft für die soziale Integration von Schülerinnen und Schülern im Mathematikunterricht der Sekundarstufe [The importance of individual characteristics and constructive support for the social integration of students in secondary mathematics classes]. *Zeitschrift für Pädagogische Psychologie*, *36*(1–2), 85–100. https://doi.org/10.1024/1010-0652/a000329

Decristan, J., Kunter, M., Fauth, B., Büttner, G., Hardy, I., & Hertel, S. (2016). What role does instructional quality play for elementary school children's science competence? A focus on students at risk. *Journal for Educational Research Online*, *8*(1), 66–89. https://doi.org/10.25656/01:12032

Demaray, M. K., Malecki, C. K., Rueger, S. Y., Brown, S. E., & Summers, K. H. (2009). The role of youth's ratings of the importance of socially supportive behaviors in the relationship between social support and self-concept. *Journal of Youth and Adolescence*, *38*(1), 13–28. https://doi.org/10.1007/s10964-007-9258-3

Dickhäuser, O., & Meyer, W.-U. (2006). Gender differences in young children's math ability attributions. *Psychology Science*, *48*(1), 3–16.

Duffy, J., Warren, K., & Walsh, M. (2001). Classroom interactions: Gender of teacher, gender of student, and classroom subject. *Sex Roles: A Journal of Research*, *45*(9), 579–593. https://doi.org/10.1023/A:1014892408105

Eccles, J. S. (2005). Subjective task value and the Eccles et al. model of achievement-related choices. In A. J. Elliott & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105–121). Guilford.

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, Article 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*(3), 830–847. https://doi.org/10.2307/1131221

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127. https://doi.org/10.1037/a0018053

Fast, L. A., Lewis, J. L., Bryant, M. J., Bocian, K. A., Cardullo, R. A., Rettig, M., & Hammond, K. A. (2010). Does math self-efficacy mediate the effect of the perceived classroom environment on standardized math test performance? *Journal of Educational Psychology*, *102*(3), 729–740. https://doi.org/10.1037/a0018863

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, *29*, 1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001

Fogliati, V. J., & Bussey, K. (2013). Stereotype threat reduces motivation to improve: Effects of stereotype threat and feedback on women's intentions to improve mathematical ability. *Psychology of Women Quarterly*, *37*(3), 310–324. https://doi.org/10.1177/0361684313480045

Fredricks, J. A., Hofkens, T., Wang, M.-T., Mortenson, E., & Scott, P. (2018). Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *Journal of Research in Science Teaching*, *55*(2), 271–298. https://doi.org/10.1002/tea.21419

Frenzel, A. C., Goetz, T., Pekrun, R., & Watt, H. M. G. (2010). Development of mathematics interest in adolescence: Influences of gender, family, and school context. *Journal of Research on Adolescence*, *20*(2), 507–537. https://doi.org/10.1111/j.1532-7795.2010.00645.x

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, *95*(1), 148–162. https://doi.org/10.1037/0022-0663.95.1.148

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences*, *47*, 182–193. https://doi.org/10.1016/j.lindif.2016.01.002

Gaspard, H., & Lauermann, F. (2021). Emotionally and motivationally supportive classrooms: A state-trait analysis of lesson- and classroom-specific variation in teacher- and student-reported teacher enthusiasm and student engagement. *Learning and Instruction*, *75*, Article 101494. https://doi.org/10.1016/j.learninstruc.2021.101494

Goldman, A. D., & Penner, A. M. (2016). Exploring international gender differences in mathematics self-concept. *International Journal of Adolescence and Youth*, *21*(4), 403–418. https://doi.org/10.1080/02673843.2013.847850

Göllner, R., Fauth, B., & Wagner, W. (2021). Student ratings of teaching quality dimensions: Empirical findings and future directions. In W. Rollett, H. Bijlsma, & S. Röhl (Eds.), *Student feedback in schools – Using perceptions for the development of teaching and teachers* (pp. 111–122). Springer.

Gong, J., Lu, Y., & Song, H. (2018). The effect of teacher gender on students' academic and noncognitive outcomes. *Journal of Labor Economics*, *36*(3), 743–778. https://doi.org/10.1086/696203

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700–717. https://doi.org/10.1037/a0026659

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, *76*(5), 949–967. https://doi.org/10.1111/j.1467-8624.2005.00889.x

Hausen, J. E., Möller, J., Greiff, S., & Niepel, C. (2022). Students' personality and state academic self-concept: Predicting differences in mean level and within-person variability in everyday school life. *Journal of Educational Psychology*, *114*(6), 1394–1411. https://doi.org/10.1037/edu0000760

Herbert, J., & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of Applied Developmental Psychology*, *26*(3), 276–295. https://doi.org/10.1016/j.appdev.2005.02.007

Hochweber, J., & Vieluf, S. (2018). Gender differences in reading achievement and enjoyment of reading: The role of perceived teaching quality. *The Journal of Educational Research*, *111*(3), 268–283. https://doi.org/10.1080/00220671.2016.1253536

Hornstra, L., van der Veen, I., & Peetsma, T. (2016). Domain-specificity of motivation: A longitudinal study in upper primary school. *Learning and Individual Differences*, *51*, 167–178. https://doi.org/10.1016/j.lindif.2016.08.012

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, *28*(1), 1–35. https://doi.org/10.1007/s10212-011-0097-y

Hughes, J. N., & Chen, Q. (2011). Reciprocal effects of student–teacher and student–peer relatedness: Effects on academic self efficacy. *Journal of Applied Developmental Psychology*, *32*(5), 278–287. https://doi.org/10.1016/j.appdev.2010.03.005

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494–495. https://doi.org/10.1126/science.1160364

Inzlicht, M., & Schmader, T. (2012). *Stereotype threat: Theory, process, and application*. Oxford University Press.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

Kieft, M., Rijlaarsdam, G., & van den Bergh, H. (2008). An aptitude–treatment interaction approach to writing-to-learn. *Learning and Instruction*, *18*(4), 379–390. https://doi.org/10.1016/j.learninstruc.2007.07.004

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*(4), 457–474. https://doi.org/10.1007/BF02296338

Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras study. Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in* *investigating teaching and learning in the classroom* (pp. 137–160). Waxmann.

Köller, O., Baumert, J., & Schnabel, K. (2001). Does interest matter? The relationship between academic interest and achievement in mathematics. *Journal for Research in Mathematics Education*, *32*(5), 448–470. https://doi.org/10.2307/749801

Kollmayer, M., Schober, B., & Spiel, C. (2018). Gender stereotypes in education: Development, consequences, and interventions. *European Journal of Developmental Psychology*, *15*(4), 361–377. https://doi.org/10.1080/17405629.2016.1193483

Koul, R., Lerdpornkulrat, T., & Poondej, C. (2016). Gender compatibility, math-gender stereotypes, and self-concepts in math and physics. *Physical Review Physics Education Research*, *12*(2), Article 020115. https://doi.org/10.1103/PhysRevPhysEducRes.12.020115

Lapointe, J. M., Legault, F., & Batiste, S. J. (2005). Teacher interpersonal behavior and adolescents’ motivation in mathematics: A comparison of learning disabled, average, and talented students. *International Journal of Educational Research*, *43*(1), 39–54. https://doi.org/10.1016/j.ijer.2006.03.005

Lazarides, R., Gaspard, H., & Dicke, A.-L. (2019). Dynamics of classroom motivation: Teacher enthusiasm and the development of math interest and teacher support. *Learning and Instruction*, *60*, 126–137. https://doi.org/10.1016/j.learninstruc.2018.01.012

Lazarides, R., & Ittel, A. (2012). Instructional quality and attitudes toward mathematics: Do self-concept and interest differ across students’ patterns of perceived instructional quality in mathematics classrooms? *Child Development Research*, *2012*, Article 813920. https://doi.org/10.1155/2012/813920

Lazarides, R., & Ittel, A. (2013). Mathematics interest and achievement: What role do perceived parent and teacher support play? A longitudinal analysis. *International Journal of Gender, Science and Technology*, *5*(3), 207-231. https://genderandset.open.ac.uk/index.php/genderandset/article/view/301

Lesperance, K., Hofer, S., Retelsdorf, J., & Holzberger, D. (2022). Reducing gender differences in student motivational-affective factors: A meta-analysis of school-based interventions. *British Journal of Educational Psychology*, *92*(4), 1502–1536. https://doi.org/10.1111/bjep.12512

Lin, L.-C., Huang, P.-H., & Weng, L.-J. (2017). Selecting path models in SEM: A comparison of model selection criteria. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(6), 855–869. https://doi.org/10.1080/10705511.2017.1363652

Liu, R.-D., Zhen, R., Ding, Y., Liu, Y., Wang, J., Jiang, R., & Xu, L. (2018). Teacher support and math engagement: Roles of academic self-efficacy and positive emotions. *Educational Psychology*, *38*(1), 3–16. https://doi.org/10.1080/01443410.2017.1359238

López, J., Robles, I., & Martínez-Planell, R. (2016). Students’ understanding of quadratic equations. *International Journal of Mathematical Education in Science and Technology*, *47*(4), 552–572. https://doi.org/10.1080/0020739X.2015.1119895

Ma, L., Luo, H., & Xiao, L. (2021). Perceived teacher support, self-concept, enjoyment and achievement in reading: A multilevel mediation model based on PISA 2018. *Learning and Individual Differences*, *85*, Article 101947. https://doi.org/10.1016/j.lindif.2020.101947

Makarova, E., Aeschlimann, B., & Herzog, W. (2019). The gender gap in STEM fields: The impact of the gender stereotype of math and science on secondary students' career aspirations. *Frontiers in Education*, *4*, Article 60. https://doi.org/10.3389/feduc.2019.00060

Malecki, C. K., & Demaray, M. K. (2006). Social support as a buffer in the relationship between socioeconomic status and academic performance. *School Psychology Quarterly*, *21*(4), 375–395. https://doi.org/10.1037/h0084129

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal*, *35*(4), 705–738. https://doi.org/10.3102/00028312035004705

Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research*, *79*(1), 327–365. https://doi.org/10.3102/0034654308325583

Martin, A., & Marsh, H. (2005). Motivating boys and motivating girls: Does teacher gender really make a difference? *Australian Journal of Education*, *49*(3), 320–334. https://doi.org/10.1177/000494410504900308

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, *52*(4), Article 1119. https://doi.org/10.2307/1129498

Maslowsky, J., Jager, J., & Hemken, D. (2015). Estimating and interpreting latent variable interactions: A tutorial for applying the latent moderated structural equations method. *International Journal of Behavioral Development*, *39*(1), 87–96. https://doi.org/10.1177/0165025414552301

McFarland, L., Murray, E., & Phillipson, S. (2016). Student–teacher relationships and student self-concept: Relations with teacher and student gender. *Australian Journal of Education*, *60*(1), 5–25. https://doi.org/10.1177/0004944115626426

Mihaly, K., Klieme, E., Fischer, J., & Doan, S. (2021). Questionnaire scale characteristics. In OECD (Ed.), *Global Teaching InSights technical report.* Section IV: Analysis (pp. 1–22). OECD Publishing. https://web-archive.oecd.org/2021-03-10/578338-GTI-TechReport-Chapter18.pdf

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90*(3), 376–419. https://doi.org/10.3102/0034654320919354

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, Article 267. https://doi.org/10.2307/271070

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nagy, G., Watt, H. M. G., Eccles, J. S., Trautwein, U., Lüdtke, O., & Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: Different countries, different trajectories? *Journal of Research on*

Adolescence, *20*(2), 482–506. https://doi.org/10.1111/j.1532-7795.2010.00644.x
National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. NCTM.
Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, *7*(2), 133–144. https://doi.org/10.1177/1477878509104318
OECD. (2013). *Mathematics self-beliefs and participation in mathematics-related activities. In PISA 2012 results: Ready to learn (volume III)*. OECD Publishing.
OECD. (2019). *PISA 2018 assessment and analytical framework*. OECD Publishing. https://www.oecd-ilibrary.org/education/pisa-2018-assessment-and-analytical-framework_b25efab8-en
OECD. (2020a). *Global Teaching InSights: A video study of teaching*. OECD Publishing. https://www.oecd-ilibrary.org/education/global-teaching-insights_20d6f36b-en
OECD. (2020b). *Global Teaching InSights: A video study of teaching*. [Dataset]. https://www.oecd.org/education/school/global-teaching-insights-technical-documents.htm
Pajares, F. (2005). Gender differences in mathematics self-efficacy beliefs. In A. Gallagher & J. Kaufman (Eds.), *Gender differences in mathematics: An integrative psychological approach* (pp. 294–315). Cambridge University Press.
Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, *34*(1), 29–48. https://doi.org/10.1080/01443410.2013.797339
Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, *99*(1), 83–98. https://doi.org/10.1037/0022-0663.99.1.83
Pennington, C. R., Heim, D., Levy, A. R., & Larkin, D. T. (2016). Twenty years of stereotype threat research: A review of psychological mediators. *PLoS ONE*, *11*(1), Article e0146487. https://doi.org/10.1371/journal.pone.0146487
Pintrich, P. R. (2003). Motivation and classroom learning. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (Vol. 7, pp. 103–122). John Wiley & Sons Inc.
Plenty, S., & Heubeck, B. G. (2013). A multidimensional analysis of changes in mathematics motivation and engagement during high school. *Educational Psychology*, *33*(1), 14–30. https://doi.org/10.1080/01443410.2012.740199
Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*(1), 539–569. https://doi.org/10.1146/annurev-psych-120710-100452
Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM*

Mathematics Education, *50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4

Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, *31*, 2–12. https://doi.org/10.1016/j.learninstruc.2013.12.002

Preckel, F., Goetz, T., Pekrun, R., & Kleine, M. (2008). Gender differences in gifted and average-ability students: Comparing girls' and boys' achievement, self-concept, interest, and motivation in mathematics. *Gifted Child Quarterly*, *52*(2), 146–159. https://doi.org/10.1177/0016986208315834

Prewett, S. L., Bergin, D. A., & Huang, F. L. (2019). Student and teacher perceptions on student-teacher relationship quality: A middle school perspective. *School Psychology International*, *40*(1), 66–87. https://doi.org/10.1177/0143034318807743

Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 149–172). Springer. https://doi.org/10.1007/978-1-4614-2018-7_7

Reeve, J., Deci, E. L., & Ryan, R. M. (2004). Self-determination theory: A dialectical framework for understanding the sociocultural influences on student motivation. In D. M. McInerney & S. van Etten (Eds.), *Research on sociocultural influences on motivation and learning: Big theories revisited* (Vol. 4, pp. 31–59). Information Age Press.

Rowley, S. J., Kurtz-Costes, B., Mistry, R., & Feagans, L. (2007). Social status as a predictor of race and gender stereotypes in late childhood and early adolescence. *Social Development*, *16*(1), 150–168. https://doi.org/10.1111/j.1467-9507.2007.00376.x

Rudasill, K. M., Capper, M. R., Foust, R. C., Callahan, C. M., & Albaugh, S. B. (2009). Grade and gender differences in gifted students' self-concepts. *Journal for the Education of the Gifted*, *32*(3), 340–367. https://doi.org/10.4219/jeg-2009-862

Rueger, S. Y., Malecki, C. K., & Demaray, M. K. (2008). Relationship between multiple sources of perceived social support and psychological and academic adjustment in early adolescence: Comparisons across gender. *Journal of Youth and Adolescence*, *39*(1), 47. https://doi.org/10.1007/s10964-008-9368-6

Ruzek, E. A., Hafen, C. A., Allen, J. P., Gregory, A., Mikami, A. Y., & Pianta, R. C. (2016). How teacher emotional support motivates students: The mediating roles of perceived peer relatedness, autonomy support, and competence. *Learning and Instruction*, *42*, 95–103. https://doi.org/10.1016/j.learninstruc.2016.01.004

Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal*, *38*(2), 437–460. https://doi.org/10.3102/00028312038002437

Ryan, R. M., & Powelson, C. L. (1991). Autonomy and relatedness as fundamental to motivation and education. *The Journal of Experimental Education*, *60*(1), 49–66. http://www.jstor.org/stable/20152311

Ryan, R. M., Stiller, J. D., & Lynch, J. H. (1994). Representations of relationships to teachers, parents, and friends as predictors of academic motivation and self-esteem. *The Journal of Early Adolescence*, *14*(2), 226–249. https://doi.org/10.1177/027243169401400207

Sabol, T. J., & Pianta, R. C. (2012). Recent trends in research on teacher–child relationships. *Attachment & Human Development*, *14*(3), 213–231. https://doi.org/10.1080/14616734.2012.672262

Sakiz, G., Pape, S. J., & Hoy, A. W. (2012). Does perceived teacher affective support matter for middle school students in mathematics classrooms? *Journal of School Psychology*, *50*(2), 235–255. https://doi.org/10.1016/j.jsp.2011.10.005

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Scherer, R., & Nilsen, T. (2016). The relations among school climate, instructional quality, and achievement motivation in mathematics. In T. Nilsen & J.-E. Gustafsson (Eds.), *Teacher quality, instructional quality and student outcomes: Relationships across countries, cohorts and time* (Vol. 2). Springer. https://doi.org/10.1007/978-3-319-41252-8

Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, *26*(3–4), 299–323. https://doi.org/10.1080/00461520.1991.9653136

Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. OECD Publishing. https://www.oecd.org/pisa/PISA%202018%20Insights%20and%20Interpretations%20FINAL%20PDF.pdf

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 16-32). Academic Press. https://doi.org/10.1016/B978-012750053-9/50003-6

Shapiro, J. R., & Williams, A. M. (2012). The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles: A Journal of Research*, *66*(3), 175–183. https://doi.org/10.1007/s11199-011-0051-0

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, *42*(1), 70. https://doi.org/10.1037/0012-1649.42.1.70

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, *85*(4), 571–581. https://doi.org/10.1037/0022-0663.85.4.571

Smith, E., & Farkas, G. (2022). Gender and mathematics achievement: The role of gender stereotypical beliefs of classroom peers. *European Sociological Review*, *39*(2), 161–176. https://doi.org/10.1093/esr/jcac043

Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology*, *43*(1), 583–626. https://doi.org/10.1146/annurev.ps.43.020192.003055

Soland, J., Kuhfeld, M., Wolk, E., & Bi, S. (2019). Examining the state-trait composition of social-emotional learning constructs: Implications for practice, policy, and evaluation. *Journal of Research on Educational Effectiveness*, *12*(3), 550–577. https://doi.org/10.1080/19345747.2019.1615158

Thoman, D. B., Smith, J. L., Brown, E. R., Chase, J., & Lee, J. Y. K. (2013). Beyond performance: A motivational experiences model of stereotype threat. *Educational Psychology Review*, *25*(2), 211–243. https://doi.org/10.1007/s10648-013-9219-1

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, *92*(1), 144–151. https://doi.org/10.1037/0022-0663.92.1.144

Vekiri, I. (2010). Boys' and girls' ICT beliefs: Do teachers matter? *Computers & Education*, *55*(1), 16–23. https://doi.org/10.1016/j.compedu.2009.11.013

Wang, M.-T. (2012). Educational and career interests in math: A longitudinal examination of the links between classroom environment, motivational beliefs, and interests. *Developmental Psychology*, *48*(6), 1643–1657. https://doi.org/10.1037/a0027247

Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy-value perspective to understand individual and gender differences in STEM fields. *Developmental Review* , *33*(4), 304–340. https://doi.org/10.1016/j.dr.2013.08.001

Wang, M.-T., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development*, *83*(3), 877–895. https://doi.org/10.1111/j.1467-8624.2012.01745.x

Wang, M.-T., Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice: Individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological Science*, *24*(5), 770–775. https://doi.org/10.1177/0956797612458937

Wasserman, T., & Wasserman, L. (2020). Motivation: State, trait, or both. In T. Wasserman & L. Wasserman (Eds.), *Motivation, effort, and the neural network model* (pp. 93–101). Springer. https://doi.org/10.1007/978-3-030-58724-6_8

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th- through 11th-grade Australian students. *Child Development*, *75*(5), 1556–1574. https://doi.org/10.1111/j.1467-8624.2004.00757.x

Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, *89*(3), 411–419. https://doi.org/10.1037/0022-0663.89.3.411

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, *6*(1), 49–78. https://doi.org/10.1007/BF02209024

Wigfield, A., Battle, A., Keller, L. B., & Eccles, J. S. (2002). Sex differences in motivation, self-concept, career aspiration, and career choice: Implications for cognitive development. In A. McGillicuddy-De Lisi & R. De Lisi (Eds.), *Biology, society, and behavior: The development of sex differences in cognition* (pp. 93–124). Ablex.

Wigfield, A., Cambria, J., & Eccles, J. S. (2012). Motivation in education. In R. C.

Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 463–478). Oxford University Press.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wigfield, A., & Harold, R. D. (1992). Teacher beliefs and children's achievement self-perceptions: A developmental perspective. In D. H. Schunk & J. L. Meece (Eds.), *Student perceptions in the classroom* (pp. 95–121). Erlbaum.

Yang, X., & Kaiser, G. (2022). The impact of mathematics teachers' professional competence on instructional quality and students' mathematics learning outcomes. *Current Opinion in Behavioral Sciences*, *48*, Article 101225. https://doi.org/10.1016/j.cobeha.2022.101225

Yang, Y., Li, G., Su, Z., & Yuan, Y. (2021). Teacher's emotional support and math performance: The chain mediating effect of academic self-efficacy and math behavioral engagement. *Frontiers in Psychology*, *12*, Article 651608. https://doi.org/10.3389/fpsyg.2021.651608

Yildirim, S. (2012). Teacher support, motivation, learning strategy use, and achievement: A multilevel mediation model. *The Journal of Experimental Education*, *80*(2), 150–172. https://www.jstor.org/stable/26594348

Yu, R., & Singh, K. (2018). Teacher support, instructional practices, student motivation, and mathematics achievement in high school. *The Journal of Educational Research*, *111*(1), 81–94. https://doi.org/10.1080/00220671.2016.1204260

Zander, L., Höhne, E., Harms, S., Pfost, M., & Hornsey, M. J. (2020). When grades are high but self-efficacy is low: Unpacking the confidence gap between girls and boys in mathematics. *Frontiers in Psychology*, *11*, Article 552355. https://doi.org/10.3389/fpsyg.2020.552355

Zhou, Y., Fan, X., Wei, X., & Tai, R. H. (2017). Gender gap among high achievers in math and implications for STEM pipeline. *The Asia-Pacific Education Researcher*, *26*(5), 259–269. https://doi.org/10.1007/s40299-017-0346-1