

Gender Differences in  
Mathematically Intensive  
STEM Fields:  
Factors of Influence and  
Multiple Perspectives

Dissertation  
zur Erlangung des Doktorgrades  
der Wirtschafts- und Sozialwissenschaftlichen Fakultät  
der Eberhard Karls Universität Tübingen

vorgelegt von  
Eike Wille, M.A.  
aus Pfullingen

Tübingen

2017

---

Tag der mündlichen Prüfung:

14.11.2017

Dekan:

Prof. Dr. rer. soc. Josef Schmid

1. Gutachter:

Prof. Dr. Ulrich Trautwein

2. Gutachterin:

Prof. Dr. Jessika Golle

---

# ACKNOWLEDGMENTS

I would like to express my great appreciations to my supervisors Prof. Dr. Ulrich Trautwein, Prof. Dr. Benjamin Nagengast, and Prof. Dr. Kerstin Oschatz for their support and guidance throughout this dissertation project. I benefitted a lot from their different expertise and greatly appreciate the opportunity to work with and learn from all of them.

Special thanks go also to Dr. Hanna Gaspard for her reliable support across large parts of this dissertation project. Your feedback was always very helpful and valuable for me.

I would also like to thank Prof. Dr. Jenna Cambria for her help and everything I learned from her, scientifically and personally. Furthermore, I would like to thank my further co-authors for their support and cooperation in the formation of the studies: Dr. Gundula Stoll, Prof. Dr. Katharina Scheiter, and Nicolas Hübner. Moreover, I express gratitude to Prof. Dr. Ulrich Trautwein and Prof. Dr. Jessika Golle for assessing this thesis.

My special thanks are extended to my former and current colleagues from the Hector Research Institute of Education Sciences and Psychology for the inspiring and supportive research environment. I enjoyed and valued the shared work as well as personal experiences with you over the last years. Particular thanks go to Dr. Norman Rose for his methodological advice and the members of the different workgroups I participated for their ideas and helpful feedback.

I owe also many thanks the (members of the) LEAD Graduate School & Research Network for the stimulating exchange, for offering me the opportunities of visiting (inter)national conferences and establishing contact to partner schools, and for funding parts of my research.

Last but not least, I would like to thank my beloved partner and my family. Thank you for all your support and for always being there for me!



---

# Abstract

Women are less likely than men to choose mathematically intensive STEM careers (Eurostat, 2017; National Science Foundation, 2017). Because there are usually high status career options in these fields, such gender differences are critical for gender equity in rights and opportunities (European Commission, 2013; Noonan, 2017). Eccles et al.'s (1983) expectancy-value theory provides a powerful framework for investigating the question of why fewer women than men enter these careers. According to Eccles et al. (1983, 2009), gender differences in mathematically intensive STEM careers can be linked to early emerging gender differences in expectancy and value beliefs in mathematics, which are themselves rooted in gendered socialization processes. There is ample empirical support for these assumptions, and a variety of individual and environmental factors have been found to be associated with women's lower participation in mathematically intensive STEM fields compared to men (see Wang & Degol, 2013; Wigfield et al., 2015). Nevertheless, important questions remain concerning the influence of critical individual and environmental factors. The present dissertation sought to address some of these questions: Using Eccles et al.'s (1983) expectancy-value theory as a guiding framework, the effects of two environmental factors that have the potential to reduce women's participation in mathematically intensive STEM careers were investigated, namely gender-math stereotypes in the media and high school coursework requirements. Furthermore, expectancy and value beliefs, the most proximal constructs associated with STEM careers in Eccles et al.'s (1983) expectancy-value theory, were compared to prominent predictors in the area of work and vocational psychology (i.e., vocational interests) in terms of their relative predictive power on gendered STEM careers.

To this end, three empirical studies were conducted as part of the present dissertation: Study 1 investigated the influence of gender stereotypes embedded in a children's television program on girls' and boys' stereotype endorsement, math motivation, and math performance. For this purpose, a randomized study was conducted with a total of 335 fifth graders. Watching a short clip of gender-math stereotypes embedded in a television program increased girls' and boys' stereotype endorsement. However, there was no effect on girls' math motivation, and only a small effect on boys' motivation. Furthermore, no effects on math performance were found for either girls or boys. Thus, there was only partial support for short-term effects of gender stereotypes embedded in a television program on girls' and boys' math motivation and performance.

Study 2 investigated whether encouraging young women to take advanced math courses in high school might bring more women into STEM careers. Young women are less

---

likely than men to choose advanced math courses in high school, which has been linked to gender differences in STEM careers (Ceci, Ginther, Kahn, & Williams, 2014; Watt & Eccles, 2008). Therefore, the effects of a statewide educational reform in Germany that required all students to take advanced math courses were investigated. To this end, data from 4,730 students who participated in high school courses before the reform were compared with data from 4,715 students who participated in high school courses after the reform. The reform was associated with different effects for young women and men: Gender differences favoring young men were smaller in math achievement, but larger in math self-concept and STEM-related vocational interests (i.e., realistic, investigative) in the cohort after the reform. Gender differences favoring young men in the choice of STEM majors at university did not differ between the two cohorts. Thus, it seems that encouraging young women to take advanced math courses in high school does not automatically increase gender equality in mathematically intensive STEM fields.

Study 3 examined the relative predictive validity of expectancy-value constructs and vocational interests for gender differences in mathematically intensive STEM careers. Both sets of constructs are highly predictive of gender differences in mathematically intensive STEM careers (Schoon & Eccles, 2014; Su, Rounds, & Armstrong, 2009), but their relative predictive validity is unclear so far. To address this question, longitudinal data from 4,984 students in Germany at the end of high school and two years later were reanalyzed. Both expectancy-value constructs and vocational interests predicted math achievement and the choice of STEM majors, but there were important differences in their predictive validity: Whereas expectancy-value constructs were more predictive of math achievement, vocational interests were more predictive of the choice of different STEM majors. Gender differences at the mean levels of expectancy-value constructs and vocational interests partly explained gender differences in math achievement and the choice of a STEM major. Furthermore, whereas the predictive power of expectancy-value constructs for math achievement and the choice of different STEM majors was invariant over gender, the predictive power of vocational interests differed slightly between young women and men. Thus, expectancy-value constructs and vocational interests seem to contribute differently to young men's and women's STEM career pathways.

The findings of the three studies are summarized and discussed within the broader research context. All three studies contribute to a better understanding of gender differences in mathematically intensive STEM careers, but also highlight the complexity of factors involved.

---

# ZUSAMMENFASSUNG

Es entscheiden sich deutlich weniger Frauen als Männer für einen Berufsweg innerhalb der mathematiknahen MINT-Fächer (Mathematik, Informatik, Naturwissenschaft, Technik; Eurostat, 2017; National Science Foundation, 2017). Da diese Branchen häufig sehr gute Karrierechancen bieten, sind solche Geschlechtsunterschiede problematisch im Hinblick auf die Chancengerechtigkeit von Frauen und Männern (European Commission, 2013; Noonan, 2017). Die Erwartungs-Wert-Theorie von Eccles et al. (1983) bietet ein umfassendes Modell, um unterschiedliche Berufswege von Frauen und Männern zu erklären. Eccles und Kollegen (1983, 2009) gehen davon aus, dass Geschlechtsunterschiede in mathematiknahen MINT-Berufsweegen durch bereits früh auftretende Geschlechtsunterschiede in Erwartungs- und Wertüberzeugungen in Mathematik entstehen, welchen wiederum geschlechtsspezifische Sozialisationsprozesse zugrunde liegen. Diese theoretischen Annahmen wurden empirisch vielfach überprüft und eine Vielzahl von individuellen und umweltbezogenen Faktoren wurden identifiziert, die mit Geschlechtsunterschieden in Berufsweegen der mathematiknahen MINT Fächer zusammenhängen (siehe Wang & Degol, 2013; Wigfield et al., 2015). Nichtsdestotrotz sind wichtige Fragen hinsichtlich des Einflusses von bedeutsamen individuellen und umweltbezogenen Faktoren auf solche Geschlechtsunterschiede noch nicht beantwortet. Im Rahmen dieser Dissertation wurde daher einigen dieser zentralen Fragen nachgegangen. Eingebettet in den theoretischen Rahmen der Erwartungs-Wert-Theorie von Eccles et al. (1983) wurden zwei Umweltfaktoren untersucht, welche junge Frauen negativ in der Wahl eines mathematiknahen MINT-Berufswegs beeinflussen könnten: Zum einen Geschlechtsstereotype über Mathematik in den Medien und zum anderen Besonderheiten des Kurssystems in der gymnasialen Oberstufe. Zudem wurden Erwartungs- und Wertüberzeugungen, welche in der Erwartungs-Wert-Theorie von Eccles et al. (1983) proximal mit der Wahl eines MINT-Berufs zusammenhängen, mit zentralen Prädiktoren aus dem Bereich der Arbeits- und Organisationspsychologie (den beruflichen Interessen) im Hinblick auf ihre Vorhersagekraft für geschlechtsspezifische MINT-Berufswege verglichen.

Im Rahmen dieser Dissertation wurden deshalb drei empirische Studien durchgeführt: In Studie 1 wurde der Einfluss von Geschlechtsstereotypen in einer Kinderfernsehsendung auf die Zustimmung zu Geschlechtsstereotypen sowie die Motivation und Leistung in Mathematik von Mädchen und Jungen untersucht. Hierfür wurde eine randomisierte Studie mit insgesamt 335 Kindern der fünften Klasse durchgeführt. Das Ansehen eines kurzen Clips mit Geschlechtsstereotypen über Mathematik im Rahmen einer Fernsehendung erhöhte sowohl bei den Mädchen als auch bei den Jungen die Zustimmung zu Geschlechtsstereotypen.

---

Die Motivation der Mädchen für das Fach Mathematik wurde nicht beeinflusst, und die Motivation der Jungen nur teilweise. Auch die Mathematikleistung der Mädchen und Jungen wurde nicht beeinflusst. Die Ergebnisse deuten somit darauf hin, dass Geschlechtsstereotype in einer Fernsehsendung die Motivation und die Leistung von Mädchen und Jungen kurzfristig nur geringfügig beeinflussen.

In Studie 2 wurde untersucht, ob die verpflichtende Belegung von Mathematikleistungskursen in der gymnasialen Oberstufe dazu führt, dass sich mehr junge Frauen für einen Berufsweg innerhalb der MINT-Fächer entscheiden. Junge Frauen belegen seltener Mathematikleistungskurse als junge Männer, was mit Geschlechtsunterschieden in der späteren Wahl von MINT-Berufen in Verbindung gebracht wird (Ceci, Ginther, Kahn, & Williams, 2014; Watt & Eccles, 2008). Daher wurden Effekte einer landesweiten Bildungsreform in Deutschland untersucht, durch welche alle Schülerinnen und Schüler dazu verpflichtet wurden Mathematikurse auf Leistungskursniveau zu belegen. Hierzu wurden Daten von 4.730 Schülerinnen und Schülern, die Oberstufenkurse vor der Reform belegt hatten, mit Daten von 4.715 Schülerinnen und Schülern verglichen, welche Oberstufenkurse nach der Reform belegt hatten. Die Reform war mit differenziellen Effekten für junge Frauen und Männer verbunden: Unterschiede in der Mathematikleistung zugunsten junger Männer waren in der Kohorte nach der Reform geringer als in der Kohorte vor der Reform. Unterschiede im mathematischen Selbstkonzept und in Interessen im MINT-Bereich (praktisch-technische und intellektuell-forschende Interessen) zugunsten junger Männer waren nach der Reform jedoch noch größer als zuvor. Hinsichtlich der Wahl von MINT-Studienfächern war das Geschlechterverhältnis in den Kohorten vor und nach der Reform vergleichbar, mit jeweils mehr jungen Männern als Frauen in diesen Studiengängen. Die Ergebnisse deuten somit darauf hin, dass die verpflichtende Belegung von Mathematikleistungskursen in der Oberstufe nicht automatisch die Anzahl von Frauen in mathematiknahen MINT-Fächern erhöht.

In Studie 3 wurde die relative prädiktive Validität von Erwartungs- und Wertüberzeugungen sowie beruflichen Interessen für Geschlechtsunterschiede in mathematiknahen MINT-Berufswegen untersucht. Separat untersucht, weisen die Konstrukte alle eine gute Prognose der Geschlechtsunterschiede in Berufswegen innerhalb der mathematiknahen MINT-Fächer auf (Schoon & Eccles, 2014; Su, Rounds, & Armstrong, 2009), ihre relative Vorhersagekraft ist bislang allerdings unklar. Um dieser Frage nachzugehen, wurden längsschnittliche Daten von 4.984 Schülerinnen und Schülern in Deutschland am Ende der Oberstufe und zwei Jahre später reanalysiert. Sowohl Erwartungs-



---

und Wertüberzeugungen als auch berufliche Interessen sagten die Mathematikleistung und die Wahl von MINT-Studienfächern vorher, wenn auch mit bedeutsamen Unterschieden: Während Erwartungs- und Wertüberzeugungen die Mathematikleistung besser vorhersagten, lieferten beruflichen Interessen eine bessere Prognose für die Wahl von verschiedenen MINT-Studienfächern. Mittelwertsunterschiede zwischen Frauen und Männern in Erwartungs- und Wertüberzeugungen sowie beruflichen Interessen erklärten teilweise die Geschlechtsunterschiede in der Mathematikleistung und der Wahl eines MINT-Studienfachs. Die Vorhersagekraft von Erwartungs- und Wertüberzeugungen für die Mathematikleistung und die Wahl eines MINT-Studienfachs war für junge Frauen und Männer vergleichbar, die Vorhersagekraft der beruflichen Interessen unterschied sich jedoch leicht zwischen jungen Frauen und Männern. Erwartungs- und Wertüberzeugungen und berufliche Interessen scheinen somit unterschiedlich zu Berufswegen im MINT Bereich von jungen Frauen und Männern beizutragen.

Die Ergebnisse der drei Studien werden abschließend zusammengefasst und im Rahmen ihres breiteren Forschungskontextes diskutiert. Alle drei Studien tragen zu einem besseren Verständnis von Geschlechtsunterschieden in Berufswegen der mathematiknahen MINT-Fächer bei, betonen allerdings auch die Komplexität der beteiligten Faktoren.



---

# TABLE OF CONTENTS

<b>1. INTRODUCTION AND THEORETICAL FRAMEWORK.....</b>	<b>1</b>
1.1 Explanations for Gender Differences in Mathematically Intensive STEM Fields.....	7
1.2 The Expectancy-Value Theory.....	11
1.2.1 A powerful framework for understanding gendered careers.....	11
1.2.2 Origins and central assumptions.....	12
1.2.3 Theoretical conceptualization of expectancy and value beliefs.....	14
1.3 Expectancy and Value Beliefs and Gendered STEM Careers.....	21
1.3.1 Associations of expectancy and value beliefs with STEM careers.....	21
1.3.2 Gender and associations of expectancy and value beliefs with STEM careers.....	24
1.4 Factors Influencing Gender Differences in Expectancy and Value Beliefs.....	31
1.4.1 Stereotypes about gender and mathematics.....	31
1.4.2 Different educational choices of girls and boys.....	38
1.5 Expectancy-Value Constructs, Vocational Interests, and Gendered STEM Careers.....	43
1.5.1 Holland’s theory of vocational personalities and working environments.....	44
1.5.2 Contrasting expectancy and value beliefs and vocational interests.....	47
1.5.3 Gender differences in vocational interests and associations with STEM careers.....	49
1.6 Research Questions of the Present Dissertation.....	51
<b>2 STUDY 1: MATH = MALE AND GEEKY: DO GENDER STEREOTYPES IN A TELE- VISION PROGRAM AFFECT GIRLS’ AND BOYS’ MATH PERFORMANCE AND MOTIVATION?.....</b>	<b>55</b>
<b>3 STUDY 2: MAXIMIZING GENDER EQUALITY BY MINIMIZING COURSE CHOICE OPTIONS? EFFECTS OF OBLIGATORY COURSEWORK IN MATH ON GENDER DIFFERENCES IN STEM.....</b>	<b>105</b>
<b>4 STUDY 3: IT TAKES TWO: EXPECTANCY-VALUE CONSTRUCTS AND VOCATIONAL INTERESTS PREDICT STEM CAREERS DIFFERENTLY AND DIFFER BETWEEN MEN AND WOMEN.....</b>	<b>151</b>
<b>5 GENERAL DISCUSSION.....</b>	<b>209</b>
5.1 Discussion of General Findings.....	210
5.1.1 Summary of central findings.....	210
5.1.2 Complexity of gendered career pathways.....	211
5.2 Strengths and Limitations.....	217
5.3 Implications and Future Directions.....	222
5.3.1 Implications and directions for future research.....	222
5.3.1 Implications for educational policy and practice.....	225
<b>REFERENCES.....</b>	<b>228</b>



Introduction and  
Theoretical Framework



---

# 1 Introduction and Theoretical Framework

Gender<sup>1</sup> differences in mathematically intensive STEM (science, technology, engineering, and mathematics) fields exist throughout education and work in many countries around the world such that women are less represented than men. Women currently constitute more than 50% of the total enrollment in bachelor and master education in the EU and US (Eurostat, 2017; National Center for Education Statistics, 2017). However, in engineering, computer sciences, and physics, only about one quarter of students in tertiary education are women in these countries (DG Research and Innovation, 2016; Eurostat, 2017; National Science Foundation, 2017). Relatively more women are represented in mathematics and statistics such that more than 40% of bachelor and master students and about 30% of doctoral students are female, but there are still fewer women than men in these subjects (DG Research and Innovation, 2016; Eurostat, 2017; National Science Foundation, 2017). The gender gap seems to increase at the occupational level. Only about 15% of engineers and 25% of computer scientists are women (DG Research and Innovation, 2016; National Science Board, 2016), and even larger gender differences have been found in academia, where women hold 20% or even fewer of the full time research positions in engineering, technology, and natural sciences in most countries of the EU and the US (DG Research and Innovation, 2016; National Science Board, 2016).

The “leaky STEM pipeline” is a metaphor that is frequently used to describe these gender differences across all educational and attainment levels in mathematically intensive STEM areas, indicating that women “drop out” of this pipeline at various points and more frequently than men do (e.g., Blickenstaff, 2005; Steffens, Jelenec, & Noack, 2010; Watt, Eccles, & Durik, 2006). In politics, economics, and research, major concerns about the “leaky STEM pipeline” have been raised in terms of gender equity in rights and opportunities and in terms of the efficiency of the larger economy (Ceci, Williams, & Barnett, 2009; European Commission, 2013; Hunt, Layton, & Prince, 2015). For example, a more gender-balanced

---

<sup>1</sup> In the literature on differences between women and men in mathematically intensive STEM fields, the terms “sex” and “gender” are both commonly used, sometimes even interchangeably (see Ceci & Williams, 2010; Ceci et al., 2009; Eccles, Wigfield, et al., 1993; Wigfield, Battle, Keller, & Eccles, 2002). Although the terms do not refer to constructs that can be clearly separated, “sex” and “gender” have been defined as describing related but different parts of masculinity and femininity (Deaux, 1985; Lips, 2008). Whereas “sex” has been defined as referring to biological aspects of being male or female, “gender” has been defined as referring to a social category such as cultural expectations of females and males (Deaux, 1985; Lips, 2008). Because the present dissertation focused on differences between men and women from a sociocultural perspective the term “gender” is used throughout.

participation in mathematically intensive STEM fields would contribute to gender equity in income because there are usually high-status career options in these fields (i.e., higher earnings and lower unemployment rates compared with other occupations; Langdon, McKittrick, Beede, Khan, & Doms, 2011; Noonan, 2017). Furthermore, gender differences in mathematically intensive STEM fields are also crucial to the larger economy because it is economically inefficient when talented and qualified women drop out of the STEM pipeline (European Commission, 2013; Hunt et al., 2015). Moreover, a more gender-diverse population of STEM employees could also lead to higher quality in research and better performance as well as a greater possibility of creating products that fulfill the needs of the population at large (European Commission, 2013; Hunt et al., 2015). In addition, knowledge and skills in STEM are essential for understanding fundamental elements of the world and people's everyday lives (Bybee, 1997; Conference of Ministers of Education, 2009; OECD, 2016). Science and technology, for example, are of major importance in our world and affect almost every part of it, such as the production of food, health treatments, or the generation of energy, to name only a few (Federal Ministry of Education and Research, 2012; OECD, 2016; Sjøberg, 2001). A sophisticated understanding of science and technology is thus necessary for both women and men to participate as responsible citizens in a society that is determined by science and technology (Jones, Wheeler, & Centurino, 2015; OECD, 2016).

A powerful theoretical framework for understanding gender differences in choices and outcomes in mathematically intensive STEM areas is provided by Eccles et al.'s (1983) expectancy-value theory. In this theory, gender differences in STEM careers are linked to early emerging gender differences in math expectancy and value beliefs that relate to individuals' perceptions about how they are able to succeed in a task and how much they want to engage a task (Eccles, 2009; Eccles et al., 1983). According to Eccles and colleagues (1983, 2007a, 2009), gender differences in these beliefs are rooted in different socialization processes for girls and boys such as the gender stereotypes that children and adolescents encounter in their environments or the choices females and males make on the basis of their socialization, such as course selection in high school.

Since Eccles et al.'s (1983) expectancy-value model was first published more than 30 years ago, the model has been applied in a wide selection of research to investigate gender differences in mathematically intensive STEM fields (for recent reviews, see Wang & Degol, 2013; Wigfield et al., 2015). Such work has shown that expectancy and value beliefs are highly predictive of gender differences in mathematically intensive STEM fields over and above other motivational and cognitive constructs (e.g., Schoon & Eccles, 2014; Watt & Eccles, 2008;



Wigfield & Cambria, 2010). Also, studies have identified a variety of factors that influence females' and males' expectancy and value beliefs differently throughout childhood and adolescence (for reviews, see Eccles, 2007a; Wang & Degol, 2013). Two environmental factors stand out as particularly relevant for gender differences in mathematically intensive STEM disciplines from a sociocultural perspective: the experience of stereotypes that link mathematically intensive STEM domains to males rather than to females, and the requirements of the educational system that make it possible for young women to opt out of the mathematically intensive STEM pipeline when they are still in high school. Nevertheless, important questions still remain open: Whereas there is profound evidence for how the endorsement of gender stereotypes is linked to different socialization processes for girls and boys, not much is known about the concrete mechanism through which stereotypes experienced in the environment influence children's motivation. Furthermore, ample research has linked gender differences in high school coursework (i.e., the choice of advanced courses in mathematics) with gender differences in mathematically intensive STEM careers. However, it is unclear whether encouraging girls to take advanced courses in mathematics would be an adequate option for reducing gender differences in later STEM careers. In addition, to better understand the complex process of gendered career choices, researchers have started to investigate how expectancy and value beliefs are related to other motivational constructs that predict career choices and how they contribute differently to explaining gender differences in mathematically intensive STEM careers. However, such work has primarily focused on the constructs that have typically been used in educational psychology so far, and not much is known about associations with other constructs that are highly predictive of gendered career choices, for instance, constructs that are often investigated in the area of vocational and work psychology.

With the overarching goal to deepen the understanding of gender differences in mathematically intensive STEM domains, this dissertation extends research on the expectancy-value theory with respect to these research gaps. On the one hand, critical environmental factors that exist throughout childhood and young adulthood are investigated to determine how they might contribute to how women "drop out of the STEM pipeline," namely, gender stereotypes experienced in television programs and the encouragement of young women to choose advanced math courses in high school. On the other hand, associations between expectancy and value beliefs with vocational interests—which are among the most widely used constructs with respect to gendered career choices in work and vocational psychology—as well as their relative contribution in explaining gendered STEM careers are investigated.

The present dissertation is structured as follows: In the introductory chapter, the three empirical studies that were conducted are placed within their broader research context. In the first section of the introduction, different explanations that have been proposed to explain gender differences in mathematically intensive STEM disciplines are presented to locate Eccles et al.'s (1983) expectancy-value theory in a broader research context. Afterwards, Eccles et al.'s (1983) expectancy-value theory is described with its strengths, origins, and central assumptions. In this section, theoretical conceptualizations of expectancy and value beliefs are also summarized, as these are the most proximal constructs for explaining achievement and achievement-related choices in the expectancy-value theory. In the third section, associations between expectancy and value beliefs with gender differences in mathematically intensive STEM careers are discussed in depth. Environmental factors that might reduce women's participation in mathematically intensive STEM disciplines are presented in the fourth section. This section focuses on the role of stereotypes in which mathematics and science are more strongly associated with males than with females as well as the roles of the different educational choices that female and male students make in school. In the fifth section, associations of expectancy and value beliefs with vocational interests as well as differential associations of all of these constructs with gender differences in mathematically intensive STEM disciplines are presented. At the end of the first chapter, the research questions that guide the three empirical studies are introduced.

In the three chapters that follow the introduction, the three empirical studies conducted as part of this dissertation are presented. In the first study, effects of stereotypes about gender and mathematics embedded in a children's television program on girls' and boys' stereotype endorsement, math achievement, and math motivation are investigated. In the second study, effects of a school reform on gender differences in mathematics are examined. More specifically, the second study investigates how a school reform that encouraged all students to participate in advanced math courses in high school influenced gender differences in math achievement, math motivation, and vocational interests. In the third study, differential associations between expectancy and value beliefs, vocational interests, and gender differences in mathematically intensive STEM disciplines (i.e., math achievement and the choice of STEM majors) are examined.

The findings of the three empirical studies are summarized and discussed within the broader conceptual framework in the final chapter of this dissertation. Implications for future research and educational practice are also discussed in this chapter—they mark the end of this dissertation.

## 1.1 Explanations for Gender Differences in Mathematically Intensive STEM Fields

In the present dissertation, Eccles et al.'s (1983) expectancy-value theory is used as a guiding framework by which to investigate gender differences in mathematically intensive STEM fields. Eccles et al.'s (1983) expectancy-value theory is a particularly powerful framework for understanding and investigating different career pathways for females and males in many respects. This theory, along with its strengths, assumptions, and constructs, is described in more detail in Section 1.2. In order to locate the assumptions of the expectancy-value theory within the variety of proposed explanations for gender differences in mathematically intensive STEM fields, research on gender differences in mathematically intensive STEM disciplines is first discussed from a broader point of view.

Gender differences in mathematically intensive STEM fields have been the focus of researchers from a broad range of disciplines such as psychology, sociology, genetics, or economics (see Ceci et al., 2009; Halpern et al., 2007; Wang & Degol, 2013; Yazilitas, Svensson, & Vries, 2013). Corresponding to the broad range of disciplines, various explanations for the underrepresentation of women in these fields have been provided, spanning from individual to broad contextual factors such as differences in biology, ability, cultural influences, and the different preferences and choices made by females and males (for reviews, see Ceci & Williams, 2010; Ceci et al., 2009; Wang & Degol, 2013). A comprehensive review of explanations for gender differences in mathematically intensive STEM domains is provided by Ceci et al. (2009), who synthesized evidence from psychology, education, endocrinology, genetics, sociology, and economics. On the basis of evidence included in this review, Ceci et al. (2009) argued that research on differences in biological factors (e.g., different hormone levels, brain sizes, or brain structures) between the genders is rather insufficient for explaining effects of biological characteristics on gender differences in mathematically intensive STEM disciplines because the results are mixed and limited and the causal relations are unclear. Nevertheless, according to Ceci et al. (2009), biological differences may influence gender differences in mathematically intensive STEM fields. On the basis of the evidence considered in the review, however, the authors were not able to disentangle the biological and environmental effects nor were they able to disentangle the causal links between the two (see also Wang & Degol, 2013, for another review with a similar conclusion). In addition, research on biological differences has often been conducted with a focus on gender differences in math abilities rather than on gender differences in mathematically intensive STEM fields directly (Ceci et al., 2009). The assumption here is that gender differences in such abilities would

mediate the associations between biological features and gender differences in STEM fields (Ceci et al., 2009; Wang & Degol, 2013).

Broad cultural and sociocultural factors, however, are likely to influence gender differences in mathematically intensive STEM fields according to Ceci et al. (2009). This conclusion is based on variation in gender differences in mathematically intensive STEM disciplines in terms of participation and achievement across and within countries and across cohorts (for further support, see also reviews by Ceci & Williams, 2010; Cheryan, Ziegler, Montoya, & Jiang, 2016; Wang & Degol, 2013, 2017). Nevertheless, on the basis of evidence included in the review, Ceci et al. (2009) found no direct link between such factors and gender differences in mathematically intensive STEM disciplines. Rather, the authors found that these factors influence females and males differently throughout their socialization, leading to gender differences in preferences, choices, and math achievement in the end (for further support, see also reviews by Ceci, Ginther, Kahn, & Williams, 2014; Ceci & Williams, 2010; Wang & Degol, 2013). By contrast, differences in men's and women's preferences and choices have been identified as the primary factor that explains gender differences in STEM disciplines (Ceci et al., 2009). In addition, gender differences in math achievement have also been found to contribute to gender differences in mathematically intensive STEM domains, although these tend to be a secondary factor (Ceci & Williams, 2010; Ceci et al., 2009; Cheryan et al., 2016; Wang & Degol, 2013, 2017). In the following, the roles of preferences and choices as well as math achievement as the most proximal factors for gender differences in mathematically intensive STEM disciplines are discussed further.

### ***The roles of preferences and choices***

The evidence considered in Ceci et al.'s (2009) review indicated that differences in the career pathways of women and men are mainly based on different preferences and choices. This finding was also supported by another Ceci et al. (2014) review in which the authors found that different career preferences already emerge at early ages: Even in preschool, girls are less likely than boys to report career aspirations for mathematically intensive STEM domains, with increasing gender differences thereafter. In line with these findings, Su, Rounds, and Armstrong (2009) documented in a recent meta-analysis that women tend to prefer working with people and thus prefer socially oriented occupations (e.g., teaching professions or occupations in the medical or biological sciences). Men, however, tend to report preferences for working on "things" (i.e., technical and mechanical activities) and therefore prefer occupations involving such activities (i.e., engineering or mathematical and physical sciences).

As additional support for the roles of preferences and choice, Ceci et al. (2009) found that women prefer careers outside of mathematically intensive STEM domains even if they have the abilities needed to participate in certain disciplines (i.e., high math ability). Conversely, men with high math ability tend to choose careers in mathematically intensive STEM domains (Ceci et al., 2009). Different intraindividual patterns of achievement in various domains might be one factor that contributes to such different preferences between male and female high performers. In a longitudinal study, Wang, Eccles, and Kenny (2013) found that individuals who have both high math and high verbal abilities are more likely to choose careers outside of mathematically intensive STEM fields. Because women are more likely to be in this group, it has been argued that women's choices are steering them away from STEM (see also Riegle-Crumb, King, Grodsky, & Muller, 2012). Nevertheless, in this study, Wang and colleagues (2013) also supported the importance of preferences and choices for gender differences in mathematically intensive STEM disciplines.

### *The role of math achievement*

In addition to preferences and choices, Ceci et al. (2009) identified gender differences in math achievement as an important factor that contributes to gender differences in mathematically intensive STEM disciplines (for further support, see also the review by Ceci & Williams, 2010). Gender differences in math achievement have often been reported and linked to gender differences in mathematically intensive STEM careers (e.g., Benbow, Lubinski, Shea, & Eftekhari-sanjani, 2000; Hyde, Fennema, & Lamon, 1990). Indeed, math achievement is a key predictor of STEM career choices and also acts as a gatekeeper to such careers because specific math prerequisites are often required in such careers (Ackerman, Kanfer, & Beier, 2013; Harackiewicz, Barron, Tauer, & Elliot, 2002; Parker et al., 2012; Sells, 1980). Historically, males have tended to outperform females on math achievement tests (e.g., Hyde et al., 1990). More recent meta-analytic research, however, has indicated rather small advantages for males compared with females in math achievement at the mean level (e.g., Else-Quest, Hyde, & Linn, 2010; OECD, 2015; Reilly, Neumann, & Andrews, 2015). Furthermore, boys seem to display such advantages only on achievement tests (e.g., Reilly et al., 2015), whereas girls have been shown to have an advantage in teacher-assigned school marks (Voyer & Voyer, 2014).

Nevertheless, in the group of high performers, males still tend to outperform females in math achievement (for reviews, see Ceci & Williams, 2010; Ceci et al., 2009). The group of high performers is particularly important when careers in mathematically intensive STEM

fields are investigated because individuals in this group are most likely to enter such careers (see Ceci et al., 2009). There has been a discussion about the threshold that marks the most important group of high performers for STEM careers (see Ceci & Williams, 2010). However, actually noticeable differences between females and males have been found among the top 1% of performers, for instance, such that there are twice as many males as females in this group (see Ceci & Williams, 2010). Such gender differences are especially important with respect to gatekeeper tests (e.g., the mathematics SAT<sup>2</sup>), which are used to limit access to STEM careers (e.g., due to the requirements of university majors; see Ceci & Williams, 2010; Harackiewicz et al., 2002). But gender differences among high performers do not reflect gender differences in mathematically intensive STEM fields in which women represent only about 25% of physics, engineering, or informatics majors (DG Research and Innovation, 2016; National Science Foundation, 2017). Thus, gender differences in math achievement contribute to gender differences in mathematically intensive STEM disciplines as secondary factors (see Ceci & Williams, 2010; Ceci et al., 2009; Cheryan et al., 2016).

To summarize Section 1.1, a variety of factors have been proposed to be related to gender differences in mathematically intensive STEM domains. Different preferences and choices have been identified as the most proximal factors. In addition, gender differences in math achievement have also been found to be associated with differences in the careers of men and women but as a secondary factor. In order to understand and investigate gender differences in mathematically intensive STEM careers, it therefore seems promising to focus on gender differences in preferences and choices and also in math achievement.

---

<sup>2</sup> The SAT Subject Tests are standardized multiple-choice tests for different subjects (including mathematics), which are provided by The College Board of the US. The SAT Subject Tests are college admission tests that students usually take to improve their credentials for admission to colleges in the US (The College Board, 2017).

## 1.2 The Expectancy-Value Theory

Eccles et al.'s (1983) expectancy-value theory is a particularly powerful theoretical framework for explaining gender differences in academic achievement and achievement-related choices such as career choices. In this section, the theory is described in detail. The strengths of this framework are presented first, followed by descriptions about the main assumptions and central constructs.

### 1.2.1 A powerful framework for understanding gendered careers

Eccles et al.'s (1983) expectancy-value theory is designed to explain academic achievement and achievement-related choices and has often been applied to investigate gender differences in mathematically intensive STEM disciplines. In this regard, however, the expectancy-value theory is not the only theoretical framework. In fact, gender differences in occupational preferences and choices as well as in math achievement have been the focus of a wide range of prominent motivation and interest theories, as such theories in general attempt to explain individual behavior including choices and achievement (Hidi, 2006; Pintrich, 2003). Next to expectancy-value theories (see Eccles et al., 1983), these include interest theories, which focus on the role of interests for gendered career paths such as Holland's model of vocational interests (see Holland, 1997; Renninger, Nieswandt, & Hidi, 2015; Su & Rounds, 2015). Other examples are attribution theories, which focus on interpretations of success and failure (see Graham & Williams, 2009; Stipek & Gralinski, 1991); self-efficacy theories, which emphasize the role of self-efficacy beliefs (see Bandura, 1997; Bussey & Bandura, 1999); and achievement goal theories, which stress the role of goal orientations (see Maehr & Zusho, 2009). All of these theories apply to the understanding of gendered academic behaviors and achievement.

Of these theories, Eccles et al.'s (1983) expectancy-value theory provides a particularly powerful framework for understanding gendered academic choices and outcomes in mathematically intensive STEM fields for several reasons: First, in this theory, multiple theoretical perspectives on motivation are synthesized, and key motivational components that are included in numerous motivation and interest theories are integrated to describe achievement and achievement-related behaviors (see Barron & Hulleman, 2015). Furthermore, in the expectancy-value theory, Eccles et al. refer to the influence of a variety of individual and contextual factors that are assumed to shape and manifest gender differences in motivation, leading to gender differences in STEM careers in the end (see Wigfield et al., 2015). Thus, they present a comprehensive theoretical model for studying gender differences in mathematically

intensive STEM fields that takes into account motivational components and antecedents. In addition, Eccles et al.'s (1983) expectancy-value theory has traditionally been applied to focus on the development of motivational beliefs throughout childhood, adolescence, and young adulthood (see Schoon & Eccles, 2014). The period before individuals actually enter different occupational pathways at university or vocational training is particularly important for gender differences in mathematically intensive STEM disciplines, as career preferences have been shown to develop and manifest during this timespan (for a review, see Ceci et al., 2014). Finally, the expectancy-value framework has been widely applied in research on gender differences such that gender differences in motivational beliefs, their role in gendered career pathways, and factors related to the different motivational beliefs of females and males have been broadly studied all over the world (for recent reviews, see Eccles, 2009; Wang & Degol, 2013; Wigfield et al., 2015). Thus, there is ample evidence for the associations proposed in this comprehensive theoretical framework.

In the following, Eccles et al.'s (1983) expectancy-value theory is described. Its historical origins and main assumptions are presented first, followed by a section on the conceptualization of expectancy and value beliefs as the most proximal influences of choices and achievement.

### **1.2.2 Origins and central assumptions**

Eccles et al.'s (1983) expectancy-value theory is built on previous expectancy-value theories, which have a long tradition in motivation research (e.g., Atkinson, 1957). In the late 1950s, Atkinson (1957, 1964) developed the first formal expectancy-value model, which included motives for success, expectancies for success, and values as determinants of achievement behaviors. In this model, gender differences in achievement-related behaviors were therefore explained by different motives for success, including motives for avoiding failure (Meece, Glienke, & Askew, 2009). In addition to being criticized for methodological biases, however, this work has also been criticized for judging females' achievement behavior against males'. It has been argued that gender differences should instead be understood as outcomes of different educational and occupational choices influenced by gendered socialization and education processes (see Eccles, 1994; Meece et al., 2009).

Building on Atkinson's (1957, 1964) work, Eccles et al. (1983) developed a comprehensive social cognitive model to explain gender differences in educational and vocational choices. Compared with earlier expectancy-value theories, expectancy and value components were more elaborately defined in Eccles et al.'s (1983) expectancy-value theory,



which thus allows gender differences in STEM outcomes to be linked more concretely to specific components of the model. Furthermore, a wide range of psychological and sociocultural determinants are included in this model, and these are assumed to influence achievement-related choices and performance mediated by individuals' task perceptions and interpretations of previous experiences (see Wigfield et al., 2015). The complex process of gendered career choices is therefore reflected in the model, and multiple sets of factors related to gender differences in STEM careers can be investigated by employing this model. In addition, whereas research based on earlier expectancy-value models mostly focused on explaining behavior in laboratory settings, research that has been based on Eccles et al.'s model has tended to apply the model to real-world settings (see Wigfield et al., 2015).

Eccles et al.'s (1983) expectancy-value theory was initially developed to explain gender differences in the domain of mathematics, and therefore, many researchers have applied the model to gender differences in mathematically intensive STEM domains (e.g., Schoon & Eccles, 2014; Wang & Degol, 2013; Watt & Eccles, 2008). Nevertheless, the model has meanwhile also been applied to various other domains such as languages, arts, sports, and health (e.g., Eccles & Harold, 1991; Wigfield et al., 1997).

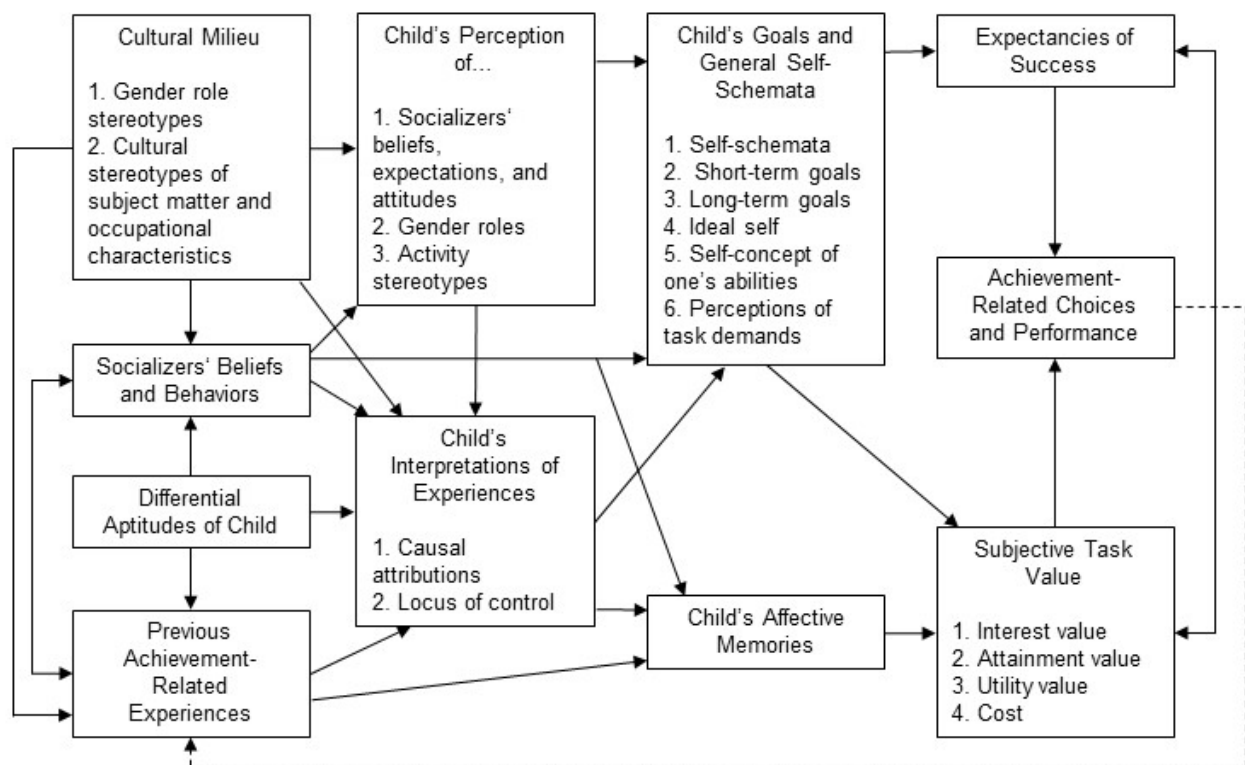


Figure 1. Eccles et al.'s (1983) expectancy-value model of achievement-related choices and performance (adapted from Wigfield et al., 2015).

Figure 1 presents an updated version of Eccles et al.'s (1983) model (see Wigfield et al., 2015). In this model, the expectancy of success and subjective task values are the most proximal determinants of achievement-related choices and performance, pictured on the right side of the model. Moving from right to left in the model, both expectancy and value beliefs in turn are influenced by goals and self-schemata, such as short-term and long-term goals, perceptions of the self and of task demands, or competence beliefs. Values are also influenced by affective memories. These goals, self-schemata, and affective memories themselves are influenced by a person's perceptions of others' attitudes and expectations of them, such as socializers' beliefs, gender roles, and activity stereotypes, along with a person's interpretation of previous achievement experiences. Finally, a person's perceptions and interpretations are influenced by a wide range of social and cultural factors, such as the cultural milieu, including stereotypes, socializers' beliefs and behaviors, or previous achievement-related experiences.

In the expectancy-value theory, gender differences in achievement and achievement-related choices are thus explained by gender differences in expectancy and value beliefs, which are influenced by different socialization processes for girls and boys, shaped by the surrounding environment and its gender norms and roles, individuals' beliefs, and the choices females and males make on the basis of their socialization (Eccles, 2009; Schoon & Eccles, 2014; Wigfield & Eccles, 2000).

### **1.2.3 Theoretical conceptualization of expectancy and value beliefs**

As expectancies for success and subjective task values are the most proximal factors that determine achievement-related choices and performance in the expectancy-value model, Eccles and colleagues considered motivation with respect to two questions, namely, "Can I do this task?" and "Do I want to do this task?" (Eccles et al., 1983; Wigfield et al., 2015). In doing so, two broad categories of beliefs that are emphasized by a variety of social cognitive models are integrated into the expectancy-value model. Such social cognitive models generally focus on different beliefs, values, goals, or interests as proximal determinants of students' motivation, which in turn influence later achievement and achievement-related choices (Wigfield et al., 2015). According to Pintrich, Marx, and Boyle (1993), these constructs refer to two broad categories: beliefs about one's abilities to do tasks on the one hand and reasons for doing these tasks on the other hand, both of which are integral components of Eccles et al.'s (1983) expectancy-value model.

Both components are described in more detail in the following. First, the conceptualization of expectancy and value beliefs in the expectancy-value theory and its

associations to related constructs are presented, followed by a section in which evidence on the structure of expectancy and value beliefs is summarized.

### *Expectancies for success*

Eccles et al. (1983) defined *expectancies for success* as individuals' beliefs about personal success in a future task in either the immediate or the long-term future. Thus, in their expectancy-value theory, expectancies for success are typically measured by questions about how well a person expects to do in a specific domain in the next year or how good a person would be at learning something new in a specific domain (Wigfield & Eccles, 2000).

Expectancies for success are conceptually closely related to other competence beliefs such as self-efficacy (see Bandura, 1997) and academic self-concept (see Marsh, 2007; Marsh & Shavelson, 1985). Similar to expectancies for success, self-efficacy also refers to personal beliefs about succeeding in upcoming tasks. More specifically, Bandura (1997) defined self-efficacy as an individual's beliefs about his or her own ability to successfully perform a specific task or to reach a specific goal. However, expectancies for success and self-efficacy differ in the level of specificity they typically refer to. Although sometimes also measured with respect to specific tasks within domains, expectancies for success are usually assessed as general expectancies with respect to specific domains, often at the level of school subjects such as mathematics (Wigfield & Eccles, 2000). By contrast, self-efficacy has mostly been assessed at task-specific levels by asking how confident an individual feels he or she will be on a specific upcoming task such as "solving an equation such as  $2(x+3) = (x+3)(x-3)$ " (OECD, 2012 p. 89; Bandura, 1997; Bong & Skaalvik, 2003; Pajares, 1996).

Academic self-concept broadly refers to a person's evaluation of his or her current ability in a domain on the basis of the interpretation of his or her previous ability (Marsh & Shavelson, 1985). Comparable to expectancies for success in the expectancy-value model, academic self-concept is typically considered to be a domain-specific construct, mostly studied with respect to school subjects such as mathematics or English (see Marsh & Shavelson, 1985). The time orientation, however, is one major conceptual difference between academic self-concept and expectancies for success. In this regard, Eccles et al. (1983) clearly differentiated between expectancies for success—referring to individuals' expectancies to do well on an upcoming task—and self-concept—referring to individuals' evaluation of their current ability on the basis of their interpretations of their previous ability. Moreover, Eccles and colleagues (Wigfield & Eccles, 2000) suggested that expectancies for success are directly influenced by self-concept, as described above (see also Figure 1 in Section 1.2.2). Nonetheless, researchers

have shown that expectancies for success and academic self-concept are often highly correlated, and Eccles and Wigfield (2002) actually concluded that although they are conceptualized in different ways, the two constructs cannot be distinguished in real-world achievement situations. Thus, in empirical studies, expectancies for success and self-concept have often been collapsed or used interchangeably (see e.g., Lauermann, Tsai, & Eccles, 2017; Nagengast et al., 2011; Trautwein, Lüdtke, Schnyder, & Niggli, 2006; Watt, Shapka, et al., 2012). Because distinctions between different ability and expectancy beliefs are outside the scope of the present dissertation, the term expectancies is used throughout to capture all ability- and expectancy-related beliefs.

### ***Subjective task values***

Besides expectancies for success, individuals' *subjective task values* are also proximal determinants of achievement and achievement-related choices in the expectancy-value model (Eccles, 2005). Individuals' subjective task values refer to reasons for choosing and engaging in different tasks (Eccles & Wigfield, 2002). More specifically, they are related to a person's perception of task quality and how this perception influences a person's desire to engage in the task (Eccles, 2005). Eccles et al. (1983, 2005) emphasized the subjective part of task values in particular because values assigned to a task can differ between individuals. Four different task value components are differentiated in the expectancy-value theory: intrinsic value, attainment value, and utility value—each capturing a person's positive values for a task or domain—and cost, which is related to negative aspects of engaging in a task or domain (Eccles, 2005).

*Intrinsic value* is defined as the enjoyment a person has when performing an activity or the person's subjective interest in a domain (Eccles, 2005). This means that intrinsic value is similar to other motivational constructs such as intrinsic motivation defined by Deci and Ryan (1985) and interest defined by Renninger and Hidi (2011). In Deci and Ryan's (1985) self-determination theory, intrinsic motivation also refers to reasons for engaging in a task or activity such as the inherent satisfaction a person has while performing the task or activity, as discussed for intrinsic value. However, a major difference between intrinsic value and the intrinsic motivation construct from self-determination theory is that the association between intrinsic value and engagement in tasks or activities is always considered in relation to the other value components (see Wigfield & Cambria, 2010). Intrinsic value is also similar to interest as defined by Renninger and Hidi (2011) or Schiefele (2001). In their four-phase model of interest development, Renninger and Hidi (2011) distinguished between situational interest (referring to a person's emotional state while engaging in an activity) and individual interest (referring to a relatively stable disposition toward specific domains). Both situational and individual interest

are related to intrinsic value because intrinsic value is also context-specific and can thus differ between situations and tasks, like situational interest. Furthermore, intrinsic value is also described in relation to long-term engagement with activities, like individual interest (see Wigfield & Cambria, 2010). Compared with intrinsic value, however, Renninger and Hidi's (2011) theoretical conceptualization of interest is more complex, also capturing cognitive components in addition to affective ones (see Wigfield & Cambria, 2010).

*Attainment value* indicates how personally important it is for a person to do well on a task (Eccles, 2005). The definition of this component was built on Battle's (1965, 1966) work on attainment value as well as self-schema and identity theories (e.g., Markus & Wurf, 1987) and is thus closely linked to identity and self-schema. Eccles and colleagues argued that the confirmation or disconfirmation of specific tasks can also act as a demonstration of desired self-schemata (e.g., being masculine or feminine), and individuals thus report high attainment value in a domain that enables them to confirm such aspects of their identity (Eccles, 2005; Eccles & Wigfield, 2002). In this respect, attainment value is similar to the construct of integrated regulation in self-determination theory, which refers to the behaviors people display in order to manifest their own sense of self (i.e., integrating their behavior with their own values and goals; Deci & Ryan, 2004; see Wigfield & Cambria, 2010).

*Utility value* specifies how useful a person feels a task is for the person's subjective goals (e.g., educational or career goals) in the short or long term (Eccles, 2005). As this value component is linked to future goals, utility value is related to more extrinsic reasons for valuing a domain (i.e., for achieving a desired aim). It is thus possible for individuals to have high utility value in domains or tasks without enjoying such tasks but because of their importance for a personal goal, such as a career plan (Eccles & Wigfield, 2002). In capturing more extrinsic reasons for engaging in a task, utility value is related to the construct of identified regulation as defined in self-determination theory (Eccles, 2005). Similar to utility value, identified regulation refers to behavior that is due to the conscious personal valuing of a specific goal (Deci & Ryan, 2004) and describes situations in which people engage in a task or activity not for its own sake but in order to achieve a desired aim (see Wigfield & Cambria, 2010).

Finally, *cost* refers to a negative component of values, as it is defined as the perceived negative aspects of performing a task (Eccles, 2005). This component includes negative emotions such as performance anxiety and fear of failure or success, costs that are related to the amount of effort that is necessary to succeed in the task and the costs of losing other opportunities as a consequence of choosing one option instead of others (Eccles & Wigfield, 2002). Compared with the other value components, cost has been the least studied component

in the expectancy-value theory (see Wigfield & Cambria, 2010). Nonetheless, cost has recently received more attention in theory and research, with some researchers already suggesting that cost should be viewed as a distinct component in the expectancy-value theory next to expectancies for success and subjective task values, rather than as one subcomponent of the latter one (Barron & Hulleman, 2015). This discussion, however, is not within the scope of the present dissertation and is therefore not taken any further.

### ***Structure of expectancy and value beliefs***

The proposed structure of expectancy and value beliefs in the expectancy-value theory has been tested in research. Such work has mainly centered around two aspects, that is, the structure of expectancy and value beliefs within and between domains (see Wigfield et al., 2015). The results on both aspects are summarized in the following.

With respect to the structure of expectancy and value beliefs within domains, expectancy and value beliefs have been found to be related but distinct factors as early as the first grade (Eccles & Wigfield, 1995; Eccles, Wigfield, Harold, & Blumenfeld, 1993). Eccles, Wigfield, et al. (1993), for instance, found that students in Grades 1 to 4 could clearly differentiate between expectancies and values in the domains of both mathematics and reading. However, students could increasingly differentiate between the different value components only within each domain. Eccles, Wigfield, et al. (1993) found that elementary school children did not distinguish between the value components. Instead, the intrinsic, attainment, and utility values were found to build one combined value factor. For secondary school students, however, Eccles and Wigfield (1995) found separate factors for the intrinsic, attainment, and utility values from Grade 5 on (see also Gaspard, Häfner, Parrisius, Trautwein, & Nagengast, 2016). Such differentiations between expectancy and value beliefs within domains beginning in the early grades and differentiations between different value components beginning in Grade 5 have been found in various domains, one of which is mathematics, which is one of the most often studied domains (e.g., Eccles & Wigfield, 1995; Gaspard, Häfner, et al., 2016; for reviews, see also Wigfield et al., 2015; Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006).

Concerning associations between expectancy and value beliefs within domains, Eccles et al. (1983) suggested that cumulative experiences lead to an increase in associations. The rationale behind this assumption is that children will value domains in which they feel competent and will place a lower value on the activities and domains in which they have difficulties (Wigfield et al., 2015). Similar assumptions have also been proposed by other theorists such as Bandura (1997) in his self-efficacy theory, for instance. Several studies have

supported such assumptions, and positive associations within domains from first grade on have been identified and have been found to increase over time (Denissen, Zarrett, & Eccles, 2007; Eccles & Wigfield, 1995; Eccles, Wigfield, et al., 1993; Fredricks & Eccles, 2002; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Wigfield et al., 1997). For example, Denissen et al. (2007) investigated the longitudinal development of associations between self-concept and intrinsic value within domains from Grades 1 to 12 using a U.S. sample. They found correlations of  $r = .20$  between self-concept and intrinsic value in the first and second grades and correlations of  $r = .63$  between the 11<sup>th</sup> and 12<sup>th</sup> grades. Furthermore, using longitudinal data from two large samples in Germany, Marsh, Trautwein, Lüdtke, Köller, and Baumert (2005) found that prior self-concept subsequently influenced later interest, whereas the effects of interest on later self-concept were rather small.

With respect to expectancy and value beliefs in different domains, Eccles et al. (1983) considered these beliefs to be highly domain-specific constructs. Therefore, expectancy and value beliefs have traditionally been examined by aligning domain-specific motivation constructs with domain-specific outcomes, such as associations between expectancy and value beliefs in mathematics with math achievement or the choice of mathematical or mathematically intensive STEM university majors (e.g., Parker, Nagy, Trautwein, & Lüdtke, 2014; Riegler-Crumb, Moore, & Ramos-Wada, 2011; Wang & Kenny, 2014). Also, the domain-specificity of expectancy and value beliefs has been supported as researchers have found that even at early ages, students clearly differentiate between expectancy and value beliefs in different domains (Bong, 2001a; Eccles, Midgley, et al., 1993; Wigfield et al., 1997). Low correlations have been found between beliefs in verbal and mathematical domains, whereas beliefs in mathematics and physics, for instance, have been found to be more closely related (Eccles, Midgley, et al., 1993; Gaspard, Dicke, et al., 2016). Such patterns have been found for different beliefs such as expectancy beliefs, self-concept, and self-efficacy as well as for each of the task value components (e.g., Bong, 2001a; Eccles, Midgley, et al., 1993; Gaspard, Dicke, et al., 2016; Marsh, 1986).

To summarize and conclude Section 1.2, in the expectancy-value theory, prominent motivational constructs are integrated as the most proximal factors for achievement-related choices and achievement, namely, expectancy and value beliefs. Expectancies for success are related to beliefs about one's ability to succeed in a task, and subjective task values (i.e., intrinsic value, attainment value, utility value, and cost) refer to reasons for choosing tasks. Expectancy and value beliefs are considered domain-specific constructs, and students

differentiate between the constructs in different domains beginning in elementary school. Students also differentiate between expectancy and value beliefs within domains from elementary school on, although only secondary school students seem to be able to differentiate between the four different value components. To further build on this conceptualization, the role of expectancy and value beliefs for gender differences in STEM careers is described in the next section.



### **1.3 Expectancy and Value Beliefs and Gendered STEM Careers**

Expectancy and value beliefs are highly predictive of gendered STEM careers. Both expectancy and value beliefs are associated with math achievement and the choice of STEM majors and occupations as well as gender differences in mathematically intensive STEM careers (see Schoon & Eccles, 2014). In the following, associations of expectancy and value beliefs with mathematically intensive STEM careers are first discussed in general. Afterwards, the role of gender in such associations is focused on.

#### **1.3.1 Associations of expectancy and value beliefs with STEM careers**

Expectancy and values beliefs in mathematics and science are closely linked to mathematically intensive STEM careers for both females and males. Many studies have documented that expectancy and value beliefs are associated with career choices as well as with math achievement. Such findings are summarized in the following. Associations of expectancy and value beliefs with mathematically intensive STEM career choices are presented first, followed by a section on associations between such beliefs and math achievement. Finally, the specific components and features of such associations are summarized.

##### ***Associations with mathematically intensive STEM career choices***

Expectancy and value beliefs predict educational and occupational choices over and above achievement (e.g., Simpkins, Davis-Kean, & Eccles, 2006; Wang et al., 2013), intelligence (e.g., Guo, Marsh, Morin, Parker, & Kaur, 2015; Spinath, Spinath, Harlaar, & Plomin, 2006), family characteristics (e.g., Eccles & Wang, 2016; Wang, Degol, & Ye, 2015), and personality (Ackerman et al., 2013). With respect to specific choices, expectancy and value beliefs in mathematics and science are linked to educational aspirations and choices in school, such as the intention to enroll in math courses in high school as well as the actual choice of such courses (e.g., Guo, Parker, Marsh, & Morin, 2015; Meece, Wigfield, & Eccles, 1990; Simpkins et al., 2006; Wang, 2012; Watt, 2005; Watt, Shapka, et al., 2012). Furthermore, math expectancy and value beliefs are associated with career aspirations in high school, such as the aspiration to choose mathematically intensive STEM university majors and the aspiration to have a career in various STEM areas (Chow, Eccles, & Salmela-Aro, 2012; Lauermann, Chow, & Eccles, 2015; Lauermann et al., 2017; Nagengast et al., 2011; Riegler-Crumb et al., 2011; Wang, 2012; Watt, Richardson, et al., 2012). Math and science expectancy and value beliefs are also associated with choices made after high school such as the choice to major in a STEM domain after graduation (e.g., Guo, Parker, et al., 2015; Musu-Gillette, Wigfield, Haring, &

Eccles, 2015; Parker et al., 2014; Wang & Kenny, 2014). Furthermore, math expectancy and value beliefs in high school are also associated with later STEM occupations, for instance, with having a mathematically intensive STEM occupation in one's mid-30s (Eccles & Wang, 2016; Schoon, 2015; Wang et al., 2015).

### ***Associations with math achievement***

Expectancy and value beliefs are also associated with achievement in different domains, including mathematics and science (Ackerman & Heggestad, 1997; Denissen et al., 2007; Parker et al., 2014; Sells, 1980). Expectancy and value beliefs are associated with different achievement outcomes in the same domain, such as teacher ratings, grades, test scores on final exams, or achievement tests (Bong, Cho, Ahn, & Kim, 2012; Denissen et al., 2007; Guay, Marsh, & Boivin, 2003; Marsh et al., 2005; Marsh & Yeung, 1998). Such associations have been found across all educational periods, from elementary school through middle and high school, up to university, across samples in different countries (Bong, 2001b; Bong et al., 2012; Cole, Bergin, & Whittaker, 2008; Denissen et al., 2007; Guay et al., 2003; Marsh et al., 2005; Marsh & Yeung, 1997). It is interesting that expectancy and value beliefs are associated with achievement not only in the short term such as across a time span of 1 year (e.g., Denissen et al., 2007) but also in the long term such as across a time span of several school years (Simpkins et al., 2006).

### ***Specific components and features***

Associations with achievement and career choices have been found for expectancy beliefs as well as for the four value beliefs. Extensive work on how expectancy beliefs and related constructs are associated with achievement has been conducted in research on self-concept (for reviews, see Marsh & Craven, 2006; Marsh & Martin, 2011). However, associations with achievement as well as choices have also been documented for other measures of expectancy beliefs, such as for self-efficacy (Bong, 2001b; Bong et al., 2012) or ability perceptions in general (Meece et al., 1990; Spinath et al., 2006). Researchers have demonstrated that all of the four task value components are associated with achievement and choices in mathematically intensive STEM disciplines, although cost has been the least studied component in this regard (Bong et al., 2012; Cole et al., 2008; Guo, Parker, et al., 2015; Simpkins et al., 2006; Spinath et al., 2006; Trautwein et al., 2012; Wang, 2012; Watt, Shapka, et al., 2012). In comparing expectancy and value beliefs, there is evidence that value beliefs are more closely linked to students' choices, whereas expectancies are more strongly related to students'

achievement (Denissen et al., 2007; Lauermann et al., 2015; Marsh et al., 2005; Meece et al., 1990).

In addition, researchers have investigated additional features of how expectancy and value beliefs are associated with achievement and choices. For example, only recently, Musu-Gillette et al. (2015) found that the change in expectancy and value beliefs during secondary school predicted the choice of university majors (for further support, see also Wang, Chow, Degol, & Eccles, 2017). Furthermore, in more recent studies, the Expectancy  $\times$  Value interaction has also been found to predict achievement and educational choices (Guo et al., 2016; Guo, Parker, et al., 2015; Lauermann et al., 2017; Nagengast et al., 2011; Trautwein et al., 2012). The Expectancy  $\times$  Value interaction was a central component in earlier expectancy-value models such as in Atkinson's model (see Feather, 1982), although in research based on Eccles et al.'s (1983) expectancy-value model, additive associations between expectancy and value beliefs have been considered instead. Nevertheless, in some recent research, the Expectancy  $\times$  Value interaction has also been included to predict choices and achievement (Guo et al., 2016; Guo, Parker, et al., 2015; Musu-Gillette et al., 2015; Nagengast et al., 2011; Trautwein et al., 2012).

How expectancy and value beliefs in multiple domains are associated with STEM outcomes has also been investigated (e.g., Chow et al., 2012; Lauermann et al., 2015). Although expectancy and value beliefs in mathematics and science are highly predictive of mathematically intensive STEM careers, researchers have also found that expectancy and value beliefs in multiple domains predict STEM outcomes (e.g., Chow et al., 2012; Lauermann et al., 2015). For instance, verbal expectancy and value beliefs have been found to be negatively associated with the choices of mathematically intensive STEM careers (Lauermann et al., 2015; Nagy et al., 2008; Parker et al., 2014), and intraindividual patterns of expectancy and value beliefs in a variety of domains have been found to predict the aspiration to choose a mathematically intensive STEM career (Chow et al., 2012; Chow & Salmela-Aro, 2014).

It can thus be concluded that expectancy and value beliefs are powerful predictors of mathematically intensive STEM careers. More specifically, expectancy and value beliefs in mathematics and science are associated with math achievement and the aspiration to choose and the actual choice of mathematically intensive STEM careers in the short term but also in the long run. In comparing expectancy and value beliefs, expectancy beliefs are more strongly associated with achievement outcomes, whereas value beliefs are more strongly associated with choices. Furthermore, changes in expectancy and value beliefs across different ages, the

interaction of expectancy and value beliefs, as well as expectancy and value beliefs in different domains are also associated with mathematically intensive STEM careers.

### **1.3.2 Gender and associations of expectancy and value beliefs with STEM careers**

Associations of expectancy and value beliefs with achievement and choices have often been studied with a special focus on the role of gender (see Schoon & Eccles, 2014; Watt & Eccles, 2008). Most of this research has focused on gender differences in the mean levels of expectancy and value beliefs and how these explain gender differences in achievement outcomes as well as STEM career choices (e.g., Denissen et al., 2007; Marsh et al., 2005; Simpkins et al., 2006; Wang, 2012). In addition, researchers have also investigated how gender moderates the associations between the constructs and STEM outcomes, although to a much smaller extent (e.g., Guo, Parker, et al., 2015; Simpkins et al., 2006; Spinath et al., 2006; Tynkkynen, Tolvanen, & Salmela-Aro, 2012). Both mean-level differences in expectancy and value beliefs between girls and boys and the role of gender as a moderator of associations between expectancy and value beliefs in predicting career pathways are described in the following.

#### ***Gender differences in mean levels***

Gender differences in the mean levels of expectancy and value beliefs have often been reported. These tend to follow a stereotypical pattern with more positive beliefs in mathematics and science for boys, whereas girls tend to show more positive beliefs in verbal domains (for a review, see Wigfield et al., 2015). Such stereotypical gender differences have been found to emerge already in early ages (Fredricks & Eccles, 2002; Steinmayr & Spinath, 2008; Wigfield et al., 1997) and apply to various motivational outcomes including expectancy and value beliefs in mathematics (e.g., Jacobs et al., 2002; Watt, 2004) or different interest constructs such as domain-specific math and science interest and also interest with respect to activities in mathematics and science on a more general level (e.g., Frenzel, Goetz, Pekrun, & Watt, 2010; Su et al., 2009). In the following, gender differences in math and science expectancy and value beliefs are described in more detail. First, the general pattern of gender differences in expectancy and value beliefs in mathematics and science is presented. Then, the development of such gender differences across age is described in two sections: First, the general development of students' expectancy and value beliefs across age is presented in order to better describe differences in the development of girls and boys. Second, gender differences in this general development of expectancy and value beliefs are described.

*Gender differences in expectancy and value beliefs*

With respect to expectancy beliefs, consistent patterns of gender differences in mathematics and science have been found in large-scale and meta-analytic research across Western countries, with boys rating themselves higher than girls (Else-Quest et al., 2010; OECD, 2015). Such gender differences have been found for various competence beliefs such as academic self-concepts (e.g., Else-Quest et al., 2010; Marsh & Yeung, 1998), self-efficacy (Huang, 2013; OECD, 2015), and other self-rated ability measures (Schoon, Ross, & Martin, 2007) in mathematics and related domains such as science. Else-Quest, Mineo, and Higgins (2013), for instance, meta-analyzed TIMSS and PISA data to explore gender differences in math self-concept and self-efficacy across 69 countries. Overall, they found effect sizes<sup>3</sup> for gender differences favoring boys in math self-concept of  $d = 0.33$  and in self-efficacy of  $d = 0.15$  (in TIMSS data) and  $d = 0.33$  (in PISA data). These differences could not be explained by gender differences in math achievement as similar gender differences have been found even when students' achievement was controlled for (Marsh & Yeung, 1998; OECD, 2015).

Although higher expectancy beliefs for boys than girls in mathematics and science and related constructs are consistently reported across Western countries, there is variation in the magnitudes of these differences (Else-Quest et al., 2010; Nagy et al., 2010; OECD, 2015). Nagy et al. (2010), for instance, compared gender differences in math self-concept in samples of secondary school students in Australia, the US, and Germany. In all of these groups, boys showed higher math self-concept than girls, but the gender differences varied: Whereas gender effects appeared to be similar for seventh-grade students in the US and Australia ( $d = 0.16/0.22$ ), differences between girls and boys in the German sample were significantly larger ( $d = 0.52$ ;

---

<sup>3</sup> As suggested by Cohen (1988),  $d$  indicates the effect size calculated as the mean for males minus the mean for females, divided by the pooled within-gender standard deviation. Therefore, values of  $d > 0.00$  indicate higher means for males than for females, whereas values of  $d < 0.00$  indicate lower means for males than for females. According to Cohen (1988), effect sizes of  $d = 0.20$  can be considered small,  $d = 0.50$  medium, and  $d = 0.80$  large. This taxonomy, however, was constructed on the basis of very broad generalizations covering many types of different interventions as opposed to differences between groups in educational settings. This taxonomy should therefore be interpreted with caution (see Hill, Bloom, Black, & Lipsey, 2008). Instead, gender differences in expectancy and value beliefs can be interpreted in comparison with gender differences in other psychological constructs (see Hill et al., 2008; Hyde, 2005). Hyde (2005) reviewed the results of 46 meta-analyses of gender differences in cognitive, social, and personality variables, communication styles and psychological well-being, motor behaviors, and miscellaneous constructs (e.g., moral reasoning). In this review, Hyde (2005) ordered the 124 included effect sizes by range of magnitude. For 30% of the effect sizes, they found  $d = 0.00 - 0.10$ , for 48%  $d = 0.11 - 0.35$ , for 15%  $d = 0.36 - 0.65$ , for 6%  $d = 0.66 - 1.00$ , and for 2%  $d > 1.00$ . Thus, the gender differences in math self-concept and self-efficacy presented here are larger than one third of the effect sizes reported in this review and are in the range where effect sizes on gender differences were usually reported in this review. The same holds for gender differences in task values, which are presented on the next page.

for more support on national variations, see OECD, 2015). Thus, sociocultural factors such as the educational system, for instance, are likely to influence gender differences in expectancy beliefs and related constructs. Such factors, however, are discussed further in Section 1.4.

Similar to gender differences in expectancy beliefs, stereotypical gender differences in value beliefs have also been found (e.g., Gaspard et al., 2015; Meece et al., 1990; Steinmayr & Spinath, 2008; Watt, 2004). In general, males tend to express higher value beliefs and lower cost in mathematics and science, although there is much more work on gender differences in mathematics than in science, whereas girls tend to express higher values and lower cost than boys in verbal domains (Else-Quest et al., 2013; Gaspard et al., 2015; Steinmayr & Spinath, 2008; Watt, 2004). Evidence of gender difference in values, however, has yielded inconsistencies depending on the value component and operationalization. With respect to math intrinsic value, gender differences favoring boys have been found across different samples in Australia and Germany with effect sizes around  $d = 0.33$  (Frenzel et al., 2010; Gaspard et al., 2015; Watt, 2004), whereas for samples in the US, no or only very small gender differences have been found (Nagy et al., 2008; Wang & Kenny, 2014; Wigfield et al., 1997). Regarding math attainment, no differences between girls and boys have been found in various samples (Frenzel et al., 2007; Meece et al., 1990; Steinmayr & Spinath, 2010), and with respect to math utility value, there is again mixed evidence (Steinmayr & Spinath, 2008; Watt, 2004). Watt (2004) found no differences between girls and boys in a sample of high school students in Australia, but Steinmayr and Spinath (2008) found higher utility values for girls than boys for 11<sup>th</sup> graders in Germany ( $d = -0.49$ ). For math cost, higher scores for girls than boys have usually been reported across different samples with  $d$ s around  $-0.30$  (Frenzel, Pekrun, & Goetz, 2007; Gaspard et al., 2015; Meece et al., 1990; Watt, 2004).

Thus, in general, it seems that girls, compared with boys, express lower expectancy beliefs in mathematics and perceive this subject as rather unattractive but as useful or even more useful than boys do. Work by Gaspard et al. (2015), however, suggests that this general pattern might differ when specific subfacets of value components are considered (e.g., girls and boys rate mathematics as equally important for achievement but show differences in the personal importance of mathematics).

#### *The development of expectancy and value beliefs across different ages*

Concerning the general pattern of change in girls' and boys' expectancy beliefs during school, researchers have found first that students have relatively stable beliefs beginning at early ages about their abilities in mathematics as well as in other domains: Even moderate

stability has been found in correlations for timespans of 1 year in elementary school samples in the US and Germany, and these associations tend to become more stable over time (Rieger et al., 2017; Spinath & Steinmayr, 2008; Wigfield et al., 1997).

Second, researchers have identified a general pattern of decline in girls' and boys' expectancy and value beliefs—and an increase in cost—during the school years for students in different countries across domains, including mathematics (e.g., Fredricks & Eccles, 2002; Gaspard, Häfner, et al., 2016; Jacobs et al., 2002; Marsh, 1989; Spinath & Steinmayr, 2008; Watt, 2004). Children report relatively high levels of expectancy and value beliefs at the beginning of elementary school (Archambault, Eccles, & Vida, 2010; Fredricks & Eccles, 2002; Jacobs et al., 2002; Marsh, Craven, & Debus, 1998; Spinath & Steinmayr, 2008), but these beliefs begin to decline during elementary school and continuously decrease during middle school and high school (and vice versa for cost; Fredricks & Eccles, 2002; Jacobs et al., 2002; Spinath & Steinmayr, 2008; Steinmayr & Spinath, 2008; Watt, 2004; Wigfield et al., 1997; Wigfield, Eccles, Iver, Reuman, & Midgley, 1991). Although this pattern is similar across different beliefs and domains, there is some variation: For competence beliefs in mathematics, a study by Jacobs et al. (2002) indicated a linear decline from Grade 1 to Grade 12, contrary to competence beliefs in verbal domains, for instance, which decreased more rapidly during elementary school but only slightly during high school (for further support, see also Fredricks & Eccles, 2002; Watt, 2004). For value beliefs, the decrease in mathematics seems to be even steeper than the decrease in other domains such as languages, biology, or sports (Fredricks & Eccles, 2002; Jacobs et al., 2002). For positive task values in mathematics, studies have indicated a linear decline in general or even an increase in the rate of decline over time (Fredricks & Eccles, 2002; Jacobs et al., 2002; Watt, 2004), whereas value beliefs in verbal domains seem to decrease more rapidly during elementary school but tend to level off in high school (Fredricks & Eccles, 2002; Jacobs et al., 2002; Watt, 2004).

A variety of explanations for this general pattern of decline in students' expectancy and value beliefs have been provided (for a review, see Wigfield et al., 2015). For example, it has been argued that young children are quite optimistic about their competence across all domains, but they judge their achievement more realistically and pessimistically when they grow up (for reviews, see Dweck & Elliott, 1983; Eccles, Wigfield, & Schiefele, 1998). As expectancy beliefs are associated with value beliefs (see Section 1.2), this process should also affect students' value beliefs (see Wigfield et al., 2006). In interest theory, it has been suggested that students increasingly differentiate between their interests in different domains as part of their identity formation (see Daniels, 2008; Krapp, 2002). Such processes should also lead to more

differentiated interest profiles among students (i.e., a decline of a students' interest in some domains, whereas the interest in others remain high), which explains why the mean is declining (see Krapp, 2002). It has also been argued that the school environment contributes to the decline in students' expectancy and value beliefs, due to an increasing mismatch between students' needs and personal goals in adolescence (e.g., the need to control their behavior more and take on more responsibilities) and the school environment (see Eccles & Midgley, 1989; Eccles & Roeser, 2009). However, explanations for this general pattern of decline in students' expectancy and value beliefs are beyond the scope of the present dissertation and are therefore not discussed in detail.

#### *How gender moderates the development of expectancy and value beliefs across age*

In addition to this general trend toward a decrease in students' expectancy and value beliefs—and an increase in cost—in mathematics and other domains during school, different trajectories have been reported for girls and boys (e.g., Fredricks & Eccles, 2002; Marsh, 1989; Watt, 2004). Regarding the development of gender differences during childhood and adolescence, Eccles et al. (1987, 2011) proposed that differences between girls' and boys' expectancy and value beliefs in mathematics should increase over time due to gendered socialization processes.

Indeed, gender differences in expectancy beliefs and some value beliefs in mathematics have been found to already emerge at early ages. Even in elementary school, boys show higher expectancy beliefs in mathematics than girls (Jacobs et al., 2002; Wigfield et al., 1997). Furthermore, in cross-sectional and longitudinal studies, researchers have found variation in gender differences in expectancy and value beliefs over time (e.g., Fredricks & Eccles, 2002; Huang, 2013; Jacobs et al., 2002; Watt, 2004). With respect to expectancy beliefs, Watt (2004) found increasing gender differences over time for a sample of secondary school students in Australia. This increase was driven by a decrease in girls' expectancy beliefs, whereas boys' expectancies remained almost stable during this time. In this study, Watt (2004) thus provided support for the assumption that gender differences in mathematics expectancy beliefs increase over time. Other researchers, however, did not replicate these results. Studies that used other samples in Australia and the US found stable gender differences across preadolescence and middle adolescence (Marsh, 1989; Wigfield et al., 1997). Furthermore, Fredricks and Eccles (2002), for instance, found a more rapid decrease in boys' expectancy beliefs than those of girls between Grades 1 and 12 in a U.S. sample, which thus indicated that gender differences even decreased over time (see also Jacobs et al., 2002).



With respect to gender differences in the development of math value beliefs, there is also only a little support for increasing differences between girls and boys throughout school. Gaspard, Häfner, et al. (2016) found increasing gender differences favoring boys in some subfacets of attainment and utility value in cross-sectional data of students from Grade 5 to 12 in Germany. Furthermore, Watt (2004) found a divergent curvilinear pattern for girls' and boys' perceived cost in mathematics in a sample in Australia, indicating the largest gender differences favoring girls in Grade 9, whereas no gender differences were found in Grades 7 and 11. Both studies, however, found similar changes in girls' and boys' value beliefs with respect to the other components and subfacets that were studied (Gaspard, Häfner, et al., 2016; Watt, 2004). These findings are in line with other studies that also found similar trajectories for girls' and boys' value beliefs in U.S. samples (Fredricks & Eccles, 2002; Jacobs et al., 2002; Wigfield et al., 1997). The development of gender differences in math expectancy and value beliefs is therefore not yet clear and might vary according to environmental factors, such as the educational system. These factors will be further explained in Section 1.4.

### ***Gendered associations of expectancy and value beliefs with STEM careers***

In addition to mean-level differences between girls and boys, there is also work on how associations of expectancy and value beliefs with STEM outcomes differ between girls and boys, although to a smaller extent. In general, associations of expectancy and value beliefs with STEM outcomes have been found to be invariant across gender (Dickhäuser & Stiensmeier-Pelster, 2002; Guo, Parker, et al., 2015; Marsh & Craven, 2006; Marsh et al., 2005; Marsh & Yeung, 1998; Sáinz & Eccles, 2012; Simpkins et al., 2006; Spinath et al., 2006; Tynkkynen et al., 2012). Nevertheless, slightly different associations have also been found. Denissen et al. (2007) investigated associations between self-concept, intrinsic value, and achievement from Grades 1 to 12 in a sample of students in the US and found that self-concept and intrinsic value had smaller associations with achievement for girls compared with boys. Furthermore, Watt, Shapka, et al. (2012) found different associations of expectancy and value beliefs with educational and occupational STEM outcomes in samples of high school students in the US, Canada, and Australia. In this study, a combined attainment-utility value scale predicted the aspiration to choose a mathematically intensive STEM career only for girls but not for boys in Australian and Canadian samples. Furthermore, Korhonen, Tapola, Linnanmäki, and Aunio (2016) found that interest in mathematics was predictive only of girls' educational aspirations, whereas only boys' aspirations were predicted by previous achievement. The reverse pattern, however, was found in a study by Watt et al. (2016), with intrinsic value being the most

important predictor of male students' career preferences, and achievement and self-concept being the most important predictors of females students' career preferences. Guo, Parker, et al. (2015) also found some differential associations between girls and boys for expectancy and value beliefs and STEM outcomes, although most of the investigated associations were similar for girls and boys in their study (for further support, see also Nagy, Trautwein, Baumert, Köller, & Garrett, 2006). Thus, the pattern of how gender might moderate the associations between expectancy and value constructs with STEM careers is still unclear.

To summarize Section 1.3, expectancy and value beliefs are important predictors of mathematics achievement as well as educational and career choices throughout the mathematically intensive STEM pipeline, for both females and males. All of the expectancy-value constructs are related to math achievement and educational and career choices within STEM, although expectancy beliefs are more strongly related to achievement outcomes, whereas value beliefs are more closely linked to choices. Associations between expectancy and value beliefs and STEM careers in general seem to be invariant across gender, but there are also some indications of different associations across gender, although the pattern of the moderating role of gender is not yet clear. Mean-level differences in expectancy and value beliefs between girls and boys, however, tend to be more consistent such that expectancy beliefs in mathematics and science tend to be lower for girls than boys, and girls also tend to score lower in these domains than boys, although such differences seem to depend on the investigated value component.

## **1.4 Factors Influencing Gender Differences in Expectancy and Value Beliefs**

Expectancy and value beliefs do not develop in a “vacuum.” Rather, their development is driven by the interaction of the individual with the surrounding environment, including family, school, or the media. According to expectancy-value theory, both self-socialization as well as external socialization processes (e.g., through socializers’ beliefs and behavior) influence individuals’ task perceptions, their interpretations of previous academic achievement, and their goals and values, leading to specific expectancy and value beliefs in the end (Eccles, 2009; Eccles et al., 1983; see also Section 1.2.2). In their explanations of gender differences in expectancy and value beliefs, Eccles et al. assumed that boys and girls undergo different socialization processes, which they suggested are shaped by the surrounding environment and its gender norms and roles as well as individuals’ beliefs and the choices females and males make on the basis of their socialization (Eccles, 2009; Schoon & Eccles, 2014; Wigfield & Eccles, 2000).

In this regard, gender stereotypes play a major role in the development of gender differences in expectancy and value beliefs as they influence self-socialization processes as well as socializers’ expectations, beliefs, and behaviors (e.g., those of parents, teachers, or peers; Eccles, 2007b, 2011, 2014). In addition, the constraints and opportunities of the environment are also important as they determine the extent to which girls and boys are able to make choices, for instance, in the educational system through coursework requirements in high school (see Wang & Degol, 2013; Wigfield et al., 2015). The roles that stereotypes and high school coursework play in explaining gender differences in math expectancy and value beliefs as well as mathematically intensive STEM careers is described in more detail in the following.

### **1.4.1 Stereotypes about gender and mathematics**

Stereotypes can be broadly defined as associations of group members with specific abilities and attributes (Eagly & Mladinic, 1989; Greenwald et al., 2002). Regarding gender, there are specific stereotypes about the traits, abilities, and motivations of females and males, specifically in the domains of mathematics and science (see Leaper, 2015). Mathematics and science are male-typed domains, and gender stereotypes in these subjects include assumptions that females have lower abilities and less talent in these domains than males (e.g., Spencer, Steele, & Quinn, 1999; Wigfield & Eccles, 2002). Conversely, stereotypes in female-typed domains such as languages include assumptions that females are better at reading and writing than males (e.g., Retelsdorf, Schwartz, & Asbrock, 2015; Steffens et al., 2010). In addition,

stereotypes exist about people who work in STEM professions or people who like STEM subjects (Diekman, Brown, Johnston, & Clark, 2010; Kessels, 2015; Mercier, Barron, & Connor, 2006). Such stereotypes about mathematicians, scientists, or engineers include assumptions that these people are mostly male and geeky; that is, they are seen as socially awkward and as loving technical information and mechanical tasks (Kessels, 2015; Mercier et al., 2006). There are also stereotypes that mathematicians, scientists, and engineers tend to work in isolation rather than in a team or on problems related to communal goals such as helping society (Diekman et al., 2010; Diekman, Clark, Johnston, Brown, & Steinberg, 2011). And there are also stereotypes that people who like mathematically intensive STEM subjects are rather unfeminine and unpopular (see Kessels, 2015).

Such stereotypes are associated with gender differences in mathematically intensive STEM careers. Miller, Eagly, and Linn (2014), for instance, used data from about 350,000 individuals in more than 60 countries to investigate gender differences in tertiary science education. They found that such gender differences were positively associated with national gender-science stereotypes (i.e., stronger stereotypes were associated with larger gender differences; for further support, see also Cundiff, Vescio, Loken, & Lo, 2013).

An explanation for how stereotypes might contribute to gender differences in mathematically intensive STEM careers is provided by the expectancy-value theory. In fact, stereotypes play a major role in the expectancy-value model in which associations between stereotypes and gender differences in mathematically intensive STEM careers are described by different gendered socialization processes (Eccles, 2009, 2011; see also Figure 1 in Section 1.2.2). These gendered socialization processes include differences in socializers' beliefs and behaviors based on stereotypes as well as self-socialization processes, which are both described in the following. Unless stated otherwise, the term gender-math stereotypes is used in the following to describe the stereotype that females have lower abilities and less talent in mathematics compared with males.

### ***Stereotypes and socializers' behaviors and beliefs***

According to Eccles and colleagues, gender-math stereotypes influence teachers' and parents' expectations and beliefs of their students and children, respectively, as well as their behavior toward them. These expectations, beliefs, and behaviors, in turn, are assumed to influence children's goals and self-schemata as well as their expectancy and value beliefs (Eccles, 2009, 2011; Wigfield et al., 2015; see also Figure 1 in Section 1.2.2). There is a large body of evidence for such assumptions. Socializers' beliefs and behaviors have been found to

be related to children's expectancy and value beliefs, and stereotypes have been found to influence parents' and teachers' beliefs and expectations of children as well as parents' and teachers' behaviors toward them (for reviews, see Eccles, 2007a, 2014; Wang & Degol, 2013; Wigfield et al., 2015). For example, parents who endorse gender-math stereotypes are biased in estimating their children's math ability as they tend to overestimate the ability of a son but underestimate the ability of a daughter (for reviews, see Eccles, 2007a; Wigfield et al., 2015). Moreover, parents provide more math-related opportunities for their sons than their daughters (Jacobs & Bleeker, 2004; Jacobs, Davis-Kean, Eccles, & Malanchuk, 2005; Jodl, Michael, Malanchuk, Eccles, & Sameroff, 2001). Furthermore, teachers have differential expectations and beliefs about their male and female students, and such differences in expectations can be related to treating them differently in the classroom such as boys receiving more attention from teachers than girls in mathematics (for reviews, see Eccles, 2007a; Wang & Degol, 2013; Wigfield et al., 2015). Also socializers' expectations and behaviors have been found to be associated with children's expectancy and value beliefs as well as the careers they choose later (for reviews, see Eccles, 2007a; Wang & Degol, 2013; Wigfield et al., 2015). For example, Jacobs and Bleeker (2004) found that mothers' purchases of math and science toys and their involvement in their children's math and science activities are positively associated with children's later math interest. And Bleeker and Jacobs (2004) found that teachers' ratings of students' math ability in Grade 6 and mothers' predictions of their children's success in math careers in Grade 7 were positively associated with individuals' self-efficacy in math-science careers 2 years after high school.

### ***Stereotypes and self-socialization processes***

According to Eccles (1987, 1994), gender-math stereotypes also influence children's and adolescents' self-socialization processes, both consciously and nonconsciously (see also Figure 1 in Section 1.2.2). Eccles (1987, 1994) assumed that individuals develop images of their identity and their ideal selves when they mature and also perceive stereotypes about gender and domains. The two are brought together when individuals develop expectancy and value beliefs in specific domains as well as career choices: Eccles (1987) assumed that individuals assess how their self-images and images about their ideal self, gender, and domains fit together. If individuals perceive a good fit, they develop higher beliefs in such domains and consider careers in such domains as possible options. But if they perceive a mismatch, they develop lower beliefs and tend not to consider such careers as possible options (Eccles, 1987, 2009). Accordingly, girls perceive a mismatch between themselves and stereotypes about people who

are good at mathematically intensive STEM disciplines on which basis they develop career goals that are related to occupations outside of mathematically intensive STEM fields. Conversely, Eccles (2007b) assumed that boys perceive a better fit between themselves and stereotypes about males being good at STEM, which makes it more likely for them to pursue STEM careers as a career goal.

There is also empirical support for the role of stereotypes in self-socialization processes, which is presented along with three questions in the following: How do children endorse gender-math stereotypes? How is the endorsement of gender-math stereotypes associated with expectancy and value beliefs? Where do children experience gender-math stereotypes that influence their expectancy and value beliefs?

### *Children's endorsement of gender-math stereotypes*

With respect to children's endorsement of gender-math stereotypes, evidence has indicated that children are aware of their own gender even at early ages and become increasingly aware of gender stereotypes during childhood and adolescence (see Leaper, 2015; Signorella, Bigler, & Liben, 1993). Several studies have indicated that from elementary school on, children rate boys as being better in mathematics than girls (Cvencek, Kapur, & Meltzoff, 2015; Cvencek, Meltzoff, & Greenwald, 2011; Muzzatti & Agnoli, 2007; Steffens et al., 2010). Cvencek et al. (2011), for instance, investigated the explicit and implicit gender-math stereotypes of children in Grades 1 to 5 in a U.S. sample. They found that girls and boys from second grade on rated boys as being better in mathematics than girls when measured both explicitly and implicitly. In two studies that used a sample of students in Germany, Steffens et al. (2010) found that girls and boys in Grades 4, 7, and 9 rated boys as having higher math abilities than girls (for further support, see also Muzzatti & Agnoli, 2007). Others, however, have not found such stereotype endorsement in children (Ambady, Shih, Kim, & Pittinsky, 2001; Passolunghi, Rueda Ferreira, & Tomasetto, 2014).

These inconsistent findings have been discussed with respect to the adequate assessment of stereotype endorsement, as many different measures have been used in previous research including different explicit and implicit measures (see Greenwald et al., 2002; Nosek, Greenwald, & Banaji, 2007). Furthermore, prior evidence has been criticized because most studies did not differentiate between children's awareness of gender-math stereotypes and their own gender-math stereotype endorsement (for a review, see Régner, Steele, Ambady, Thinus-Blanc, & Huguet, 2014). Thus, findings on children's endorsement of math-gender stereotypes might depend on how these are measured.

*Associations between stereotype endorsement and expectancy and value beliefs*

In some recent studies, associations between stereotype endorsement and expectancy and value beliefs have been documented (Passolunghi et al., 2014; Plante, De la Sablonnière, Aronson, & Théorêt, 2013; Smith, Brown, Thoman, & Deemer, 2015). Passolunghi et al. (2014), for instance, documented that the explicit endorsement of gender-math stereotypes (but not the implicit one) is negatively associated with ability beliefs for girls but positively associated with ability beliefs for boys in a sample of third, fifth, and eighth graders in Italy. However, no associations were found between students' gender stereotypes and their task values in this study for girls or for boys. Furthermore, Plante et al. (2013) found that math expectancy and value beliefs mediated the associations between gender-math stereotypes and grades as well as career aspirations in mathematics for girls and boys in a sample of seventh and eighth graders (ages 11 to 14) in Canada. In this study, such associations were negative for girls and positive for boys. Using samples of 15-year-old students in China, Song and colleagues (2016, 2017) also documented negative associations between the endorsement of gender-math stereotypes and ability and value beliefs in mathematics for girls, which also mediated the association of gender-math stereotypes and career intentions. For boys, they found no associations between the endorsement of gender-math stereotypes and expectancy and value beliefs. In addition, also for female university students, negative associations between the endorsement of gender-math stereotypes and expectancy and value beliefs have been found (Bonnot & Croizet, 2007; Smith et al., 2015).

Taken together, there is evidence for negative associations between the endorsement of gender-math stereotypes and math expectancy and value beliefs for girls from early ages on, whereas boys' stereotype endorsement seem to be not or positive associated with math expectancy and value beliefs.

*Environmental gender-math stereotypes that influence children's expectancy and value beliefs*

With respect to the question of where children experience gender-math stereotypes that influence their expectancy and value beliefs, Eccles et al. assumed that it might be repeated experiences that cause effects to accumulate and might sustainably affect girls and boys, leading to gender differences in mathematics expectancy and value beliefs later on (Eccles, 2009; Wigfield & Eccles, 2000). Yet, there is not much evidence for such mechanisms. There is evidence, however, for potential sources of stereotypes in the environment as well as for short-term effects of the experience of such stereotypes in the environment, which is described in the following.

*Potential sources of stereotypes.* Besides the influence of parents' and teachers' stereotypes as described above, the media seem to be a powerful source of gender stereotypes for children (Steinke, 1997; Steinke et al., 2007). Mass media are central components of children's socialization as more than 70% of children and adolescents watch television every day, on average for about 2.5 hr (Rideout, 2015; Rideout, Foehr, & Roberts, 2010). Television programs with STEM content are increasingly available (National Research Council, 2009) and very popular. For example, the television sitcom *The Big Bang Theory* on the American TV network CBS is one of the highest rated and viewed television shows in the US and many other countries such as Germany (Kirsch, 2011; Patten, 2013). In such programs, certain beliefs and stereotypes about gender roles in the STEM field are presented such as the idea that females have lower math ability compared with males or that scientists are nerdy (Collins, 2011; Heyman, 2008; Rideout, 2015; Steinke et al., 2007). It is therefore not surprising that children state that television and movies are their primary sources of information about how scientists are (Steinke et al., 2007).

*Short-term effects of stereotypes.* There is evidence for how stereotypes experienced from different environmental sources might influence children's and adolescents' motivation in laboratory settings in the short term, provided by research in the area of stereotype threat. The concept of stereotype threat originated in Steele and Aronson's work in the 1990s. They defined stereotype threat as a situational experience in which group members feel concerned about confirming a negative stereotype of their own group (Steele, 1997; Steele & Aronson, 1995). They suggested that such concerns might compromise a person's behavior and performance (Steele, 1997; Steele & Aronson, 1995). Originally, research on stereotype threat was conducted to explain the underperformance of African Americans (Steele & Aronson, 1995), but ample research has also been conducted to address gender differences in math performance and motivation (e.g., Schmader, 2002; Spencer et al., 1999; Tomasetto, Alparone, & Cadinu, 2011). Such research has demonstrated that women show lower math performance if they are reminded of negative stereotypes about women in math, but they perform as well as males if such stereotypes are not made salient before they take a math test (Doyle & Voyer, 2016; Nguyen & Ryan, 2008).

Much of this work has been conducted to explore effects of stereotypes on performance (Doyle & Voyer, 2016; Nguyen & Ryan, 2008), but there is also research that has indicated that experiencing salient stereotypes in the environment can affect motivational outcomes, such as women's expectancy beliefs (Cadinu, Maass, Frigerio, Impagliazzo, & Latinotti, 2003) or interest (Smith, Sansone, & White, 2007; see also Thoman, Smith, Brown, Chase, & Lee, 2013



for a review). Most research on stereotype threat, however, has been conducted on university students or older adults, and less is known about effects of stereotype threat on children or adolescents (e.g., Ambady et al., 2001; Flore & Wicherts, 2015). There is some evidence for effects of stereotype threat on children's and adolescents' performance in mathematics (Ambady et al., 2001; Muzzatti & Agnoli, 2007; Neuville & Croizet, 2007; Tomasetto, Alparone, & Cadinu, 2011). In a meta-analysis that included samples up to the age 18, Flore and Wicherts (2015) found that girls who were reminded of gender-math stereotypes showed slightly lower math performance compared with girls who were not reminded of such stereotypes. Only a little research, however, has been conducted to explore effects of stereotypes on children's motivation. Muzzatti and Agnoli (2007) found effects of stereotype threat on students' expectancy beliefs in eighth graders (but not for third or fifth graders). In addition, Master, Cheryan, and Meltzoff (2015) found that making stereotypes in the field of computer science salient in the classroom environment of a high school computer course reduced female adolescents' interest in enrolling in computer courses compared with female adolescents in a control group.

In addition, a wide range of situations has been identified in which stereotypes about females' underperformance in math can affect girls and women, although mostly in research in laboratory settings. Thereby, stereotypes presented in the media have been found to be powerful in affecting women's performance and motivation, for instance, presented in newspaper articles (Cheryan, Plaut, Handron, & Hudson, 2013), images in schoolbooks (Good, Woodzicka, & Wingfield, 2010), photographs (Muzzatti & Agnoli, 2007), or videos (Davies, Spencer, Quinn, & Gerhardstein, 2002; M. C. Murphy, Steele, & Gross, 2007). Davies et al. (2002), for instance, showed that women avoided tasks related to mathematics after watching gender-stereotyped commercials compared with women who watched gender-neutral commercials. After watching gender-stereotyped commercials, women also showed lower interest in educational and vocational areas that are typically male-stereotyped domains and performed worse on a mathematics test compared with women in the control group. However, there are indications of publication bias in research on stereotype threat (Flore & Wicherts, 2015), and such previous results might thus be questionable.

To summarize Section 1.4.1, gender-math stereotypes seem to be a major factor for explaining gender differences in expectancy and value beliefs, which are important precursors of careers. According to expectancy-value theory, gender-math stereotypes influence socializers' expectations, beliefs, and behaviors as well as the self-socialization processes of girls and boys, both of which in turn are related to gender differences in expectancy and value

beliefs. There is empirical support for such associations, although most work in this area has focused on associations between parents' and teachers' stereotypes with children's expectancy and value beliefs. Also associations between children's stereotype endorsement and their expectancy and value beliefs have been found. Nevertheless, the concrete mechanisms through which stereotypes experienced in the environment might influence children's stereotype endorsement as well as their expectancy and value beliefs have mostly been unclear so far. Some insights have been provided in research in the area of stereotype threat, where the media have been identified as a powerful source of gender-math stereotypes for children. But only a little research has been conducted in samples of children and adolescents. Furthermore, research on stereotype threat has mainly been conducted in laboratory settings, and there are indications of publication bias. Thus, little is known about how stereotypes experienced in real-life settings contribute to children's expectancy and value beliefs.

#### **1.4.2 Different educational choices of girls and boys**

Next to stereotypes, the different educational choices of girls and boys are also assumed to influence gender differences in expectancy and value beliefs as well as later career choices (Eccles, 2007a, 2009). Children and adolescents spend a great deal of their time in school, and the school environment guides the learning process by determining what and how students learn through curricula or the classroom structure or by providing opportunities for students to interact with teachers and peers (for reviews, see Eccles, 2007a; Wang & Degol, 2013; Wigfield et al., 2015). These factors determine students' achievement-related experiences as well as their individual interpretations of such experiences, which in turn influence the development of expectancy and value beliefs (Wigfield et al., 2015; see also Figure 1 in Section 1.2.2). Therefore, if girls' and boys' different educational choices influence such learning environments, they are also likely to influence gender differences in expectancy and value beliefs as well as later career choices. This assumption has also been supported by findings on variation in gender differences in math expectancy and value beliefs across Western countries, as described in Section 1.3. In this regard, one important finding in previous research is related to differences in high school coursework between young men and young women (see Ceci et al., 2014; Meece & Scantlebury, 2006) as presented in the following.

#### ***Advanced math courses and gender differences in math expectancy and value beliefs***

In Germany, young women are less likely to choose advanced math courses in high school than young men, as in most Western countries (Budde, 2009; Kennedy, Lyons, & Quinn,

2014; National Science Board, 2016; The Further Mathematics Support Programme, 2016). This appears to hold at least when they are able to choose different courses and when there are no university requirements for choosing specific advanced math courses in high school. In the US, however, girls participate in advanced math courses at rates that are similar to boys since most universities require advanced courses in mathematics for different majors (Snyder & Hoffmann, 2001; Watt, Shapka, et al., 2012).

As described in Section 1.3, gender differences in math course selection in high school can be explained by differences in expectancy and value beliefs, which are important predictors of such educational choices (e.g., Guo, Parker, et al., 2015; Meece et al., 1990; Watt, 2005; Watt, Shapka, et al., 2012). However, the choice of advanced math courses in high school is not only an outcome but also an important factor that influences later career pathways as such math courses are required by many university majors (e.g., Sells, 1980; Tyson, Lee, Borman, & Hanson, 2007). Because not choosing advanced courses in mathematics can limit access to mathematically intensive STEM careers, it has therefore been argued that gender differences in course-taking behaviors mark one important step where women drop out of the mathematically intensive STEM pipeline (see Ceci et al., 2014; Watt & Eccles, 2008).

Indeed, advanced and basic courses provide different learning experiences, as such courses usually differ in their benefits for and constraints on students (Köller, Baumert, & Schnabel, 2001; Marsh, 2005). These differences usually lead to higher achievement levels for students in advanced courses compared with students in basic courses, even when prior achievement is controlled for (Gamoran & Mare, 1989; Köller et al., 2001). From an expectancy-value perspective, however, effects of gender differences in math coursework in high school on gender differences in mathematically intensive STEM careers are not very clear, as course level is associated with achievement and class composition, both of which are differentially associated with expectancy and value beliefs (e.g., Byun, Irvin, & Bell, 2015; Marsh, 1986; Schurtz, Pfof, Nagengast, & Artelt, 2014). In the following, associations between gender differences in expectancy and value beliefs and high school coursework are described first with a focus on the role of achievement, followed by a section that focuses on the role of class composition in these courses.

### *The role of achievement*

As described in the previous section, advanced and basic courses in high school differ with respect to the course level, that is, advanced courses are usually more demanding than basic courses in terms of the curricular content and the teaching time, and grades in such courses

have a higher impact on final GPA (Brunello & Checchi, 2007; Gamoran & Mare, 1989; Hanushek & Wössmann, 2006; Kelly, 2004; Lucas, 2001). All of these factors are related to students' achievement, and thus, students in advanced courses typically show higher achievement than students in basic courses, even when prior achievement is controlled for (e.g., Byun et al., 2015; Gamoran & Mare, 1989; Köller et al., 2001; Patall, Cooper, & Allen, 2010; Scheerens & Hendriks, 2014).

Expectancy and value beliefs are related to achievement in different ways in the expectancy-value model. On the one hand, expectancy and value beliefs in one domain are strongly related to later achievement in the same domain, as described in Section 1.3.1. Such associations have been explained by the idea that individuals engage more frequently and intensely in tasks and activities that are related to a domain in which they have high expectancy and value beliefs (i.e., show high levels of persistence and effort), which in turn lead to high achievement later on (see Wigfield, Tonks, & Klauda, 2009). As described in Section 1.3.1, there is ample evidence for the proposed associations (Bong et al., 2012; Denissen et al., 2007; Guay et al., 2003; Harackiewicz et al., 2002; Marsh & Craven, 2006; Marsh et al., 2005; Marsh & Yeung, 1998).

On the other hand, achievement in turn also influences individuals' expectancy and value beliefs in the expectancy-value model. It is assumed that prior achievement is positively related to expectancy and value beliefs in the same domain, as achievement-related experiences are assumed to influence children's interpretations of such experiences, their goals, and self-schemata, all of which in turn are associated with expectancy and value beliefs (Eccles & Wigfield, 2002; see also Figure 1 in Section 1.2.2). These associations have also been supported by ample research (e.g., Denissen et al., 2007; Harackiewicz et al., 2002; Marsh et al., 2005; Skaalvik & Skaalvik, 2008). Researchers have documented that prior achievement predicts different measures of expectancy beliefs (see Skaalvik & Skaalvik, 2008), although the most research has been conducted on self-concept (see Marsh & Craven, 2006; Möller, Pohlmann, Köller, & Marsh, 2009). Furthermore, prior achievement has been found to be associated with a variety of interest-related constructs, including subjective task values, although intrinsic interest is the most studied of the task values (Denissen et al., 2007; Harackiewicz et al., 2002; Marsh et al., 2005).

Taken together, advanced and basic courses in mathematics are positively associated with math achievement. Prior achievement predicts the choice of such courses, and the courses also predict later achievement, with higher achievement for students in advanced courses.

*The role of class composition*

In evaluations of their own ability and interest, students do not rely on only their own previous achievement. Rather, their expectancy and value beliefs are also affected by the composition of the class, that is, their peers in class (e.g., Cambria, Brandt, Nagengast, & Trautwein, 2017; Marsh, 1987; Trautwein, Köller, Lüdtke, & Baumert, 2005). Extensive research on the influence of class composition on individuals' motivation has been conducted in research on self-concept, also referred to as the *big-fish-little-pond effect* (Chmielewski, Dumont, & Trautwein, 2013; Liem, Marsh, Martin, McInerney, & Yeung, 2013; Marsh, 1987; Marsh & Hau, 2003; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006). Here, researchers have shown that individuals refer not only to their own prior achievement in a domain when evaluating their abilities but also to their perceptions of achievement in their surroundings. Students who are surrounded by other students with low achievement—as is the case in basic courses—therefore usually judge their own achievement as relatively higher, feeling like a “big fish in a little pond.” Conversely, students who are surrounded by high performers—as in advanced courses—judge their own achievement as relatively lower (Marsh, 1987, e.g., 2005; Niepel, Brunner, & Preckel, 2014; Trautwein, Lüdtke, Marsh, et al., 2006). Thus, students in advanced courses usually show lower self-concept than students in basic courses when achievement is controlled for (Chmielewski et al., 2013; Trautwein, Lüdtke, Marsh, et al., 2006). Such negative effects of the mean level of a class or school achievement on individuals' self-concept have been found for various subjects including mathematics and for different age groups, although effects seem to be stronger for students in secondary than in primary school (Marsh et al., 2014). Furthermore, such negative effects have found to be substantial even 2 years after graduation (Marsh, Trautwein, Lüdtke, Baumert, & Köller, 2007).

Effects of social comparison with respect to task values have been studied relatively less often (Cambria et al., 2017; Köller, Trautwein, Lüdtke, & Baumert, 2006; Schurtz et al., 2014; Trautwein, Lüdtke, Marsh, et al., 2006). Nevertheless, Cambria et al. (2017) recently investigated effects of social comparisons on math task values in a large sample of high school students in Germany. They found negative effects of the average level of achievement in a school on utility value and a combined intrinsic- and attainment-value scale as well as positive effects on cost (for further support for such effects on interest and intrinsic motivation, see also Marsh et al., 2014; Schurtz et al., 2014).

To summarize Section 1.4.2, high school coursework is likely to be associated with gender differences in STEM careers. On the one hand, achievement has been found to be positively associated with both expectancy and value beliefs in the same domain as well as with

course level. On this basis, it would make sense to also expect positive associations between course level and expectancy and value beliefs. On the other hand, negative associations between the mean level of achievement in a class and students' motivation have been found. As girls are less likely to choose advanced mathematics courses in high school than boys, differences in course-taking behavior can be expected to differentially effect their achievement and motivation. Given positive associations between course level and achievement, one could argue that girls' lower participation rates in advanced math courses contribute to gender differences in STEM careers. On the basis of negative associations between course level and expectancy and value beliefs, however, one could argue that young women's lower participation rates in math advanced courses in high school might actually have a positive effect on their motivation. However, effects of gender differences in high school coursework in mathematics on different predictors of STEM careers have also been studied less often, and it is not clear how these differences might contribute to gender differences in STEM careers.

To conclude Section 1.4, expectancy and value beliefs are influenced by different environmental factors. Here, stereotypes about gender and mathematics as well as the different educational choices of young women and men in high school are related to gender differences in expectancy and value beliefs in various ways.

## 1.5 Expectancy-Value Constructs, Vocational Interests, and Gendered STEM Careers

As described in Sections 1.2 and 1.3, the expectancy-value theory is a powerful framework that can be applied to further the understanding of gender differences in mathematically intensive STEM careers. Expectancy and value beliefs are highly predictive of central STEM outcomes such as math achievement and the choice of mathematically intensive STEM majors at university as well as gender differences in these areas (see Section 1.3). Furthermore, there are comprehensive theoretical assumptions as well as empirical support for how (gendered) expectancy and value beliefs develop during childhood and adolescence (see Section 1.3.2). However, as introduced in Section 1.2, a variety of other motivation- and interest-based theories have also been successfully applied to explain gender differences in mathematically intensive STEM careers. Holland's (1997) theory of vocational personalities and working environments, for instance, is one of the most widely used theories in the area of work and vocational psychology, and there is ample evidence indicating that vocational interests are highly predictive of STEM careers as well as gender differences in mathematically intensive STEM fields (for some recent meta-analyses, see Nye, Su, Rounds, & Drasgow, 2017; Su, Golubovich, & Robbins, 2015; Su et al., 2009).

In order to understand the complex process of gendered career pathways, it is essential to understand how the expectancy-value constructs and other central predictors of STEM careers are associated with each other and how they relatively predict gender differences in STEM careers. According to P. K. Murphy and Alexander (2000), such knowledge can deepen the understanding of the similarities and differences between the expectancy-value constructs and other motivational variables and how they impact young women's and men's career paths in relation to each other. Given the complex process of career choices, such knowledge is the basis for counteracting influences that lead females to drop out of the STEM pipeline at various points and may provide helpful ideas for interventions that are designed to enhance women's motivation to work in mathematically intensive STEM domains.

More recently, researchers have started to increasingly investigate how expectancy and value beliefs are associated with other motivational variables as well as how they are differentially associated with achievement outcomes (Ackerman et al., 2013; Gao, 2007; Wang, 2012; Wigfield & Cambria, 2010). For example, there is work on associations of expectancy and value beliefs with motivational variables in the area of self-determination theory (see Wang, 2012; Zhang, Solmon, & Gu, 2012), self-regulation (see Ackerman et al., 2013), self-efficacy theory (see Gao, 2007), sense of belonging (see Goodenow, 1993), and achievement goal theory

(see Conley, 2012; Hulleman, Durik, Schweigert, & Harackiewicz, 2008; Wigfield & Cambria, 2010). There is also work on associations of expectancy and value beliefs with domain-specific interest (see Hulleman et al., 2008; Wigfield & Cambria, 2010). However, such work has tended to focus on variables in the area of educational psychology so far, and less is known about other motivation- and interest-based theories in other psychological areas, such as in vocational and work psychology in which gender differences in mathematically intensive STEM careers are also often studied.

Holland's (1997) theory of vocational personalities and working environment is one of the most prominent theories in and outside the area of vocational and work psychology. The Occupational Information Network (O\*NET) of the U.S. Department of Labor ETA, for instance, has been using Holland's classification of vocational interests for almost 20 years (Mariani, 1999). This model has also been widely applied to investigate career pathways, including career choices where vocational interests have been found to be highly predictive of gendered STEM careers (for more recent meta-analyses, see Humphreys & Yao, 2002; Nye et al., 2017; Rounds & Su, 2014; Su & Rounds, 2015; Su et al., 2009). In explanations of gendered career choices, there are important similarities and differences between the expectancy-value theory and Holland's theory, which are described in the following. As a basis for such comparisons, the main assumptions of Holland's theory are described first, followed by a section on the most important similarities and differences between the theories. Finally, the role of gender in associations of expectancy and value beliefs and vocational interests with STEM careers is described.

### **1.5.1 Holland's theory of vocational personalities and working environments**

In the theory of vocational personalities and working environments (for simplicity, the term *theory of vocational interests* is used throughout), Holland (1997) proposed primary (key) assumptions as well as secondary assumptions, which are described one after the other in the following.

#### *Key assumptions of the theory of vocational interests*

Four key assumptions are central to Holland's theory of vocational interests and describe why people choose different career pathways: First, Holland (1997) postulated that people in Western cultures can be categorized into six interest orientations, which are assumed to represent preferences for activities across the whole range of occupations, namely, Realistic, Investigative, Artistic, Social, Enterprising, and Conventional. Holland (1966) defined



vocational interests as “the expression of personality in work, hobbies, recreational activities, and preferences” (p. 3). Each of these interest orientations represents specific attitudes, skills, and preferences for specific activities or situations in which these activities are located (Holland, 1997). A summary of the interest orientations in tabular form is presented in Table 1. As indicated in this table, investigative interests, in particular, are related to the STEM fields as they refer to preferences for activities in science and research. In addition, realistic interests are also associated with disciplines in the natural sciences and engineering, as such interests describe preferences for typical work activities in mechanical and technical areas. Finally, social interests are also important for STEM disciplines in the field of life sciences (e.g., medicine) because social interests reflect preferences for social activities that include helping other people.

Table 1

*Characteristics of Interest Orientations and Environments*

	<b>The person with this interest orientation prefers, and the environment entails and reinforces...</b>
<b>Realistic</b>	... explicit or systematic manipulation of objects, tools, machines, and animals; ...activities that require strength and coordination and lead to visible outcomes, including being outdoors or working with one’s hands
<b>Investigative</b>	...observational, symbolic, systematic, and creative investigation of physical, biological, and cultural phenomena
<b>Artistic</b>	...manipulation of physical, verbal, or human materials to create art forms or products ...ambiguous, free, and unsystematized activities, which involve creativity, sensitivity, or expression
<b>Social</b>	...manipulation of others to inform, train, develop, cure, or enlighten
<b>Enterprising</b>	...manipulation of others to attain organizational goals or economic gain ...activities in leadership and business
<b>Conventional</b>	...explicit, ordered, systematic manipulation of data, such as counting, statistics, or office work

*Note.* Adjusted from Stoll and Trautwein (2017) and Holland (1997).

Second, there are six environmental models in which work and life environments are categorized as corresponding to the six interest orientations: Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (Holland, 1997). These environmental models are

characterized by the demands and opportunities that are dominant in a work environment. Furthermore, the environmental models are characterized by the people who typically work in these environments. The investigative environment, for instance, is dominated by people who show high investigative interest orientations (Stoll & Trautwein, 2017). A summary of the environmental models is presented in Table 1.

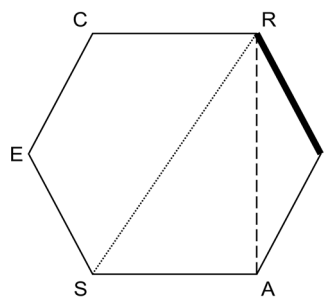
The third assumption describes the association between people and environments: People strive to be in environments in which they can exercise their skills and abilities, express their attitudes and values, and take on roles they agree with (Holland, 1997). Accordingly, people characterized by realistic interests seek realistic environments, people characterized by investigative interests seek investigative environments, and so forth. This assumption thus describes how young people strive for specific careers.

Finally, the fourth assumption describes the idea that behavior is determined by the interaction between interest orientations and environment. Thus, young people's career choices depend on the person's interest orientation and the model of the environment (Holland, 1997; for recent meta-analytic support, see Nye, Su, Rounds, & Drasgow, 2012; Nye et al., 2017; Van Iddekinge, Roth, Putka, & Lanivich, 2011).

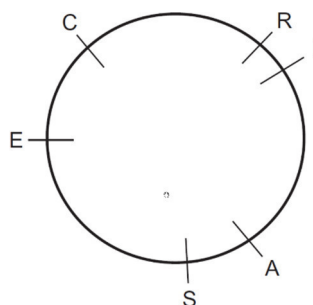
#### *Secondary assumptions in the theory of vocational interests*

In addition to the four key assumptions, Holland (1997) proposed several secondary assumptions in which he further described relations within the six interest orientations as well as relations between the interest orientations and environmental models. Holland (1997) assumed that individuals usually express interests in more than only one of the described interest orientations and can be assigned to some degree to all of the different interests (and work environments can also be characterized by different environmental models). The extent to which individuals can be associated with the six interest orientations differs, with some interests typically being more pronounced than others. Consequently, the expression of the six different interest orientations indicates an individual interest profile (Holland, 1997). According to Holland (1997), the relations between the six interests can be displayed in a hexagonal model (see Figure 2). In such a model, each interest orientation is presented at one angle, and the distances between the angles symbolize the similarity between the particular interests. Consequently, the strongest relations are postulated between interests that are next to each other (e.g., realistic and investigative interests), and the weakest relations are proposed between interests that are opposite one another (e.g., realistic and social interests).

There is empirical support for the postulated relations of the vocational interests in research on the structure of vocational interest, although only with respect to the order of the six orientations (e.g., Anderson, Tracey, & Rounds, 1997). The hexagonal structure has not been found in previous research; rather, the six interest orientations have been found to be best represented in a circumplex model (e.g., Darcy & Tracey, 2007; Tracey & Rounds, 1993; see Figure 3). In such a model, all vocational interest orientations are located on a circular line, with the distances between the vocational interests on the line representing the similarity between the orientations.



*Figure 2.* The RIASEC hexagon, representing the postulated order of the R-I-A-S-E-C dimensions and the assumption of equal distances between the six interest dimensions. Line width represents similarity. The strongest relations are expected for adjacent interests (R-I), smaller relations are expected for nonadjacent interests (R-A), and the weakest relations are expected for opposite interests (R-S; from Stoll & Trautwein, 2017<sup>4</sup>).



*Figure 3.* The RIASEC circumplex, representing the postulated order of the R-I-A-S-E-C dimensions without the assumption of equal distances between the six interest dimensions; this demonstration is not based on empirical data (from Stoll & Trautwein, 2017<sup>4</sup>).

### 1.5.2 Contrasting expectancy and value beliefs and vocational interests

Although there is only a little empirical support for associations of expectancy and value beliefs with vocational interests and their relative importance for gender differences in STEM careers, similarities and differences between the expectancy and value beliefs and vocational interests can be derived on the basis of theoretical considerations:

<sup>4</sup> Reprinted from Stoll, G. & Trautwein, U. (2017). Vocational interests as personality traits. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 401-417). London, San Diego, Cambridge, Oxford: Elsevier, Page No. 403, Copyright 2017, with permission from Elsevier

First, in both theories, the motivational functions of the incorporated constructs (i.e., expectancy and value beliefs and vocational interests, respectively) are stressed, and these influence a person's behavior and outcomes (see Eccles et al., 1983; Holland, 1997). More specifically, Holland (1997) proposed that based on preferences for specific activities, individuals actively engage with their environment, for instance, in seeking or avoiding specific activities or environments, such as specific university majors. This motivational function is thus similar to expectancy-value theory, which explains achievement-related choices on the basis of individuals' motivational constructs (i.e., expectancy and value beliefs; see Section 1.2.2).

Second, Eccles et al. (1983) and Holland (1997) both assume relations between specific values, goals, ability beliefs, and abilities, all of which together determine individuals' preferences for specific careers. As formulated in the first key assumption, Holland (1997) assumed that vocational interest orientations represent a specific set of preferences, attitudes, and skills. Such relations, however, are less explicitly formulated in Holland's model than in the expectancy-value model, which differentiates between different motivational variables (i.e., expectancy and value beliefs), which are influenced by other motivational (e.g., ability beliefs) and cognitive variables (e.g., previous achievement; Wigfield & Eccles, 2000; see also Section 1.2).

Furthermore, in expectancy-value theory and in the theory of vocational interests, the importance of psychological and social factors is emphasized. These influence vocational interests as the most proximal determinants of career choices. As described in Section 1.2.2, Eccles et al. formulated the influence of a wide range of variables on expectancy and value beliefs in a complex model. Holland (1997) also assumed that the active interaction of individuals with their environment leads to a specific disposition later on, and this guides the individual's thinking, perceptions, and actions, such as choosing a specific career instead of another. Such associations, however, are less explicitly formulated in Holland's theory of vocational interest than in the expectancy-value theory.

However, differences occur with respect to the domain-specificity of the constructs and the context that they are typically applied to. As described in Section 1.2, expectancy-value theory includes the influences of different domain-specific components (i.e., expectancies for success and different subjective task values) as predictors of academic decision making, which furthermore leads to specific educational outcomes. Expectancies and subjective task values are traditionally considered at a domain-specific level, such as expectancy and value beliefs in mathematics, which have been linked to careers in mathematically intensive STEM disciplines (see Wigfield et al., 2015; and see also Section 1.3). In addition, intraindividual hierarchies of

expectancy and value beliefs in different domains also predict students' development of expectancy and value beliefs as well as the extent to which these beliefs are associated with achievement and educational choices (e.g., Chow & Salmela-Aro, 2014; Lauermann et al., 2015; see also Section 1.3).

Unlike Eccles et al. (1983), whose primary focus was on several components in one domain, Holland (1997) focused on interests in different domains to explain a person's preference for specific careers. Holland (1997) assumed that the six vocational interest orientations represent preferences for activities across the whole range of different occupations. Thereby, vocational interests typically refer to tasks and activities on a more general level than expectancy-value constructs do. Rather than incorporating different (school) subjects such as mathematics, they capture a broad set of activities and preferences, which can be associated with typical activities in various occupations and fields of professions, such as realistic and investigative interests in the STEM area (Holland, 1997). Therefore, vocational interests have often been used to predict career choices in a more comprehensive manner, such as major choices across the whole range of different subjects (Humphreys & Yao, 2002; Su & Rounds, 2015). Accordingly, the six interest orientations have been shown to be strong predictors of STEM career choices compared with choices of other majors (Humphreys & Yao, 2002; Päßler & Hell, 2012), with realistic and investigative interests being especially highly related to the choice of STEM careers (Nagy, 2006; Rolffhus & Ackerman, 1996; Roloff Henoch, Klusmann, Lüdtkke, & Trautwein, 2015).

There is only a little evidence for associations between expectancy and value beliefs and vocational interests. There is some work on associations between vocational interests and self-efficacy, although such studies have usually employed very general measures of self-efficacy (e.g., Armstrong & Vogel, 2010; Lent, Lopez, Lopez, & Sheu, 2008; Patrick, Care, & Ainley, 2011). Furthermore, there is some recent research on associations between domain-specific self-concept and vocational interests in which positive correlations between math self-concept and realistic and investigative interests have been reported (Volodina & Nagy, 2016).

### **1.5.3 Gender differences in vocational interests and associations with STEM careers**

There are meaningful gender differences in the STEM-related vocational interest orientations (e.g., Päßler & Hell, 2012; Su et al., 2009), such as in math expectancy and value beliefs (see Section 1.3). In a meta-analysis of 81 samples involving a total of 503,188 individuals, Su et al. (2009) found higher realistic ( $d = 0.84$ ) and investigative interests ( $d = 0.26$ ) for males compared with females. Conversely, females showed higher social interests ( $d$

= 0.68) than males in this meta-analysis. Gender differences in STEM-related vocational interests are therefore even more pronounced than gender differences in math expectancy and value beliefs (see Section 1.3). Also, in comparison with gender differences in other psychological characteristics, gender differences in realistic and social interests seem to be relatively large, as such effect sizes typically range from  $d = 0.00$  to  $0.35$  (see Hyde, 2005). Furthermore, there is ample evidence that gender differences in vocational interests are linked to gender differences in mathematically intensive STEM careers (Hackett, Betz, Casas, & Rocha-Singh, 1992; Lapan, Shaughnessy, & Boggs, 1996; Päßler & Hell, 2012; Su & Rounds, 2015), such as a link between gender differences in expectancy and value beliefs and gendered STEM outcomes (see Section 1.3.2).

As in research on expectancy-value constructs (see Section 1.3.2), gender has also been investigated as a moderator of the associations between vocational interests and STEM careers (Päßler, Beinicke, & Hell, 2014; Päßler & Hell, 2012), although there are fewer studies on vocational interests. Nevertheless, there are some indications that gender moderates the association between vocational interests and career choices. Päßler and Hell (2012), for instance, found stronger positive associations between realistic interests and engineering majors for women than for men, whereas social interests were negatively associated with the choice of science majors only for women (for further support on differential associations between vocational interests and STEM careers, see also Larson, Wu, Bailey, Borgen, & Gasser, 2010; Päßler et al., 2014).

To summarize Section 1.5, in order to understand the complex process of gendered STEM career choices, it is essential to consider associations between expectancy and value beliefs with other important predictors of STEM careers in the area of motivation and interest. Holland's (1997) theory of vocational interests is one of the most widely applied and studied frameworks for understanding young people's career choices in the area of work and vocational psychology. On the basis of theoretical considerations, there are important similarities and differences between expectancy and value beliefs and vocational interests as well as their associations with mathematically intensive STEM careers. However, such associations have not been tested empirically, and it is thus far unclear whether and how expectancy and value beliefs and vocational interests differentially explain gender differences in mathematically intensive STEM fields.

## 1.6 Research Questions of the Present Dissertation

In the present dissertation, the influence of individual and environmental factors on gender differences in mathematically intensive STEM careers was investigated. Given the complexity of career processes and gender differences in such pathways, only selected factors that are considered to be particularly critical for gendered career choices were explored. Using Eccles et al.'s (1983) expectancy-value theory as a guiding framework, the effects of two environmental factors that have the potential to reduce women's participation in mathematically intensive STEM careers were investigated, namely, gender-math stereotypes in the media (i.e., a public television program) and high school coursework requirements. Furthermore, expectancy and value beliefs, which are the most proximal constructs that are associated with STEM careers in Eccles et al.'s (1983) expectancy-value theory, were compared with prominent constructs in the area of work and vocational psychology (i.e., vocational interests) in terms of their differential associations with gendered STEM careers.

As outlined in the theoretical introduction, Eccles et al.'s (1983) expectancy-value theory is a powerful framework from which to investigate gender differences in mathematically intensive STEM disciplines. This theory provides a comprehensive model of how individual and contextual factors contribute to gendered achievement-related outcomes and choices. The key assumptions of the theory as well as the relevance of the constructs for career choices have been supported empirically. Expectancy and value beliefs have been shown to be highly predictive of career choices and academic achievement as well as gender differences in these areas (see Section 1.3). Moreover, a variety of factors that influence the development of gender differences in expectancy and value beliefs have also been identified, such as the gender stereotypes children and adolescents face in their environment or opportunities to self-select themselves out of the STEM pipeline throughout school (see Section 1.4). Furthermore, researchers have provided evidence for how expectancy and value beliefs are related to constructs of other motivation- and interest-based theories that are important for gendered career choices in order to broaden the understanding of the role of expectancy and value beliefs for specific gendered STEM outcomes (see Section 1.5). Nevertheless, important questions still remain unanswered, in particular, questions with respect to environmental influences on the development of gender differences in expectancy and value beliefs besides family and with respect to the role of expectancy and value beliefs in determining gendered career choices compared with other important motivation- and interest-based constructs outside the area of educational psychology.

In the present dissertation, some of these questions were addressed in order to extend prior evidence on the influence of gender stereotypes and high school coursework requirements on gender differences in mathematically intensive STEM careers as well as the differential associations of expectancy-value constructs and vocational interests with important indicators of STEM careers. Three empirical studies were conducted. They have the following characteristics in common. First, in the present dissertation, samples from specific age groups were considered so that critical phases in the development of gender differences in motivation as well as the transition into STEM careers at university could be investigated, including childhood, adolescence, and young adulthood. In doing so, the present dissertation followed the recommendations of Ceci et al. (2014), who called for a life-course perspective in investigating gender differences with a particular focus on the time period before students actually enter professions, as career aspirations develop and manifest before individuals actually enter different career pathways such as different university majors (see also Schoon & Eccles, 2014).

Second, the influence of environmental factors on the development of gender differences in mathematically intensive STEM careers was investigated in real-world situations. As described in Section 1.2, one strength of research based on the expectancy-value framework is its focus on educational outcomes in the real world. Thus, there was ample research on which the present dissertation could build. Nevertheless, as described in Section 1.4, there is limited research on how environmental factors influence gender differences in motivation in real-world situations, in particular, in terms of gender stereotypes, which have often been studied in laboratory settings. Research involving high school coursework requirements is also lacking because there is limited real-world data on such effects. Therefore, the present dissertation was based on but also designed to extend previous research in the expectancy-value framework.

Third, multiple indicators of students' academic outcomes were included to comprehensively investigate effects of environmental factors on gender differences in motivation, including various motivational constructs as well as achievement measures. In doing so, a domain-specific approach was chosen with the primary focus on motivation in mathematics as this domain is an important prerequisite for STEM careers as described in Section 1.3, and in contrast to the other science disciplines, it is also one of the core subjects in school (OECD, 2013). Also, different STEM university majors need to be considered when associations between motivation- and interest-based constructs with such majors are investigated, so this dissertation explored mathematically intensive STEM majors and majors in the area of life sciences. As stated by Wang and Degol (2013), for instance, such distinctions within the STEM fields are important so that a more nuanced and precise understanding of



gender differences in fields such as STEM majors can be developed, and thus, it is important to include different majors.

Building on these general features, three empirical studies were conducted in this dissertation with the following specific goals and features. Study 1 (*Math = Male and Geeky: Do Gender Stereotypes in a Television Program Affect Girls' and Boys' Math Performance and Motivation?*) was designed to investigate gender stereotypes embedded in a television program as one important environmental factor that might influence gender differences in math motivation. As previously stated, little is known about how gender stereotypes encountered in the environment contribute to gender differences in mathematically intensive STEM fields. Accordingly, a randomized study with a total of 335 female and male students in the fifth grade was conducted to address this question. In order to reflect what children encounter in real-world situations, a television program that was broadcast on a national TV channel was chosen as the experimental material. The chosen program was designed to show children that math could be interesting and fun and included a section with gender-math stereotypes. To obtain a comprehensive picture of possible effects of the gender stereotypes presented in this television program, several outcome measures were assessed, including children's stereotype endorsement, their task values, self-concept, as well as their math achievement. Research predictions were preregistered before the experiment was conducted to increase research transparency as this has been identified as a problem in previous research on effects of stereotypes (see [https://osf.io/8f7y6/?view\\_only=d85b73e70f5040b5a54fcf03091811f1](https://osf.io/8f7y6/?view_only=d85b73e70f5040b5a54fcf03091811f1)).

Study 2 (*Maximizing Gender Equality by Minimizing Course Choice Options? Effects of Obligatory Coursework in Math on Gender Differences in STEM*) was designed to examine the influence of high school coursework requirements on gender differences in mathematically intensive STEM fields. As previously described, female students are less likely than male students to choose advanced math courses in high school, and such differences have been linked to gender differences in STEM careers. Study 2 therefore investigated whether encouraging young women to choose advanced math courses in high school would bring more women to choose STEM careers. To this end, effects of a statewide educational reform in Germany that required all students to take advanced math courses were investigated, and data from a large school achievement study (TOSCA; see Köller, Watermann, Trautwein, & Lüdtke, 2004; Trautwein, Neumann, Nagy, Lüdtke, & Maaz, 2010) were reanalyzed. Data from 4,730 students who participated in high school courses before the reform were compared with data from 4,715 students who participated in high school courses after the reform. Gender differences in math motivation and achievement in both cohorts were compared when students had a mean age of

19 years and were in their final year of high school (i.e., in 2002 for the first cohort and in 2006 for the second cohort). Gender differences in the choice of STEM majors in both cohorts were compared at age 21 when students had actually chosen different university majors.

Study 3 (*It Takes Two: Expectancy-Value Constructs and Vocational Interests Predict STEM Careers Differently and Differ Between Men and Women*) was designed to examine the relative predictive power of expectancy-value constructs and vocational interests for STEM careers for men and women. Next to Eccles et al.'s (1983) expectancy-value theory, Holland's (1997) theory of vocational interests is one of the most widely used frameworks for investigating gender differences in STEM careers. However, only a little is known about how expectancy-value constructs and vocational interests are differentially associated with STEM careers. Study 3 was therefore conducted with the aim to contribute to a better understanding of the complex process of gendered career choices, which likely involve a variety of different factors and constructs. To this end, the predictive validity of expectancy-value constructs in mathematics and English and vocational interests for math achievement and the choice of different STEM majors was simultaneously investigated. In doing so, the role of gender in such predictions was explored along with the following questions: How do gender differences in expectancy-value constructs and vocational interests explain gender differences in math achievement and the choice of (mathematically intensive) STEM university majors? Does the relative predictive power of expectancy-value constructs and vocational interests vary between women and men? As in Study 2, data from the TOSCA study (Trautwein et al., 2010) were used. For Study 3, however, only data from the second cohort were analyzed because subjective task values were assessed only for this cohort. The sample thus consisted of 4,984 female and male students who were in their final year of high school at the first measurement point. Data for the second measurement point were collected 2 years later when participants were at university.

Study 1:

Math = Male and Geeky:

Do Gender Stereotypes in a Television Program  
Affect Girls' and Boys' Math Performance and  
Motivation?

Wille, E., Gaspard, H., Trautwein, U., Oschatz, K., Scheiter, K., & Nagengast, B. (2017).  
Math = male and geeky: Do gender stereotypes in a television program affect girls' and boys'  
math performance and motivation? *Manuscript submitted for publication.*

This research was funded by the LEAD Graduate School & Research Network [GSC1028], a  
project of the Excellence Initiative of the German federal and state governments.

**The following manuscript has not been accepted and published yet. This version of the  
manuscript has been submitted to the *Journal of Educational Psychology* on July 7, 2017.  
*Journal of Educational Psychology* will have the copyright of the final version of the  
article, if it will be accepted. The version displayed here may not exactly replicate the  
final version published in the journal. It is not the copy of record.**

### Abstract

Television programs are a central part of children's everyday lives. These programs often transmit stereotypes about gender roles, such as "math is not for girls" or "boys who are good at math are geeky", but so far, it is unclear whether such stereotypes embedded in television programs affect girls' and boys' motivation and performance. On the basis of research on expectancy-value theory and stereotype threat, we conducted a randomized controlled study with a total of 335 fifth-grade students to address this question. We used a television program that was broadcast on a national TV channel as the experimental material. We investigated effects on boys' and girls' stereotype endorsement, task values, self-concept, and sense of belonging in math as well as their feelings about math and their math performance. We preregistered our research predictions and analyses before conducting the experiment. Watching the stereotypes embedded in the video increased boys' and girls' stereotype endorsement. Furthermore, boys reported a higher sense of belonging but lower utility value after watching the video with the stereotypes. However, boys' other outcome variables were not affected, and there were also no effects on performance, task values, self-concept, and sense of belonging for girls. Thus, our results provide only partial support for short-term effects of gender stereotypes embedded in television programs on children's math performance and motivation. We discuss the theoretical and practical implications of the findings.

*Keywords:* stereotypes, gender differences, television, math motivation, math sense of belonging

## **Math = Male and Geeky: Do Gender Stereotypes in a Television Program Affect Girls' and Boys' Math Performance and Motivation?**

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

Television programs are one potential source of gender stereotypes for children, as television broadcasts are central components of children's socialization (Rideout, 2015; Rideout, Foehr, & Roberts, 2010). Television shows and programs with STEM content have increased in availability (National Research Council, 2009) and popularity (Kirsch, 2011; Patten, 2013). They also transmit certain beliefs and stereotypes about gender roles in the STEM field, such as showing females as underperforming in math and science or presenting scientists as nerdy (see e.g., Collins, 2011; Heyman, 2008). It is not yet clear whether stereotypes in television programs affect girls' and boys' motivation and performance in math. So far, research on expectancy-value theory has focused primarily on the role of stereotypes presented by parents, teachers, or peers (see Wigfield et al., 2015), and research on stereotype threat has traditionally investigated effects of stereotypes presented as isolated stimuli in laboratory settings (see Spencer et al., 2016).

In the present study, we aimed to contribute to closing this gap in the literature by examining effects of traditional gender stereotypes in a math television program for children that was broadcast on a German national TV channel. Specifically, the end of this program showed girls as not performing well in math and (male) mathematicians as geeky. To examine the effects of these stereotypes, we conducted an experimental study in which fifth graders watched this television program about math either with or without the section that showed these

gender stereotypes. We studied effects on both girls' and boys' stereotype endorsement as well as their motivation and performance in math.

### ***Gender Differences in Math Motivation and Achievement***

Eccles et al.'s (1983) expectancy-value theory is one of the most widely used frameworks for investigating gender differences in math motivation and achievement and has been highly effective in explaining women's underrepresentation in the STEM fields (Schoon & Eccles, 2014; Watt & Eccles, 2008). Eccles et al. (1983) suggested that the expectation of success in a specific domain as well as several aspects of subjective task values would predict academic decision making and thereby also specific educational outcomes, such as later achievement or educational choices. Young people should thus choose math-intensive STEM careers if they expect to be good at math and science activities and have high values in these domains.

There is ample research on gender differences in math motivation that has consistently indicated that girls are less motivated in this domain than boys (for a review, see Wigfield et al., 2015). Such research has also shown that these gender differences emerge at an early age (Fredricks & Eccles, 2002; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002) and apply to various motivational outcomes, including math self-concept (Marsh & Yeung, 1998; Nagy et al., 2010), sense of belonging in math (C. Good, Rattan, & Dweck, 2012), and task values (Gaspard et al., 2015).

On the basis of these motivational differences, one might expect boys to also perform better than girls because motivation is an important prerequisite for high performance. However, meta-analyses investigating gender differences in math achievement have shown rather small advantages for boys compared with girls (e.g., Else-Quest, Hyde, & Linn, 2010; Reilly, Neumann, & Andrews, 2015). Moreover, these analyses have indicated that such gender differences seem to occur only on math achievement tests (Reilly et al., 2015), whereas girls even show an advantage in teacher-assigned school marks (Voyer & Voyer, 2014).

### ***The Role of Stereotypes in Children's Socialization***

According to expectancy-value theory, socializers' beliefs and behaviors as well as cultural milieu influence individuals' task perceptions and interpretations of previous academic achievement (Eccles et al., 1983). In explaining gender differences in motivation and achievement, expectancy-value theory thus indicates that boys and girls have different socialization processes, which are shaped by the surrounding environment and its gender norms and roles, the individuals' beliefs, and the choices females and males make on the basis of their

socialization (Eccles, 2009; Schoon & Eccles, 2014; Wigfield & Eccles, 2000). In particular, gendered socialization refers to specific gender roles or the gender-stereotypical attitudes and expectancies of parents, teachers, and other socializing influences such as the media, all of which transmit gender stereotypes (Wigfield et al., 2015). Stereotypes can be broadly defined as associations of group members with specific attributes (Greenwald et al., 2002). Regarding gender, there are specific stereotypes about the traits, abilities, and motivation of males and females, specifically in the domain of math (see Leaper, 2015). Math and science are male-typed domains, and gender stereotypes in these domains include assumptions about lower abilities and less talent in math for females compared with males (e.g., Spencer, Steele, & Quinn, 1999; Wigfield & Eccles, 2002). According to expectancy-value theory, as a result of the gender stereotypes children face in their socialization, boys develop higher competence beliefs and values in male-typed domains such as math and math-intensive STEM domains, whereas girls develop higher competence beliefs and values in female-typed domains such as languages and arts (e.g., Wigfield et al., 2015). It is assumed that such gender differences in math competence beliefs and values may lead to gender differences in math achievement in the long run (Wigfield & Eccles, 2000). Previous studies have supported this idea by showing that women's gender stereotypes reduced their future expectancies of success (Smith, Brown, Thoman, & Deemer, 2015) and their future task values (Plante, De la Sablonnière, Aronson, & Théorêt, 2013; Smith et al., 2015). Expectancy and task values, in turn, have been shown to be important predictors of later achievement (e.g., Denissen, Zarrett, & Eccles, 2007; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005).

### *Short-Term Effects of Stereotypes*

The repeated experience of stereotypes is one potential mechanism that may explain how stereotypes of others can influence girls' and boys' math motivation and achievement as well as their attitudes toward math. According to expectancy-value theory, such experiences might lead to the internalization of gender-role stereotypes. As a consequence, girls devalue math and disidentify with the subject in the long run, whereas boys may especially value math (Eccles et al., 1983; Wigfield et al., 2015). Research on stereotype threat has provided support for this idea by showing that the activation of traditional gender stereotypes—referred to as stereotype threat—reduces women's and girls' motivation and performance in the short term (for a review, see Spencer, Logel, & Davies, 2016; C. M. Steele et al., 2002). The concept of stereotype threat originated from Steele and Aronson's work in the 1990s. They defined stereotype threat as a situational experience in which group members feel concerned about

confirming a negative stereotype of their own group (C. M. Steele, 1997; C. M. Steele & Aronson, 1995). They suggested that such concerns might compromise a person's behavior and performance.

Originally, research on stereotype threat focused on explaining the underperformance of African Americans in performance (C. M. Steele & Aronson, 1995), but ample research has also been conducted to address gender differences in math-intensive domains (e.g., Schmader, 2002; Spencer et al., 1999; Tomasetto, Alparone, & Cadinu, 2011). Such research has demonstrated that women show lower math performance if they are reminded of negative stereotypes about women in math, but they perform equally as well as males if such stereotypes are not made salient before they take a math test (Doyle & Voyer, 2016; Nguyen & Ryan, 2008). Effects of stereotype threat have also been shown for motivational outcomes, such as women's evaluation of their math performance (Cadinu, Maass, Frigerio, Impagliazzo, & Latinotti, 2003), women's sense of belonging in math and science and their domain identification (e.g., Cheryan, Plaut, Davies, & Steele, 2009; see also Thoman, Smith, Brown, Chase, & Lee, 2013, for a review), how much they would like to engage in math-related activities (J. R. Steele & Ambady, 2006), their motivation to improve in math (Fogliati & Bussey, 2013), and their interest (Smith, Sansone, & White, 2007; see also Thoman et al., 2013, for a review).

Effects of the stereotype of the male advantage in math on men's and boys' achievement and motivation are less clear. In the literature on stereotype threat, a meta-analysis by Walton and Cohen (2003) indicated positive effects of traditional gender stereotypes for men on math performance, whereas a more recent meta-analysis by Doyle and Voyer (2016) indicated no effects. No effects of traditional gender stereotypes for men have also been reported with respect to their interest and belonging in computer science (Cheryan et al., 2009) or their motivation to improve in math (Fogliati & Bussey, 2013). In addition, there is research on the perceived fit between the stereotype of math and science and students' self-image. Such research has indicated that favoring these subjects reduces students' social competence and belonging. A study by Hannover and Kessels (2004) showed that students who favored science were judged as less popular, less attractive, less socially competent, and less integrated than students who did not like science.

### ***Effects of Stereotypes on Children***

Although most research on gender stereotypes has been conducted on college students or older adults, multiple studies have also been conducted on children or adolescents (e.g., Ambady, Shih, Kim, & Pittinsky, 2001; Flore & Wicherts, 2015). These studies have



demonstrated that children in elementary school are already aware of their own gender and show gender-stereotypical views in the domain of math as they attribute lower math ability and talent to girls and women than to boys and men (e.g., Ambady et al., 2001; Passolunghi, Rueda Ferreira, & Tomasetto, 2014; Signorella, Bigler, & Liben, 1993). There is also research on short-term effects of stereotypes on children with respect to math performance (Ambady et al., 2001; Muzzatti & Agnoli, 2007; Neuville & Croizet, 2007; Tomasetto, Alparone, & Cadinu, 2011). A meta-analysis by Flore and Wicherts (2015), for instance, indicated that girls who are reminded of typical gender stereotypes in math show slightly lower math performance compared with girls who were not reminded of such stereotypes. Such effects were consistently found for girls younger than 13 years old or even younger than 8 years old. However, the authors also suggested that the results might be questionable due to indications that publication bias may have occurred.

Less research has investigated the effects of stereotypes on students' motivation and their attitudes toward STEM. In two studies, Master et al. (2015) examined the effects of stereotypes on adolescents' sense of belonging in STEM courses and their interest in a sample of 15-year-old high school students. Their findings showed that when stereotypes in the field of computer science were made salient in the classroom environment of a high school computer course, female adolescents' sense of belonging and their interest in enrolling in computer courses were reduced in comparison with female adolescents in a control group. In both studies, male adolescents' sense of belonging and interest in enrolling in computer courses were not affected. Furthermore, Muzzatti and Agnoli (2007) investigated the effects of stereotypes on children's evaluation of their own math performance in different age groups. Their results indicated no effects for children in Grades 3 and 5, but students in the stereotype condition in Grade 8 showed lower self-confidence in their math ability than students in the control group. Nevertheless, these two studies provide only initial insights into the effects of stereotypes on motivational outcomes in children.

### *Effects of Stereotypes Presented in the Media*

Research on expectancy-value theory has focused primarily on the influence of parents, teachers, or peers on children's endorsement of stereotypes and their motivation (see Wigfield et al., 2015), but research on stereotype threat has indicated a wide range of situations in which stereotypes about females' underperformance in math can affect girls and women, at least in laboratory settings. These include examples of how stereotypes can be induced by the media, such as newspaper articles (Cheryan, Plaut, Handron, & Hudson, 2013), images in schoolbooks

(J. J. Good, Woodzicka, & Wingfield, 2010), and photographs (Muzzatti & Agnoli, 2007). There are also initial findings with respect to videos and television advertising (Davies, Spencer, Quinn, & Gerhardstein, 2002; Murphy, Steele, & Gross, 2007). Murphy et al. (2007) investigated the effects of reminding women of their underrepresentation in math-intensive STEM fields via video by presenting groups of people working in the STEM fields. Women who watched a video in which the male-female ratio reflected the proportion of women in these fields showed a lower sense of belonging, a lower intention to participate in a STEM-related conference, and remembered the contents of the video less correctly, compared with women who watched a video with a gender-balanced proportion.

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

The findings by Murphy et al. (2007) and Davies et al. (2002) indicate that stereotypes in videos can have negative effects on women compared with men. However, these findings provide only initial insights into the effects of television for adults. So far, it is unclear if television programs can affect children in similar ways. Furthermore, both studies investigated stereotypes that were presented in isolated situations. Thus, they were not able to provide insights into how stereotypes might affect children when experienced in their daily lives in more complex situations, for instance, as one part of a whole television program.

### ***The Present Study***

In the present study, we investigated effects of gender stereotypes in a STEM television program on girls' and boys' math motivation and performance. Despite the importance of television programs in children's everyday lives and the relevance of such programs for children's informal science learning, there is a lack of research on how girls' and boys' reception of STEM television programs might be affected by presentations of traditional gender stereotypes in such programs. Research on expectancy-value theory and stereotype threat has

provided initial insights into how stereotypes might affect children. However, research on expectancy-value theory has mainly focused on the role of stereotypes presented by parents, teachers, or peers (see Wigfield et al., 2015), and research on stereotype threat has traditionally investigated effects of stereotypes presented as isolated stimuli in laboratory settings (see Spencer et al., 2016). Furthermore, there are indications of publication bias in the stereotype threat literature (Flore & Wicherts, 2015). Accordingly, it is unclear whether and how stereotypes embedded in children's daily activities such as in a television program might affect girls and boys.

Therefore, we conducted an experiment in which fifth-grade students watched a children's television program about math that either contained or did not contain a clip in which traditional gender stereotypes were made salient. In order to link the study as closely as possible to what children are likely to watch in their everyday lives, we used a television program that was broadcast on a national TV channel in Germany as the experimental material. The chosen program was designed to show children that math could be interesting and fun and included a section with stereotypes in which girls were annoyed about math and a boy was good at math but geeky.

We preregistered our research predictions before conducting the experiment ([https://osf.io/8f7y6/?view\\_only=d85b73e70f5040b5a54fcf03091811f1](https://osf.io/8f7y6/?view_only=d85b73e70f5040b5a54fcf03091811f1)). Based on previous research, the primary focus of our study was to examine effects of gender stereotypes in the television program on girls' math performance and motivation. We investigated math self-concept, sense of belonging in math, math task values, and attitude toward math as motivational outcomes in order to obtain a comprehensive picture of possible effects on math motivation. On the basis of existing literature on effects of stereotypes on math performance (Flore & Wicherts, 2015), self-concept (Cadinu et al., 2003; Muzzatti & Agnoli, 2007), and sense of belonging (Master et al., 2015), we expected that girls who watched the gender stereotyped television program would show lower math performance, lower math self-concept, and a lower sense of belonging in math compared with girls in the control condition.

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

effects on task values and feelings about math but rather investigated possible effects of the gender-stereotyped television program on these outcomes for girls as open research questions.

We explored effects on boys' math performance and motivation as well, using the same outcomes measures. Due to the mixed findings from previous research on the effects of stereotypes on males' performance and motivation, we did not hypothesize specific effects for boys but rather investigated possible effects on these outcomes for boys as exploratory research questions.

We additionally assessed the endorsement of gender stereotypes for both girls and boys. We also did not formulate any specific hypotheses with respect to this outcome because previous research has provided mixed results on effects of gender stereotypes on children's endorsement of gender stereotypes (Ambady et al., 2001; Schmader, Johns, & Barquissau, 2004; Steffens, Jelenec, & Noack, 2010).

## **Method**

### ***Participants***

Participants were 335 fifth-grade students. Children were recruited from 18 classes of four academic track schools in Baden-Württemberg, Germany. The sample size was based on a power analysis for a randomized block trial with the treatment implemented at the student level using Optimal Design (Raudenbush et al., 2011). We calculated the required number of classrooms by aiming to achieve an acceptable level of power ( $\beta = .80$ ) to detect intervention effects of  $\delta = 0.40$  when comparing the experimental with the control condition. We assumed that 10 girls and 10 boys would participate in each class, and they would be randomly assigned to the control and experimental conditions. We furthermore assumed an effect size variability of 0.10 (for more details, see the preregistration protocol at [https://osf.io/8f7y6/?view\\_only=d85b73e70f5040b5a54fcf03091811f1](https://osf.io/8f7y6/?view_only=d85b73e70f5040b5a54fcf03091811f1)).

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

### ***Design and Procedure***

As preregistered, we collected the data using a pretest-posttest design, and we applied a randomized block design to examine effects of gender stereotypes in a television program. Girls and boys were randomly assigned to the experimental and control groups within each class

(experimental group:  $N = 87$  girls and  $N = 85$  boys; control group:  $N = 76$  girls and  $N = 87$  boys). Participants were tested in one classroom simultaneously, but every student watched the video separately on an iPad. Headphones were used so that students would not be distracted by the other participating students around them. We collected the pretest data 1 week before the experimental manipulation and the posttest data directly after the experimental manipulation. Because research on stereotype threat has shown that even small and short manipulations can influence students' performance and motivation (e.g., Master et al., 2015), we aimed to balance any effects of the order in which the instruments were presented on students' outcomes. That is, the achievement test might affect students' attitudes and motivation if assessed first, or the questionnaire might wash out any effects of the video on the achievement test if assessed first. To examine such order effects of the instruments, the order of the achievement test and the questionnaire was balanced on the class level in both phases of data collection. We randomly assigned the classes to the two conditions ( $N = 9$  classes in each condition). Data were collected in June and July 2016 by trained research assistants during school hours.

### *Experimental Manipulation*

As experimental material, we used one episode from a German children's television program, which was broadcast on a German national television channel in June 2015. The episode focused on math and was designed to show children that math could be interesting and fun even though it might be experienced as boring in school (KiKa.de, 2015). The episode had a total duration of 23 min. As preregistered, only 15 min of the episode were used in the present study due to time constraints. This included an introduction by a male television presenter (about one minute) and two different math tasks solved by fifth-grade children (about 13 minutes). In addition, the video included a clip that implied traditional gender stereotypes in math (about one minute). This part showed two girls who were very bored with math and annoyed that they had to do math homework. Instead of doing their homework, one girl copied it from a classmate, the "math nerd," and in exchange, she promised him that her friend would accompany him to the movies. Her friend was horrified about going out with this boy. The "math nerd" wore very large glasses and a shirt that was completely buttoned up, including the uppermost button. He additionally wore suit trousers and suspenders. He was holding a leather briefcase in front of his chest.

The introduction and the math tasks solved by the children were used in both conditions. The experimental manipulation depended on only the last minute of the video. In the experimental condition, participants watched the clip about the two girls who were annoyed

about their math homework and copied it from the “math nerd.” In the control condition, participants watched a neutral summary of the first 14 min of the video. The summary was comparable in length so that the total length of the video would be held constant between the groups. Consequently, participants experienced the stereotypes as a short section within the whole television program so that the ecological validity of the experiment would be high.

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

### *Instruments*

#### *Math performance*

We assessed students’ math performance with a speed test that consisted of three sections containing basic tasks involving addition, subtraction, and multiplication (basic competence test; Lambert, Dackermann, & Möller, 2017). Each part consisted of 36 tasks, and for each individual part, we asked the students to solve as many tasks as possible within 2 min. The sum score of all three parts, generated by computing the sum of correctly solved items, was used in the analyses. The test showed high internal consistency (Kuder-Richardson 20 = .93/.94 for the pretest/posttest).

#### *Questionnaire*

We assessed the following variables with a questionnaire. Unless otherwise noted, all items on the questionnaire were measured with a 4-point Likert scale ranging from 1 (*completely disagree*) to 4 (*completely agree*). The questionnaire is available at [https://osf.io/8f7y6/?view\\_only=d85b73e70f5040b5a54fcf03091811f1](https://osf.io/8f7y6/?view_only=d85b73e70f5040b5a54fcf03091811f1).

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

*Sense of Belonging.* We assessed students’ sense of belonging in math with 10 items (e.g., “I feel like a real part of my class in math”), based on the Psychological Sense of School Membership (PSSM; Goodenow, 1993). The items were translated into German and adapted to math class instead of school membership. Due to low item-scale correlations ( $r_{it} = .03/.16$  for the pretest/posttest), we excluded one item when we computed the scale. The final scale therefore consisted of nine items and showed an acceptable internal consistency ( $\alpha = .76/.84$

for the pretest/posttest). Because we did not preregister the exclusion of the item, we conducted the analysis for this outcome also using the original scale, which included all 10 items. The internal consistency for this scale was acceptable ( $\alpha = .73/.83$  for the pretest/posttest), and the results did not differ meaningfully from those computed with the reduced scale (see the Supplemental Material).

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

*Task Values.* We assessed students' value beliefs in math with a previously developed questionnaire (Gaspard et al., 2015). The items covered all four conceptual dimensions of task values as specified in the expectancy-value model (Wigfield & Eccles, 2000). Intrinsic value (e.g., "I like doing math";  $\alpha = .92/.94$  for the pretest/posttest), attainment value (e.g., "It is important to me to be good at math"; four items;  $\alpha = .87/.93$  for the pretest/posttest), and cost (emotional costs, e.g., "Studying math makes me quite nervous";  $\alpha = .78/.86$  for the pretest/posttest) were assessed with four items each. For utility value, we differentiated between two facets: utility for daily life (e.g., "Knowing about the subject of math brings me many advantages in my daily life";  $\alpha = .82/.84$  for the pretest/posttest) and social utility (e.g., "Sound knowledge in math counts for something with my classmates";  $\alpha = .68/.80$  for the pretest/posttest), which were both assessed with three items.

*Stereotype Endorsement.* We assessed stereotype endorsement with three items based on items from Schmader, Johns, and Barquissau (2004). We adapted the items for children by using "boys" and "girls" in the wording instead of "men" and "women" (e.g., "Boys have higher math abilities than girls";  $\alpha = .76/.76$  for the pretest/posttest).

We extended the scale by including two items in which the words "boys" and "girls" were interchanged (e.g., "Girls have more math abilities than boys") and preregistered this extension. We recoded these items before computing the scale score. However, the reliability of the extended scale was rather low ( $\alpha = .52/.55$  for the pretest/posttest). Consequently, we used only the original scale in our analyses.

*Manipulation Check.* As a manipulation check, we asked the students what they had seen in the last minute of the video, that is, two girls who copied their homework from the “math nerd” or a summary of the video. Thirteen students did not answer the question correctly. In our analyses, we conducted an intention-to-treat analysis by taking only the original assignment into account, but as robustness checks, we excluded the students in the experimental condition who failed the manipulation check. The results did not differ meaningfully (see the Supplemental Material).

*Additional Scales.* As preregistered, we also assessed stereotype endorsement with measures based on studies by Ambady, Shih, Kim, and Pittinsky (2001) and Steffens, Jelenec, and Noack (2010) in which the participants were asked how much they would like to engage in activities related to math and German (see [https://osf.io/8f7y6/?view\\_only=d85b73e70f5040b5a54fcf03091811f1](https://osf.io/8f7y6/?view_only=d85b73e70f5040b5a54fcf03091811f1)). Due to high rates of missing data and the low reliability of these scales, we refrained from conducting additional analyses on these instruments.

We furthermore preregistered exploratory analyses with respect to the same set of motivational outcomes in the domain of German. Dimensional comparisons of complementary domains are important in the development of students’ motivation (Möller & Marsh, 2013), and there are initial findings on how motivation in a verbal domain might be affected by traditional gender stereotypes in commercials (Davies et al., 2002). Due to space limitations, the results on girls’ and boys’ motivation in German are reported in the Supplemental Material. In summary, we found no effects of the experimental condition on girls’ and boys’ motivation in German, except for a negative effect on cost in German for girls.

### *Statistical Analyses*

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

To test whether there were any order effects of the instruments, we computed multiple-group regression analyses with the order of the instruments as the grouping variable. We tested



the difference between the models for each group with Wald  $\chi^2$  tests. If there were no significant differences between the coefficients in the models, we calculated multiple regressions for the whole sample. Although not explicitly preregistered, we used multiple-group analyses instead of including the order in the respective two-way and three-way interaction terms in the original model because results of multiple-group analyses are easier to interpret.

Because some students were absent when the pretest or posttest was administered and others did not respond to individual scales, missing data ranged from 2.1% to 9.9% for the different scales. To deal with missing data, we used the full information maximum likelihood approach as implemented in Mplus 7.4 (Muthén & Muthén, 2012). This approach takes all available information into account when estimating the model parameters (Graham, 2009).

We considered the clustered structure of the data (students nested in classes) by using the design-based correction of standard errors implemented in Mplus 7.31 (Muthén & Muthén, 2012). The advantage of this approach is the smaller number of assumptions and the easier interpretation of the estimates, as these are identical to single-level models, although the standard errors were adjusted for the clustered structure (McNeish, Stapleton, & Silverman, 2017).

## Results

### *Descriptive Statistics and Randomization Check*

The means and standard deviations for all scales are shown in Table 1 (for girls and boys) and Tables 2 and 3 (for girls and boys in each condition). Girls and boys reported relatively high levels of self-concept, sense of belonging, intrinsic value, attainment value, and utility value for daily life in all conditions. Compared with boys, girls showed significantly lower math performance and reported lower levels of math self-concept, the feeling thermometer, math intrinsic value, and social utility value in math on the pretest. The correlations for the outcome variables at the two measurement points are presented in Table 4. These indicate that the mean levels were relatively stable across the two measurement points for all outcomes ( $.60 < r < .87$ ). All correlations between the study variables were reasonable in terms of direction and size.

To test whether the randomization of girls and boys in the experimental and control groups had been successful in the baseline measures, we computed multiple regression models as preregistered. Here, we regressed the baseline values for the outcomes on the experimental condition and gender as well as the Condition  $\times$  Gender interaction. There were no significant differences between the experimental and control groups for girls and boys on the pretest values

for all variables (all  $ps > .137$ ) except for the boys with respect to sense of belonging. Here, boys in the experimental group showed lower baseline scores than those in the control group ( $\beta = 0.36, SE = 0.18, 95\% CI [0.07, 0.65], p = .044$ ). As preregistered, we controlled for the pretest scores in all analyses to estimate the effect of the experimental manipulation more precisely because of the explanatory power of this covariate.

Table 1

*Descriptive Statistics for All Study Variables on the Pretest Separated by Gender*

Variable	Girls		Boys		$d^a$	$d$ 95% CI	
	$M$	$SD$	$M$	$SD$			
Stereotype endorsement T1	2.55	0.04	2.73	0.04	0.35	0.20	0.50
Performance T1	51.88	0.61	56.09	0.76	0.48	0.33	0.62
Self-concept T1	3.15	0.07	3.40	0.06	0.37	0.20	0.55
Sense of belonging T1	3.16	0.04	3.19	0.06	0.05	-0.16	0.26
Feeling thermometer T1	1.62	3.65	15.80	3.39	0.42	0.25	0.60
Intrinsic value T1	3.12	0.07	3.27	0.09	0.20	0.03	0.37
Attainment value T1	3.50	0.05	3.46	0.07	-0.06	-0.31	0.20
Utility value – daily life T1	3.24	0.06	3.26	0.05	0.02	-0.15	0.20
Utility value – social T1	2.22	0.05	2.41	0.06	0.29	0.11	0.47
Cost T1	1.60	0.06	1.53	0.05	-0.13	-0.33	0.08

*Note.* CI = confidence interval.

<sup>a</sup>The dependent variable is standardized.

Table 2

*Descriptive Statistics for All Outcome Variables at T1 Separated by Gender and Group*

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

Table 3

*Descriptive Statistics for All Outcome Variables at T2 Separated by Gender and Group*

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

Table 4

*Correlations between all Study Variables*

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.
1. Stereotype endors. T1	—																			
2. Stereotype endors. T2	.38	—																		
3. Performance T1	(.11)	(.09)	—																	
4. Performance T2	.13	.15	.85	—																
5. Self-concept T1	(.09)	(.08)	.36	.35	—															
6. Self-concept T2	.12	.11	.30	.33	.80	—														
7. Sense of belonging T1	(.10)	(.10)	.14	.19	.49	.49	—													
8. Sense of belonging T2	(.05)	(.09)	.12	.17	.39	.50	.81	—												
9. Feeling thermo. T1	(.07)	(.01)	.19	.22	.48	.44	.33	.29	—											
10. Feeling thermo. T2	.11	(.04)	.20	.26	.42	.45	.32	.28	.87	—										
11. Intrinsic value T1	.11	(.09)	.15	.18	.75	.72	.53	.46	.51	.51	—									
12. Intrinsic value T2	(.05)	(.07)	.18	.22	.68	.76	.46	.47	.39	.41	.87	—								
13. Attainment value T1	(.10)	(.07)	(-.05)	(-.00)	.29	.22	.30	.25	.22	.17	.33	.25	—							
14. Attainment value T2	(.06)	(.05)	(-.01)	(.03)	.24	.27	.28	.32	.18	.19	.29	.32	.70	—						
15. Utility v: daily life T1	.15	(.08)	(-.06)	(-.03)	.25	.22	.32	.26	.23	.21	.38	.30	.42	.33	—					
16. Utility v: daily life T2	(.06)	(-.01)	(-.10)	(-.07)	.16	.22	.30	.36	(.11)	.15	.30	.33	.30	.45	.60	—				
17. Utility v: social T1	.16	(.07)	(.10)	.16	.34	.28	.36	.31	.26	.26	.39	.34	.43	.29	.31	.23	—			
18. Utility v: social T2	.13	(.03)	(.09)	(.07)	.24	.27	.25	.28	.21	.24	.35	.33	.34	.28	.23	.24	.70	—		
19. Cost T1	(-.01)	(-.06)	-.16	-.19	-.68	-.68	-.53	-.47	-.42	-.36	-.71	-.68	-.16	-.18	-.23	-.21	-.17	(-.10)	—	
20. Cost T2	(-.01)	(-.05)	-.11	-.19	-.59	-.73	-.45	-.51	-.31	-.31	-.56	-.68	(-.09)	-.23	-.18	-.25	-.15	(-.10)	.72	—

Note. Nonsignificant correlations are displayed in parentheses; for all other correlations,  $p < .05$ .

Stereotype endors. = Stereotype endorsement; Feeling thermo. = Feeling thermometer; Utility v = Utility value.

### *Effects of the Experimental Manipulation*

We investigated the effects of the experimental manipulation by computing multiple regression analyses. First, we tested if there were any order effects of the instruments by computing multiple-group regression analyses using the order of the instruments as the grouping variable. The Wald  $\chi^2$  tests that were used to identify differences in these models were not statistically significant with respect to any of the studied outcomes (all  $ps > .154$ ) except for social utility value. Here, the coefficients for the Condition  $\times$  Gender interaction differed significantly,  $\chi^2(1) = 11.76, p = .001$ . Consequently, we computed multiple regression analyses using the total sample for all outcomes (i.e., averaged across instrument order) except for social utility value (see Tables 5 to 7).

We specified multiple regressions to test effects of the experimental manipulation for girls and boys (see Tables 5 and 6). As girls were coded 0, the main effect of the experimental condition was equal to the simple slope for girls, whereas the Condition  $\times$  Gender interaction term indicated whether the effects differed between boys and girls. Because we were more interested in investigating effects of the experimental manipulation on girls' and boys' performance and motivation rather than on gender differences in these outcomes, we additionally estimated the simple slopes for boys for all outcomes using the model constraint in Mplus. All effects we report are fully standardized effect sizes for the continuous variables.

First, we investigated effects of the experimental condition on stereotype endorsement. However, due to mixed previous results for effects of stereotype threat on this outcome, we did not hypothesize specific effects for girls and boys. The results revealed a significant positive effect of the experimental condition for girls ( $\beta = 0.50, SE = 0.15, 95\% CI [0.03, 0.74], p = .001$ ) when pretest scores were controlled for. The Gender  $\times$  Condition interaction was not statistically significant, indicating that the effects of the experimental condition did not differ between girls and boys. Consistent with this result, the additionally calculated simple slope of the condition for boys was positive and significant ( $\beta = 0.22, SE = 0.11, 95\% CI [0.04, 0.40], p = .048$ ).

Regarding math performance, math self-concept, and sense of belonging, we hypothesized that girls in the experimental condition would score lower on these outcomes than girls in the control condition. We did not hypothesize specific effects for boys on these outcomes. For these outcomes, the results revealed no significant effect of the experimental condition for girls when the respective pretest values were controlled for. For math performance and math self-concept, the Gender  $\times$  Condition interaction and the additionally calculated simple slope of the condition for boys were also not significant. With respect to sense of

belonging, the Gender  $\times$  Condition interaction was statistically significant ( $\beta = 0.30$ ,  $SE = 0.11$ , 95% CI [0.12, 0.49],  $p = .008$ ), and the effect of the condition for boys was statistically significant and positive ( $\beta = 0.20$ ,  $SE = 0.15$ , 95% CI [0.04, 0.36],  $p = .039$ ). This indicates that in contrast to girls, boys in the experimental group showed higher values of sense of belonging than boys in the control group (see Figure 1).

Regarding task values and attitudes toward math assessed with the feeling thermometer, we did not hypothesize specific effects of the experimental condition for girls and boys. With respect to the feeling thermometer, intrinsic value, attainment value, utility value for daily life, and cost, we found no significant effects of the experimental condition for either girls or boys, as indicated by nonsignificant Gender  $\times$  Condition interactions and nonsignificant simple slopes for boys when the respective pretest scores were controlled for.

For social utility, we computed multiple-group regression analyses using the order of the instruments as a grouping variable because a Wald  $\chi^2$  test indicated effects of the order of the instruments in the assessment as described above. Because we were interested in the effects of the experimental manipulation on social utility assessed with the questionnaire, the results for the students who were given the questionnaire first in the assessment were of major interest. Results for the students who were given the achievement test first in the assessment can be understood as a robustness check because possible effects of the experimental condition might have been washed out by the test. The results for both groups are presented in Table 7.

For the students who were given the questionnaire first, there was no significant effect of the condition for girls when the pretest was controlled for. The Gender  $\times$  Condition interaction was statistically significant ( $\beta = -0.88$ ,  $SE = 0.14$ , 95% CI [-1.12, -0.64],  $p < .001$ ) and the additionally calculated simple slope for boys was statistically significant and positive ( $\beta = -0.64$ ,  $SE = 0.21$ , 95% CI [-0.98, -0.30],  $p = .002$ ), indicating higher scores for social utility for boys in the experimental group than in the control group (see Figure 1).

For the students who were given the achievement test first, there was no significant effect of the condition for girls when their pretest scores were controlled for. The Gender  $\times$  Condition interaction and the simple slope for the boys were also not significant.

Table 5

## Multiple Regression Models 1: Effects on Stereotype Endorsement, Performance, Self-Concept, Sense of Belonging, and Feeling Thermometer

Predictor	Stereotype endorsement			Performance			Self-concept			Sense of belonging			Feeling thermometer		
	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI
Pretest	0.39***	0.08	[0.26, 0.52]	0.86***	0.03	[0.81, 0.91]	0.81***	0.05	[0.73, 0.89]	0.81***	0.04	[0.75, 0.87]	0.86***	0.04	[0.80, 0.92]
Gender (boys = 1)	0.29 <sup>†</sup>	0.16	[0.03, 0.55]	0.10	0.10	[-0.06, 0.26]	-0.01	0.07	[-0.14, 0.11]	-0.20*	0.10	[-0.36, -0.04]	0.04	0.08	[-0.10, 0.18]
Condition (exp. = 1)	0.50***	0.15	[0.03, 0.74]	0.04 <sup>a</sup>	0.09	[-0.10, 0.18]	0.03 <sup>a</sup>	0.07	[-0.09, 0.15]	-0.10 <sup>a</sup>	0.08	[-0.23, 0.02]	-0.12 <sup>†</sup>	0.07	[-0.23, -0.01]
Gender × Condition	-0.28	0.18	[-0.58, 0.02]	-0.07	0.10	[-0.24, 0.09]	0.12	0.09	[-0.03, 0.26]	0.30**	0.11	[0.12, 0.49]	0.10	0.10	[-0.07, 0.27]
Effect of condition for boys	0.22*	0.11	[0.04, 0.40]	-0.03	0.07	[-0.15, 0.09]	0.14 <sup>†</sup>	0.09	[0.00, 0.29]	0.20*	0.15	[0.04, 0.36]	-0.02	0.07	[-0.13, 0.10]

Note. All continuous variables are standardized. CI = confidence interval; exp. = experimental. <sup>a</sup>We formulated a hypothesis for this effect prior to the analysis.

<sup>†</sup>  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .



Table 6  
 Multiple Regression Models 2: Effects on Intrinsic Value, Attainment Value, Utility Value for Daily Life, and Cost

Predictor	Intrinsic value			Attainment value			Utility value: daily life			Cost		
	$\beta$	SE	$\beta$ 95% CI	$\beta$	SE	$\beta$ 95% CI	$\beta$	SE	$\beta$ 95% CI	$\beta$	SE	$\beta$ 95% CI
Pretest	0.87 <sup>***</sup>	0.04	[0.80, 0.93]	0.71 <sup>***</sup>	0.05	[0.62, 0.79]	0.62 <sup>***</sup>	0.06	[0.52, 0.71]	0.71 <sup>***</sup>	0.05	[0.62, 0.80]
Gender (boys = 1)	-0.03	0.09	[-0.17, 0.11]	-0.05	0.08	[-0.19, 0.09]	-0.09	0.14	[-0.32, 0.15]	-0.11	0.14	[-0.35, 0.12]
Condition (exp. = 1)	0.05	0.09	[-0.10, 0.19]	-0.02	0.07	[-0.14, 0.10]	0.03	0.12	[-0.17, 0.23]	-0.05	0.10	[-0.21, 0.11]
Gender $\times$ Condition	0.00	0.11	[-0.18, 0.18]	0.00	0.13	[-0.21, 0.21]	-0.03	0.17	[-0.31, 0.26]	0.05	0.15	[-0.19, 0.30]
Effect of condition for boys	0.05	0.08	[-0.09, 0.18]	-0.02	0.10	[-0.18, 0.15]	0.00	0.12	[-0.19, 0.19]	0.00	0.14	[-0.23, 0.23]

Note. All continuous variables are standardized. CI = confidence interval; exp. = experimental.

<sup>†</sup>  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table 7

*Multiple-Group Multiple Regression Model: Effects on Social Utility Value*

Predictor	Questionnaire first			Achievement test first		
	$\beta$	<i>SE</i>	$\beta$ 95% CI	$\beta$	<i>SE</i>	$\beta$ 95% CI
Pretest	0.65 <sup>***</sup>	0.03	[0.60, 0.71]	0.76 <sup>***</sup>	0.04	[0.70, 0.82]
Gender (boys = 1)	0.30 <sup>**</sup>	0.10	[0.14, 0.47]	0.10	0.14	[-0.13, 0.34]
Condition (exp. = 1)	0.24 <sup>†</sup>	0.14	[0.00, 0.48]	0.21 <sup>†</sup>	0.11	[0.03, 0.40]
Gender $\times$ Condition	-0.88 <sup>***</sup>	0.14	[-1.12, -0.64]	-0.08	0.18	[-0.38, 0.22]
Effect of condition for boys	-0.64 <sup>**</sup>	0.21	[-0.98, -0.30]	0.14	0.11	[-0.05, 0.32]

*Note.* All continuous variables are standardized. CI = confidence interval; exp. = experimental.

<sup>†</sup>  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

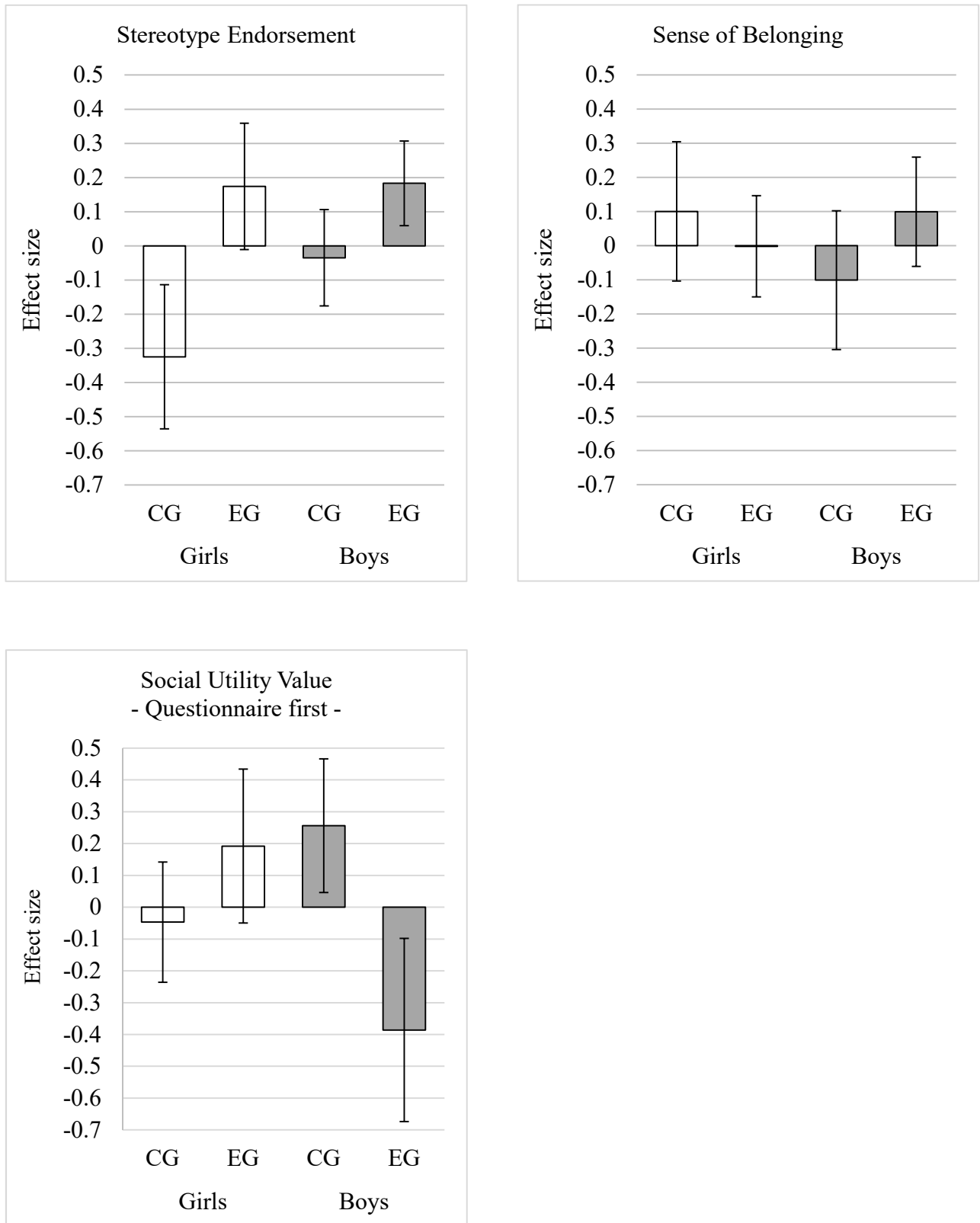


Figure 1. Effects of the experimental manipulation. Error bars represent 95% confidence intervals. CG = control group; EG = experimental group.

## Discussion

In this experimental study, we examined how stereotypes embedded in a children's television program about math influence girls' and boys' performance and motivation in math. We used a randomized controlled trial and a relatively large sample size, which enabled us to detect moderate-sized effects. The material we chose was a television program that had been broadcast on a German national television channel, thus contributing to the high validity of the study. Television programs play a central role in children's everyday lives and are an important part of their informal science learning, but such programs can also provide specific stereotypes about math such as "math is not for girls" or "boys who are good at math are geeky." Previous research has indicated that the stereotypes children encounter in their environment can impact young girls' and boys' math performance and motivation. Yet, such research has primarily been conducted in laboratory settings where stereotypes have been presented as isolated stimuli, rather than integrated into other information as would be the case in children's daily lives, for instance, in television programs.

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

### *Differential Effects of the Experimental Condition for Girls and Boys*

The stereotypes that were displayed in the video provide one possible explanation for the fact that we did not find any effects on girls' performance and motivation, although we found an effect on their stereotype endorsement. First, the girls in the video were bored with their math homework; thus, the video targeted the low motivation of these girls and not their performance or talent in math, which has typically been targeted in studies investigating the effects of stereotype threat (see e.g., Nguyen & Ryan, 2008). Furthermore, the stereotypes presented in the video also included the boy who is good at math but also geeky. To the best of our knowledge, effects of this stereotype on girls (as well as on boys) have not been investigated yet in research on stereotype threat.

However, it might be possible that the stereotype of the geeky math boy affected the girls positively because the girls experienced this stereotype as a negative stereotype against boys and consequently experienced positive effects of *stereotype lift*. Stereotype lift describes the effect of a performance boost for the nontargeted group in settings in which stereotypes are activated, such as for women after negatively stereotyping men's math performance (e.g., Johnson, Barnard-Brak, Saxon, & Johnson, 2012; Walton & Cohen, 2003). It might thus be possible that negative effects of stereotype threat (caused by the stereotype of the girls) and positive effects of stereotype lift (caused by the stereotype of the boy) cancelled each other out. However, due to the lack of previous research, this assumption is rather speculative, and future research is needed to disentangle possible differential effects of such stereotypes.

Although the results indicated that boys' performance and most of the motivational outcomes were not affected by the experimental manipulation, our findings provide initial insights into how such stereotypes in television shows might differentially affect boys. Thereby, the positive effect on boys' sense of belonging could be an indication of effects of stereotype lift on this outcome due to the traditional gender stereotypes in the video such as the stereotype that boys are very good at math. The negative effect on social utility, on the other hand, might indicate that boys are negatively affected by the stereotype that good math students are geeky. The boy in the part of the video in which the stereotypes were made salient was presented as socially incompetent, and the social utility scale directly referred to social acceptance. It is interesting that we found negative effects of the experimental condition on boys' social utility only for those students who were given the questionnaire first in the assessment (in which we assessed social utility). For those boys who were given the questionnaire after an achievement test, we did not find any effects of condition. This might indicate that such stereotypes affect boys' social utility only in the short term, and the effects were washed out after they completed the achievement test.

When interpreting the results of the present study for girls and for boys, the specific age group of the participants should be considered, as we investigated effects of stereotypes from television in fifth graders. Although previous research has indicated that children in elementary school are already aware of gender stereotypes (e.g., Ambady et al., 2001) and could be affected by them—at least with respect to math performance (Flore & Wicherts, 2015)—the participants in the present study may have been too young for the stereotypes shown in the program to affect their motivation. One reason for this assumption is provided by findings from the stereotype threat literature that indicated that group and domain identification moderate effects of stereotype threat (e.g., Lewis & Sekaquaptewa, 2016; Schmader, 2002). Given that children

increasingly identify with specific school subjects in elementary and middle school but do not differentiate much between the subjects at younger ages (see Wigfield et al., 2015), the participants in our study might have been too young and not sufficiently identified with the domain of math.

### ***Strengths and Limitations***

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

In our study, we investigated effects of two important facets of gender stereotypes in math: boys' advantages in math compared with girls and boys who are good at math being viewed as geeky. Both stereotypes are often presented simultaneously as was the case in the television program we used or in the American television sitcom *The Big Bang Theory* of the American TV network CBS, which is one of the highest rated and viewed television shows in the US and many other countries such as Germany (Kirsch, 2011; Patten, 2013). However, as described above, such stereotypes might differentially affect girls and boys, and it might be useful to investigate the stereotypes separately in future research in order to obtain a better understanding of possible influences.

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

To assess possible effects of the stereotypes embedded in the television program, we included several different outcome measures such as scales for measuring all dimensions of the

task values, for instance, or scales for assessing students' sense of belonging. The findings thus provide a comprehensive picture of possible effects on different outcomes. However, the measures we used were based on an achievement test and a questionnaire, which consisted of self-report measures. Our results thus provide no insights into how individuals might process the information presented in the video. Other assessment tools such as observational outcome measures (e.g., eye tracking) are necessary for investigating such processes.

### *Implications and Further Research*

Because our study provides only initial evidence on effects of stereotypes embedded in a television program on girls and boys, implications can be drawn only with caution. On the basis of the findings, one could argue that traditional gender stereotypes presented as one short section in a television programs do not seem to affect young girls in math, nor does the stereotype that mathematicians are geeky have a substantial effect on boys. This might be positive, particularly in light of the huge amount of time children spend watching television every day (Rideout, 2015; Rideout et al., 2010). Thus, it might not be necessary for program developers to care about stereotypes in developing television programs for children. However, in our study, we investigated effects of stereotypes in a television program in which only about one minute of the material had been manipulated. According to expectancy-value theory, it might be repeated experience that causes effects to accumulate and might sustainably affect boys and girls in the end, leading to gender differences in the math-intensive STEM fields (Eccles, 2009; Wigfield & Eccles, 2000). To address this question, more research is needed to explore effects of such stereotypes when they are presented repeatedly such as in a television series, for instance.

Our research furthermore adds to the discussion of the relevance of stereotype threat effects, particularly with respect to motivational outcomes (see Spencer et al., 2016). Despite effects of the experimental condition on girls' and boys' stereotype endorsement, we found hardly any effects on children's performance and motivation. This indicates that even if there might be effects of stereotype threat in the short term, they are not substantial in the long term. However, again, it might be repeated experience that might render effects of stereotype threat potentially harmful, and more research is needed to explore the duration of possible effects.

## References

- Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12*(5), 385–390. <https://doi.org/10.1111/1467-9280.00371>
- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectance on performance. *European Journal of Social Psychology, 33*, 267–285. <https://doi.org/10.1002/ejsp.145>
- Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology, 97*(6), 1045–1060. <https://doi.org/10.1037/a0016239>
- Cheryan, S., Plaut, V. C., Handron, C., & Hudson, L. (2013). The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women. *Sex Roles, 69*(1–2), 58–71. <https://doi.org/10.1007/s11199-013-0296-x>
- Collins, R. L. (2011). Content analysis of gender roles in media: Where are we now and where should we go? *Sex Roles, 64*(3), 290–298. <https://doi.org/10.1007/s11199-010-9929-5>
- Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images: How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin, 28*(12), 1615–1628. <https://doi.org/10.1177/014616702237644>
- Denissen, J. J. A., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development, 78*(2), 430–447. <https://doi.org/10.1111/j.1467-8624.2007.01007.x>
- Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences, 47*, 103–116. <https://doi.org/10.1016/j.lindif.2015.12.018>
- Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78–89. <https://doi.org/10.1080/00461520902832368>
- Eccles, J. S., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–121). San Francisco, CA: W. H. Freeman & Co.



- 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>
- Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology, 53*(1), 25–44. <https://doi.org/10.1016/j.jsp.2014.10.002>
- 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., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology, 38*(4), 519–533. <https://doi.org/10.1037//0012-1649.38.4.519>
- Gaspard, H., Dicke, A.-L., Flunger, B., Häfner, I., Brisson, B., Trautwein, U., & Nagengast, B. (2016). Side effects of motivational interventions? Effects of an intervention in math classrooms on motivation in verbal domains. *AERA Open, 2*(2), 1–14. <https://doi.org/10.1017/CBO9781107415324.004>
- Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*(3), 663–677. <https://doi.org/10.1037/edu0000003>
- 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>
- Good, J. J., Woodzicka, J. a, & Wingfield, L. C. (2010). The effects of gender stereotypic and counter-stereotypic textbook images on science performance. *The Journal of Social Psychology, 150*(2), 132–47. <https://doi.org/10.1080/00224540903366552>
- Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in the Schools, 30*(1), 79–90. [https://doi.org/10.1002/1520-6807\(199301\)30:1<79::AID-PITS2310300113>3.0.CO;2-X](https://doi.org/10.1002/1520-6807(199301)30:1<79::AID-PITS2310300113>3.0.CO;2-X)
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549–576. <https://doi.org/10.1146/annurev.psych.58.110405.085530>

- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*(1), 3–25. <https://doi.org/10.1037/0033-295X.109.1.3>
- Hannover, B., & Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learning and Instruction*, *14*(1), 51–67. <https://doi.org/http://dx.doi.org/10.1016/j.learninstruc.2003.10.002>
- Heyman, K. (2008). Talk nerdy to me. *Science*, *320*(5877), 740–741. <https://doi.org/10.1126/science.320.5877.740>
- 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>
- Johnson, H. J., Barnard-Brak, L., Saxon, T. F. ., & Johnson, M. K. . (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *Journal of Experimental Education*, *80*(2), 137–149. Retrieved from <http://10.0.4.56/00220973.2011.567312>
- Kessels, U., Rau, M., & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, *76*(4), 761–780. <https://doi.org/10.1348/000709905X59961>
- KiKa.de. (2015). Pur+. Mathe ist Magie [Pur+. Math is magic]. Germany. Retrieved from <http://www.kika.de/pur/sendungen/sendung84728.html>
- Kirsch, J. (2011). “The Big Bang Theory” - nächster Glückstreffer für ProSieben? [“The big bang theory” - the next lucky strike for ProSieben?]. Retrieved February 23, 2017, from <http://www.quotenmeter.de/cms/?p1=n&p2=48491>
- Lambert, K., Dackermann, T., & Möller, K. (2017). Mathematischer Basiskompetenztest für die weiterführende Schule [Math basic competence test for secondary school]. *Manuscript in Preparation*.
- Leaper, C. (2015). Gender and social-cognitive development. In *Handbook of child psychology and developmental science* (pp. 1–48). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118963418.childpsy219>
- Lewis, N. A., & Sekaquaptewa, D. (2016). Beyond test performance: A broader view of stereotype threat. *Current Opinion in Psychology*, *11*, 40–43.

- <https://doi.org/10.1016/j.copsyc.2016.05.002>
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>
- 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>
- Master, A., Cheryan, S., & Meltzoff, A. N. (2015). Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of Educational Psychology, 108*(3), 424–437. <https://doi.org/10.1037/edu0000061>
- McNeish, D., Stapleton, L. M., & Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods, 22*(1), 114–140. <https://doi.org/10.1037/met0000078>
- Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review, 120*(3), 544–60. <https://doi.org/10.1037/a0032459>
- Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat. *Psychological Science, 18*(10), 879–885. <https://doi.org/10.1111/j.1467-9280.2007.01995.x>
- Muthén, B. O., & Muthén, L. K. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Muzzatti, B., & Agnoli, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. *Developmental Psychology, 43*(3), 747–759. <https://doi.org/10.1037/0012-1649.43.3.747>
- 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 Research Council. (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: The National Academic Press. <https://doi.org/https://doi.org/10.17226/12190>
- National Science Board. (2016). *Science & engineering indicators 2016*. <https://doi.org/10.1002/ejoc.201200111>

- National Science Foundation. (2015). *Women, minorities, and persons with disabilities in science and engineering 2015. Special Report NSF*. [https://doi.org/Special Report NSF 11-309](https://doi.org/Special%20Report%20NSF%2011-309)
- Neuville, E., & Croizet, J.-C. (2007). Can salience of gender identity impair math performance among 7-8 years old girls? The moderating role of task difficulty. *European Journal of Psychology of Education, 22*(3), 307–316. <https://doi.org/10.1007/BF03173428>
- Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology, 93*(6), 1314–1334. <https://doi.org/10.1037/a0012702>
- Passolunghi, M. C., Rueda Ferreira, T. I., & Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learning and Individual Differences, 34*, 70–76. <https://doi.org/10.1016/j.lindif.2014.05.005>
- Patten, D. (2013). Full 2012-2013 TV season series rankings. Retrieved February 23, 2017, from <http://deadline.com/2013/05/tv-season-series-rankings-2013-full-list-506970/>
- 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*(3), 225–235. <https://doi.org/10.1016/j.cedpsych.2013.03.004>
- Raudenbush, S. W. (1997). Statistical analysis and optimal design for cluster randomized trials. *Psychological Methods, 2*(2), 173–185. <https://doi.org/10.1037/1082-989X.2.2.173>
- Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. -f., Martinez, A., & Bloom, H. S. (2011). Optimal design software for multi-level and longitudinal research (Version 3.01). Retrieved from [www.wtgrantfoundation.org](http://www.wtgrantfoundation.org)
- Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Educational Progress assessments. *Journal of Educational Psychology, 107*(3), 645–662. <https://doi.org/10.1037/edu0000012>
- Rideout, V. J. (2015). *The common sense census: Media use by tweens and teens*. San Francisco, CA: Common Sense Media. <https://doi.org/10.1017/CBO9781107415324.004>
- Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). *Media in the lives of 8-18 year olds*. Menlo Parc, CA. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Schmader, T. (2002). Gender identification moderates stereotype threat effects on women’s math performance. *Journal of Experimental Social Psychology, 38*(2), 194–201.

- <https://doi.org/10.1006/jesp.2001.1500>
- Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles, 50*(11/12), 835–850. <https://doi.org/10.1023/B:SERS.0000029101.74557.a0>
- Schoon, I., & Eccles, J. S. (Eds.). (2014). *Gender differences in aspiration and attainment: A life course perspective*. Cambridge, U.K.: Cambridge University Press.
- Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review, 13*(2), 147–183. <https://doi.org/10.1006/drev.1993.1007>
- Smith, J. L., Brown, E. R., Thoman, D. B., & Deemer, E. D. (2015). Losing its expected communal value: How stereotype threat undermines women's identity as research scientists. *Social Psychology of Education, 18*(3), 443–466. <https://doi.org/10.1007/s11218-015-9296-8>
- Smith, J. L., Sansone, C., & White, P. H. (2007). The stereotyped task engagement process: The role of interest and achievement motivation. *Journal of Educational Psychology, 99*(1), 99–114. <https://doi.org/10.1037/0022-0663.99.1.99>
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology, 67*, 415–37. <https://doi.org/10.1146/annurev-psych-073115-103235>
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4–28. <https://doi.org/10.1006/jesp.1998.1373>
- Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *The American Psychologist, 52*(6), 613–629. <https://doi.org/10.1037/0003-066X.52.6.613>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 69*(5), 797–811.
- Steele, C. M., Spencer, S. J., & Aronson, J. M. (2002). Contending with group image: The psychology of stereotype and identity threat. *Advances in Experimental Social Psychology, 34*, 379–440. [https://doi.org/10.1016/S0065-2601\(02\)80009-0](https://doi.org/10.1016/S0065-2601(02)80009-0)
- Steele, J. R., & Ambady, N. (2006). “Math is hard!” The effect of gender priming on women's attitudes. *Journal of Experimental Social Psychology, 42*(4), 428–436. <https://doi.org/10.1016/j.jesp.2005.06.003>
- Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and

- adolescents. *The Journal of Educational Psychology*, 102(4), 947–963.  
<https://doi.org/10.1037/a0019920>
- 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>
- Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology*, 47(4), 943–949. <https://doi.org/10.1037/a0024047>
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174–1204. <https://doi.org/10.1037/a0036620>
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology*, 39(5), 456–467. [https://doi.org/10.1016/S0022-1031\(03\)00019-2](https://doi.org/10.1016/S0022-1031(03)00019-2)
- Watt, H. M. G., & Eccles, J. S. (Eds.). (2008). *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences*. Washington, DC: American Psychological Association. <https://doi.org/10.1037/11706-000>
- 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. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 91–120). San Diego, San Francisco, New York, Boston, London, Sydney, Tokyo: Academic Press.
- Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. In R. M. Lerner, C. Garcia Coll, & M. Lamb (Eds.), *Handbook of child psychology and developmental science* (pp. 1–44). Hoboken, NJ, USA: John Wiley & Sons, Inc.  
<https://doi.org/10.1002/9781118963418.childpsy316>

## Supplemental Materials

### *1 Robustness Check Material (Tables S1-S3)*

Relevance: Because the television program was broadcast on a national TV channel in Germany, we assessed whether participants had already seen the video beforehand, which was the case for 41 students. As a robustness check, we also computed all analyses without these students.

### *2 Sense of Belonging – Full Scale (Table S4)*

Relevance: Due to low item-scale correlations, we excluded one item when we computed the scale used in the manuscript. Because we did not preregister the exclusion of the item, we conducted the analysis for this outcome also using the original scale, which included all 10 items.

### *3 Robustness Check Manipulation Check (Tables S5-S7)*

Relevance: As a manipulation check, we asked the students what they had seen in the last minute of the video, that is, two girls who copied their homework from the “math nerd” or a summary of the video. Thirteen students did not answer the question correctly. In our analyses, we conducted an intention-to-treat analysis by taking only the original assignment into account, but as robustness checks, we excluded the students in the experimental condition who failed the manipulation check.

### *4 Effects on Motivation in German (Tables S8-S15)*

Relevance: We preregistered exploratory analyses with respect to motivational outcomes in the domain of German. Due to space limitations, the results on girls’ and boys’ motivation in German are reported in the Supplemental Material.

Table S1

*Robustness Check Material 1: Results of Multiple Regression Models for Effects on Stereotype Endorsement, Performance, Self-concept, Sense of Belonging, and Feeling Thermometer*

Predictor	Stereotype endorsement			Performance			Self-concept			Sense of belonging			Feeling thermometer		
	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI
Pretest	0.44 <sup>***</sup>	0.06	[0.34, 0.54]	0.86 <sup>***</sup>	0.03	[0.81, 0.92]	0.81 <sup>***</sup>	0.05	[0.72, 0.89]	0.81 <sup>***</sup>	0.04	[0.75, 0.87]	0.85 <sup>***</sup>	0.05	[0.78, 0.92]
Gender (boys = 1)	0.26 <sup>†</sup>	0.15	[0.01, 0.52]	0.09	0.09	[-0.05, 0.23]	-0.02	0.08	[-0.15, 0.11]	-0.18 <sup>*</sup>	0.09	[-0.33, -0.03]	0.10	0.09	[-0.05, 0.24]
Condition (exp. = 1)	0.45 <sup>**</sup>	0.15	[0.20, 0.69]	-0.01	0.09	[-0.16, 0.15]	0.02	0.09	[-0.12, 0.16]	-0.08	0.08	[-0.22, 0.06]	-0.11	0.08	[-0.24, 0.03]
Gender × Condition	-0.21	0.20	[-0.54, 0.11]	-0.02	0.10	[-0.19, 0.15]	0.09	0.11	[-0.10, 0.28]	0.26 <sup>*</sup>	0.11	[0.07, 0.44]	0.05	0.12	[-0.14, 0.24]
Effect of condition for boys	0.23 <sup>†</sup>	0.12	[0.04, 0.43]	-0.03	0.07	[-0.14, 0.09]	0.11	0.11	[-0.06, 0.29]	0.18 <sup>†</sup>	0.10	[0.02, 0.34]	-0.06	0.07	[-0.18, 0.06]

*Note.* All continuous variables are standardized. Students who stated that they had already seen the video beforehand were excluded ( $N = 41$ ). The sample consisted therefore of  $N = 294$  students. CI = confidence interval; exp. = experimental.

<sup>†</sup>  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .



Table S2  
*Robustness Check Material 2: Results of Multiple Regression Models for Effects on Intrinsic Value, Attainment Value, Utility Value, Utility Value for Daily Life, and Cost*

Predictor	Intrinsic value			Attainment value			Utility value: daily life			Cost		
	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI
Pretest	0.86 <sup>***</sup>	0.05	[0.79, 0.94]	0.72 <sup>***</sup>	0.06	[0.63, 0.81]	0.63 <sup>***</sup>	0.06	[0.53, 0.73]	0.70 <sup>***</sup>	0.06	[0.60, 0.80]
Gender (boys = 1)	-0.02	0.09	[-0.17, 0.12]	-0.08	0.09	[-0.22, 0.07]	-0.08	0.14	[-0.32, 0.15]	-0.11	0.16	[-0.37, 0.15]
Condition (exp. = 1)	0.03	0.10	[-0.13, 0.18]	-0.04	0.09	[-0.19, 0.10]	-0.11	0.12	[-0.31, 0.09]	-0.05	0.10	[-0.22, 0.13]
Gender $\times$ Condition	0.01	0.13	[-0.20, 0.21]	-0.01	0.15	[-0.26, 0.24]	0.03	0.17	[-0.25, 0.31]	0.03	0.17	[-0.24, 0.31]
Effect of condition for boys	0.03	0.09	[-0.12, 0.19]	-0.05	0.01	[-0.22, 0.12]	-0.08	0.12	[-0.27, 0.11]	-0.01	0.16	[-0.28, 0.25]

*Note.* All continuous variables are standardized. Students who stated that they had already seen the video beforehand were excluded ( $N = 41$ ). The sample consisted therefore of  $N = 294$  students. CI = confidence interval; exp. = experimental.

<sup>†</sup>  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S3

*Robustness Check Material 3: Results of Multiple Group Multiple Regression Models for Effects on Social Utility Value*

Predictor	Questionnaire first			Achievement test first		
	$\beta$	<i>SE</i>	$\beta$ 95% CI	$\beta$	<i>SE</i>	$\beta$ 95% CI
Pretest	0.62 ***	0.04	[0.56, 0.69]	0.79 ***	0.05	[0.72, 0.87]
Gender (boys = 1)	0.31 **	0.10	[0.15, 0.48]	0.10	0.15	[-0.24, 0.26]
Condition (exp. = 1)	0.34 †	0.20	[0.00, 0.67]	0.16	0.10	[-0.01, 0.33]
Gender × Condition	-0.98 ***	0.13	[-1.20, -0.76]	-0.03	0.21	[-0.37, 0.31]
Effect of condition for boys	-0.64 **	0.21	[-0.98, -0.3 ]	0.13	0.13	[-0.08, 0.34]

*Note.* All continuous variables are standardized. Students who stated that they had already seen the video beforehand were excluded ( $N = 41$ ). The sample consisted therefore of  $N = 294$  students. CI = confidence interval; exp. = experimental.

†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S4

*Sense of Belonging – Full Scale: Results of Multiple Regression Models*

Predictor	$\beta$	<i>SE</i>	$\beta$ 95% CI
Pretest	0.81 ***	0.04	[0.75, 0.87]
Gender (boys = 1)	-0.19 †	0.10	[-0.36, -0.02]
Condition (exp. = 1)	-0.010	0.07	[-0.22, 0.02]
Gender × Condition	0.31 **	0.12	[0.12, 0.51]
Effect of condition for boys	0.21 *	0.10	[0.04, 0.39]

*Note.* All continuous variables are standardized. The scale for sense of belonging consisted of ten items, but because of low item-scale correlations ( $r_{it} = .03/.16$ ), we excluded one item in the computation of the scale used for the analysis in the manuscript. Here, we present results for the original scale, which included all ten items.

†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S5  
*Robustness Check Manipulation Check 1: Results of Multiple Regression Models for Effects on Stereotype Endorsement, Performance, Self-concept, Sense of Belonging, and Feeling Thermometer*

Predictor	Stereotype endorsement			Performance			Self-concept			Sense of belonging			Feeling thermometer		
	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI
Pretest	0.36***	0.09	[0.22, 0.50]	0.86***	0.03	[0.81, 0.92]	0.80***	0.05	[0.72, 0.88]	0.80***	0.04	[0.74, 0.86]	0.85***	0.04	[0.79, 0.92]
Gender (boys = 1)	0.28†	0.16	[0.02, 0.54]	0.09	0.10	[-0.07, 0.25]	-0.01	0.08	[-0.13, 0.12]	-0.20*	0.10	[-0.36, -0.03]	0.04	0.09	[-0.10, 0.18]
Condition (exp. = 1)	0.56***	0.15	[0.32, 0.80]	0.03	0.09	[-0.12, 0.18]	0.05	0.08	[-0.07, 0.18]	-0.09	0.07	[-0.21, 0.03]	-0.12†	0.07	[-0.23, -0.01]
Gender × Condition	-0.33†	0.18	[-0.62, -0.03]	-0.05	0.10	[-0.22, 0.12]	0.09	0.09	[-0.05, 0.24]	0.26*	0.12	[0.06, 0.47]	0.10	0.11	[-0.08, 0.28]
Effect of condition for boys	0.23*	0.12	[0.04, 0.43]	-0.02	0.08	[-0.15, 0.10]	0.15	0.10	[-0.01, 0.30]	0.18†	0.10	[0.01, 0.34]	-0.02	0.07	[-0.14, 0.11]

*Note.* All continuous variables are standardized. Students who failed the manipulation check were excluded ( $N = 13$ ). The sample consistent therefore of  $N = 322$  students. CI = confidence interval; exp. = experimental.

†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S6

*Robustness Check Manipulation Check 2: Results of Multiple Regression Models for Effects on Intrinsic Value, Attainment Value, Utility Value, Utility Value for Daily Life, and Cost*

Predictor	Intrinsic value			Attainment value			Utility value: daily life			Cost		
	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI
Pretest	0.86 <sup>***</sup>	0.04	[0.79, 0.93]	0.71 <sup>***</sup>	0.05	[0.62, 0.79]	0.63 <sup>***</sup>	0.06	[0.53, 0.72]	0.71 <sup>***</sup>	0.05	[0.62, 0.79]
Gender (boys = 1)	-0.03	0.09	[-0.18, 0.11]	-0.03	0.08	[-0.17, 0.10]	-0.09	0.14	[-0.33, 0.14]	-0.14	0.14	[-0.37, 0.09]
Condition (exp. = 1)	0.03	0.09	[-0.11, 0.17]	-0.02	0.08	[-0.16, 0.11]	-0.01	0.12	[-0.20, 0.19]	-0.10	0.09	[-0.24, 0.05]
Gender $\times$ Condition	0.02	0.11	[-0.17, 0.20]	0.00	0.13	[-0.21, 0.22]	0.00	0.17	[-0.29, 0.29]	0.12	0.14	[-0.11, 0.34]
Effect of condition for boys	0.05	0.09	[-0.10, 0.19]	-0.02	0.10	[0.00, 0.00]	-0.02	0.11	[-0.18, 0.14]	0.02	0.14	[-0.21, 0.25]

*Note.* All continuous variables are standardized. Students who failed the manipulation check were excluded ( $N = 13$ ). The sample consistent therefore of  $N = 322$  students. CI = confidence interval; exp. = experimental.

<sup>†</sup>  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S7

*Robustness Check Manipulation Check 3: Results of Multiple Group Multiple Regression Models for Effects on Social Utility Value*

Predictor	Questionnaire first			Achievement test first		
	$\beta$	SE	$\beta$ 95% CI	$\beta$	SE	$\beta$ 95% CI
Pretest	0.66 ***	0.04	[0.60, 0.73]	0.75 ***	0.04	[0.69, 0.82]
Gender (boys = 1)	0.30 **	0.10	[0.14, 0.47]	0.13	0.15	[-0.11, 0.37]
Condition (exp. = 1)	0.24	0.15	[-0.01, 0.49]	0.24 *	0.11	[0.07, 0.42]
Gender $\times$ Condition	-0.88 ***	0.11	[-1.07, -0.70]	-0.11	0.19	[-0.42, 0.21]
Effect of condition for boys	-0.65 ***	0.17	[-0.93, -0.37]	0.14	0.12	[-0.05, 0.33]

*Note.* All continuous variables are standardized. Students who failed the manipulation check were excluded ( $N = 13$ ). The sample consistent therefore of  $N = 322$  students. CI = confidence interval.

†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S8

*Effects on Motivation in German 1: Descriptive Statistics for All Study Variables on the Pretest Separated by Gender*

Variable	Girls		Boys		$d^a$	$d$ 95% CI
	$M$	$SD$	$M$	$SD$		
Self-concept T1	3.07	0.08	2.83	0.08	-0.32	-0.50 -0.14
Sense of belonging T1	3.14	0.06	2.30	0.05	-0.27	-0.49 -0.05
Intrinsic value T1	3.05	0.11	2.65	0.08	-0.48	-0.69 -0.23
Attainment value T1	3.57	0.07	3.33	0.05	-0.38	-0.64 -0.12
Utility value – daily life T1	3.09	0.08	2.93	0.09	-0.21	-0.43 0.00
Utility value – social T1	2.05	0.05	1.99	0.05	-0.09	-0.02 0.06
Cost T1	1.56	0.06	1.80	0.07	0.36	0.18 0.54

*Note.* CI = 95% confidence interval.

<sup>a</sup>The dependent variable is standardized.

Table S9

*Effects on Motivation in German 2: Descriptive Statistics for All Study Variables at T1 Separated by Gender and Group*

Variable	Girls						Boys					
	Experimental group			Control group			Experimental group			Control group		
	M	SD	Max	M	SD	Max	M	SD	Max	M	SD	Max
Self-concept T1	3.07	0.08	4.00	3.07	0.11	4.00	2.77	0.08	4.00	2.88	0.11	4.00
Sense of belonging T1	3.11	0.07	4.00	3.18	0.08	4.00	2.93	0.06	4.00	3.07	0.07	4.00
Intrinsic value T1	3.04	0.12	4.00	3.07	0.14	4.00	2.63	0.10	4.00	2.68	0.13	4.00
Attainment value T1	3.61	0.08	4.00	3.52	0.09	4.00	3.32	0.06	4.00	3.35	0.09	4.00
Utility Value – daily life T1	3.14	0.09	4.00	3.02	0.10	4.00	2.78	0.09	4.00	3.07	0.12	4.00
Utility value – social T1	2.07	0.08	3.67	2.03	0.10	4.00	1.92	0.08	3.67	2.07	0.07	3.67
Cost T1	1.61	0.07	3.50	1.50	0.09	3.25	1.83	0.08	3.75	1.76	0.11	3.50

Table S10  
*Effects on Motivation in German 3: Descriptive Statistics for All Study Variables at T2 Separated by Gender and Group*

Variable	Girls						Boys									
	Experimental group			Control group			Experimental group			Control group						
	<i>M</i>	<i>SD</i>	Max	<i>M</i>	<i>SD</i>	Max	<i>M</i>	<i>SD</i>	Max	<i>M</i>	<i>SD</i>	Max				
Self-concept T2	3.00	0.07	1.25	4.00	3.08	0.08	1.00	4.00	2.86	0.08	1.00	4.00	2.94	0.09	1.00	4.00
Sense of belonging T2	3.10	0.05	2.00	4.00	3.08	0.06	1.78	4.00	3.06	0.05	1.44	4.00	3.01	0.07	1.33	4.00
Intrinsic value T2	2.89	0.10	1.25	4.00	2.84	0.09	1.00	4.00	2.81	0.09	1.00	4.00	2.78	0.12	1.00	4.00
Attainment value T2	3.48	0.05	1.50	4.00	3.49	0.05	1.25	4.00	3.43	0.07	1.75	4.00	3.37	0.10	1.00	4.00
Utility Value – daily life T2	3.06	0.09	1.00	4.00	3.05	0.07	1.67	4.00	2.89	0.07	1.00	4.00	3.02	0.10	1.00	4.00
Utility value – social T2	2.12	0.06	1.00	4.00	1.90	0.05	1.00	4.00	2.01	0.07	1.00	3.67	2.10	0.07	1.00	3.67
Cost T2	1.58	0.06	1.00	3.25	1.66	0.08	1.00	3.75	1.66	0.07	1.00	4.00	1.79	0.08	1.00	4.00

Table S11

*Effects on Motivation in German 4: Correlations Between all Study Variables in German*

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.
1. Self-concept T1	—													
2. Self-concept T2	.86	—												
3. Sense of belonging T1	.64	.58	—											
4. Sense of belonging T2	.61	.66	.78	—										
5. Intrinsic value T1	.72	.69	.59	.52	—									
6. Intrinsic value T2	.73	.78	.56	.63	.86	—								
7. Attainment value T1	.35	.37	.30	.30	.45	.45	—							
8. Attainment value T2	.32	.35	.26	.38	.37	.44	.76	—						
9. Utility value – daily life T1	.28	.25	.33	.25	.46	.40	.33	.29	—					
10. Utility value – daily life T2	.35	.39	.31	.35	.45	.48	.36	.35	.75	—				
11. Utility value – social T1	.19	.18	.21	.17	.30	.25	.28	.21	.28	.30	—			
12. Utility value – social T2	.21	.17	.16	.19	.18	.22	.24	.21	.19	.25	.66	—		
13. Cost T1	-.73	-.69	-.56	-.52	-.69	-.68	-.31	-.28	-.34	-.39	-.12	-.07	—	
14. Cost T2	-.64	-.73	-.49	-.63	-.58	-.68	-.27	-.36	-.26	-.37	(-.09)	(-.03)	.78	—

*Note.* Non-significant correlations are displayed in parentheses, for all other correlations,  $p < .05$ .



Table S12

*Effects on Motivation in German 5: Results of Multiple Regression Models for Outcomes in German*

Predictor	Intrinsic value			Utility - daily life		
	$\beta$	<i>SE</i>	$\beta$ 95% CI	$\beta$	<i>SE</i>	$\beta$ 95% CI
Pretest	0.84 ***	0.04	[0.79, 0.90]	0.73 ***	0.03	[0.68, 0.79]
Gender (boys = 1)	-0.07	0.12	[-0.27, 0.13]	-0.04	0.13	[-0.25, 0.17]
Condition (exp. = 1)	0.06	0.09	[-0.08, 0.21]	0.02	0.11	[-0.16, 0.20]
Gender $\times$ Condition	-0.03	0.12	[-0.23, 0.17]	-0.17	0.16	[-0.44, 0.09]
Effect of condition for boys	0.03	0.08	[-0.20, 0.41]	-0.16	0.10	[-0.46, 0.22]

*Note.* All continuous variables are standardized. CI = confidence interval; exp. = experimental.

†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S13

*Effects on Motivation in German 6: Results of Multiple Group Multiple Regression Models for Effects on Self-concept and Sense of Belonging in German*

Predictor	Self-concept						Sense of belonging					
	Questionnaire first			Achievement test first			Questionnaire first			Achievement test first		
	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI	$\beta$	SE	95% CI
Pretest	0.84 <sup>***</sup>	0.05	[0.75, 0.93]	0.85 <sup>***</sup>	0.02	[0.83, 0.88]	0.73 <sup>***</sup>	0.07	[0.62, 0.84]	0.81 <sup>***</sup>	0.03	[0.76, 0.87]
Gender (boys = 1)	-0.05	0.13	[-0.26, 0.16]	-0.30 <sup>*</sup>	0.13	[-0.51, -0.08]	-0.05	0.17	[-0.32, 0.23]	-0.21 <sup>†</sup>	0.13	[-0.42, 0.00]
Condition (exp. = 1)	0.10	0.07	[-0.02, 0.21]	-0.30 <sup>***</sup>	0.09	[-0.46, -0.15]	0.21	0.14	[-0.03, 0.44]	-0.12	0.09	[-0.27, 0.03]
Gender × Condition	-0.15	0.11	[-0.33, 0.03]	0.17	0.15	[-0.07, 0.41]	-0.19	0.16	[-0.46, 0.08]	0.25 <sup>*</sup>	0.12	[0.06, 0.44]
Effect of condition for boys	-0.06	0.13	[-0.35, 0.35]	-0.13	0.10	[-0.13, 0.47]	0.01	0.19	[-0.46, 0.58]	0.13	0.12	[-0.02, 0.71]

*Note.* All continuous variables are standardized. CI = confidence interval; exp. = experimental.

<sup>†</sup>  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S14

*Effects on Motivation in German 7: Results of Multiple Group Multiple Regression Models for Effects on Attainment Value and Social Utility Value in German*

Predictor	Attainment value				Social utility value			
	Questionnaire first		Achievement test first		Questionnaire first		Achievement test first	
	$\beta$	SE	$\beta$	SE	$\beta$	SE	$\beta$	SE
Pretest	0.76 <sup>***</sup>	0.05 [0.68, 0.85]	0.75 <sup>***</sup>	0.08 [0.61, 0.89]	0.69 <sup>***</sup>	0.03 [0.64, 0.74]	0.66 <sup>***</sup>	0.05 [0.58, 0.74]
Gender (boys = 1)	0.10	0.14 [-0.12, 0.32]	-0.44 <sup>†</sup>	0.26 [-0.86, -0.02]	0.30 <sup>*</sup>	0.14 [0.08, 0.52]	0.27 <sup>†</sup>	0.15 [0.03, 0.51]
Condition (exp. = 1)	0.15	0.12 [-0.04, 0.34]	-0.17 <sup>†</sup>	0.10 [-0.33, -0.01]	0.13	0.13 [-0.08, 0.33]	0.50 <sup>***</sup>	0.14 [0.28, 0.73]
Gender × Condition	-0.33	0.23 [-0.71, 0.05]	0.48 <sup>*</sup>	0.24 [0.08, 0.88]	-0.38 <sup>†</sup>	0.19 [-0.7, -0.06]	-0.53 <sup>*</sup>	0.22 [-0.9, -0.17]
Effect of condition for boys	-0.18	0.20 [-0.78, 0.23]	0.31	0.20 [0.05, 1.45]	-0.25 <sup>†</sup>	0.15 [-0.98, -0.13]	-0.03	0.13 [-0.71, 0.10]

Note. All continuous variables are standardized. CI = confidence interval; exp. = experimental.

<sup>†</sup>  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table S15

*Effects on Motivation in German 8: Results of Multiple Group Multiple Regression Models for Effects on Cost in German*

Predictor	Questionnaire first			Achievement test first		
	$\beta$	<i>SE</i>	$\beta$ 95% CI	$\beta$	<i>SE</i>	$\beta$ 95% CI
Pretest	0.75 ***	0.07	[0.63, 0.86]	0.78 ***	0.05	[0.70, 0.87]
Gender (boys = 1)	0.05	0.12	[-0.15, 0.24]	0.32 †	0.17	[0.04, 0.61]
Condition (exp. = 1)	-0.31 **	0.11	[-0.48, -0.14]	0.09	0.15	[-0.15, 0.34]
Gender × Condition	0.05	0.22	[-0.31, 0.40]	-0.21	0.19	[-0.51, 0.10]
Effect of condition for boys	-0.26	0.26	[-0.83, 0.22]	-0.11	0.11	[-0.85, -0.02]

*Note.* All continuous variables are standardized. CI = confidence interval; exp. = experimental.

†  $p < .10$ . \*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

## Study 2:

# Maximizing Gender Equality by Minimizing Course Choice Options? Effects of Obligatory Coursework in Math on Gender Differences in STEM

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. <http://dx.doi.org/10.1037/edu0000183>

\*The first two authors contributed equally to this work and are listed in alphabetical order. This research was partially funded by the state government of Baden-Württemberg.

© 2017 American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors permission. The final article will be available, upon publication, via its DOI: 10.1037/edu0000183

### **Abstract**

Math achievement, math self-concept, and vocational interests are critical predictors of STEM careers and are closely linked to high school coursework. Young women are less likely to choose advanced math courses in high school, and encouraging young women to enroll in advanced math courses may therefore bring more women into STEM careers. We looked at a German statewide educational reform that required all students to take advanced math courses and examined differential effects of the reform on young men and women's math achievement, math self-concept, vocational interests, and field of study at university. We compared data from 4,730 students before the reform and 4,715 students after the reform. We specified multiple regression models and tested main effects of gender and cohort as well as the effect of the Cohort  $\times$  Gender interaction on all outcomes. All outcomes showed clear gender differences favoring young men before the reform. However, the reform was associated with different effects for young men and women: Whereas gender differences in math achievement were smaller after the reform, differences between young men and women in math self-concept and realistic and investigative vocational interests were larger after the reform than before. Gender differences in the field of study at university did not differ between before and after the reform. Results suggest that reducing course choice options in high school does not automatically increase gender equality in STEM fields.

### **Educational Impact and Implications Statement**

This study suggests that making it obligatory for young women and men to participate in advanced math courses in upper secondary school can increase their math achievement and realistic (e.g., technical) interests. However, it also seems to have the potential to negatively impact young women's self-perceptions of their math ability. The study illustrates that well-intended educational reforms might not achieve all goals (and in fact might result in unintended side effects) when psychological factors are ignored.

*Keywords:* gender differences, school reform, math achievement, math self-concept, vocational interests

## **Maximizing Gender Equality by Minimizing Course Choice Options? Effects of Obligatory Coursework in Math on Gender Differences in STEM**

Women are underrepresented in mathematically intensive STEM (science, technology, engineering, and mathematics) domains (Ceci, Williams, & Barnett, 2009; Schoon & Eccles, 2014). Gender disparities in STEM fields are crucial for the larger economy because the presence of more women would diversify the workforce and might add to a more competitive work environment with an increased number of qualified employees in this area (e.g., NSF, 2013; OECD, 2010). In addition, women's underrepresentation also matters to gender inequity in income because STEM fields provide high-status career options (e.g., Sells, 1980; Watt, Eccles, & Durik, 2006). Advanced high school coursework in math is a key predictor of STEM career choices (Ma & Johnson, 2008), and young women are less likely to choose advanced math courses than young men (Nagy et al., 2008; Updegraff, Eccles, Barber, & O'Brien, 1996). Thus, it is important to ask whether the challenge of recruiting more women into STEM careers may be addressed by mandatory enrollment in advanced math courses in high school (e.g., by changing course assignment procedures; Ma & Johnson, 2008; Sells, 1980). However, there is limited real-world data on the effectiveness of such reforms.

In the present study, we re-analyzed representative data from a large school achievement study on the effects of a major reform of upper secondary education in a large state in Germany. More specifically, the reform required all students to take an advanced math course, which successfully eliminated a prior imbalance between young men and women in these advanced courses. We studied the effects of this school reform on gender differences in math achievement, math self-concept, and interests in realistic and investigative areas because such outcomes are critical in terms of later educational choices. Furthermore, we investigated effects on students' actual field of study at university 2 years after they completed high school.

### ***Predictors of Gendered Career Choices in STEM***

#### *Academic Achievement and STEM Career Choices*

In explaining STEM career choices for young men and women, research on educational choices has traditionally focused on the role of math achievement on career interests (e.g., Parker et al., 2012; Sells, 1980). Such work has consistently shown that math achievement is a key predictor of both high school subject choices and later career choices, particularly with respect to mathematically intensive STEM careers (Parker et al., 2012; Sells, 1980). For instance, there is evidence that high school math achievement predicts career aspirations in

STEM during high school (e.g., Ma & Johnson, 2008), field of study at university (e.g., Parker et al., 2012), and university retention (Alarcon & Edwards, 2013).

The relation between academic achievement and career choice is often explained by employing rational choice models (Gottfredson, 1986; Lubinski & Benbow, 2006). First, individuals prefer careers that provide activities they expect to be good at. Second, individuals who have the required competencies gain access to the professional field, for instance, due to admission restrictions for college majors. Third, individuals tend to leave professions if their competencies are insufficient for the specific profession. Thus, young people with high math achievement have a tendency to pursue mathematically intensive STEM careers such as physics, engineering, or informatics (Humphreys & Yao, 2002; Parker et al., 2012).

### *Self-Concept and STEM Career Choices*

Above and beyond the effects of achievement, young people's career choices are also critically linked to their academic self-concept in high school (Schoon & Eccles, 2014; Watt & Eccles, 2008). Academic self-concept is defined as a person's self-evaluation of his or her own general ability in a specific domain, such as doing well in math (Bong & Skaalvik, 2003; Marsh, 1986). In developing a domain-specific self-concept, students refer to their own achievement in a domain but also compare their own ability with their interpretation of peers' achievements in the same domain (e.g., Marsh, 1986; Marsh et al., 2015).

In fact, self-concept has been shown to be related to future-oriented motivation and aspirations such as career choices (e.g., Schoon & Eccles, 2014; Watt & Eccles, 2008); math self-concept has been identified as positively related to various educational outcomes in the STEM area, such as high school students' educational aspirations within the STEM fields (Jansen, Scherer, & Schroeders, 2015; Schoon & Eccles, 2014) and choice and retention of mathematically intensive STEM university subjects (Perez, Cromley, & Kaplan, 2014; Schoon & Eccles, 2014) for both men and women.

It is important to mention that self-concept does not measure the same thing as self-efficacy, although they are closely related (e.g., Bong & Skaalvik, 2003). Furthermore, self-concept predicts educational biographies and trajectories, whereas self-efficacy is used for predicting success in a specific task (Jansen, Scherer, & Schroeders, 2015).

### *Vocational Interests and STEM Career Choices*

Next to math achievement and self-concept, vocational interests are very important in predicting STEM career choices. The role of interest for achievement-related outcomes is well-established (Schoon & Eccles, 2014; Su, Rounds, & Armstrong, 2009). Whereas educational



psychology has traditionally focused on children's and adolescents' interest in learning and achievement in the school context (Krapp, 1999; Wigfield & Cambria, 2010), research and theories in vocational psychology, such as Holland's theory of vocational interests (Holland, 1959, 1997), have been highly effective at addressing young people's postschool career choices with interests describing activities in fields of professions or university majors (Rounds & Su, 2014; Su & Rounds, 2015). Vocational interests are central predictors of vocational choices such as the selection of a college major or profession (Humphreys & Yao, 2002; Pässler, Beinicke, & Hell, 2014) and are also crucial for job performance and turnover (Nye, Su, Rounds, & Drasgow, 2012) as well as income (Huang & Pearce, 2013).

Holland (1966) defined vocational interests as "the expression of personality in work, hobbies, recreational activities, and preferences" (p. 3) and expected that they would directly influence goal-oriented behaviors. He posited that individuals should strive for educational and occupational environments that are in line with their interests, and there is a large body of research that supports this proposition (e.g., Humphreys & Yao, 2002; Strong, 1943). Vocational interests are therefore defined as trait-like preferences for activities, and these preferences are captured on a very general level (Holland, 1997; Rounds & Su, 2014). In this regard, vocational interests differ from the term interest in educational psychology. Interest in educational psychology is usually defined as a motivational variable that "refers to the psychological state of engaging or the predisposition to reengage" (Hidi & Renninger, 2006, p. 112). Contrary to conceptualizations of interest in educational psychology, which usually focus on domain-specific interest in single (school) subjects (e.g., Hidi & Ainley, 2002), vocational interests emphasize broad sets of activities and experiences that go with different kinds of professions. Thereby, Holland's model represents six interest domains, which describe activities that are related to different careers: realistic, investigative, artistic, social, enterprising, and conventional. In our study, we focused on the realistic and investigative dimensions because they have been shown to be related to mathematically intensive STEM fields (Ackerman & Heggestad, 1997; Su et al., 2009). People with high realistic interests tend to like working with things and prefer activities that involve the manipulation of objects, tools, and machines. People with high investigative interests are likely to be interested in understanding how physical and biological phenomena function and tend to prefer activities that include analyzing and problem solving on a more abstract level (Holland, 1997). Consequently, young people with realistic and investigative interests are likely to choose mathematically intensive STEM careers such as physics, engineering, or informatics (Su & Rounds, 2015; Su, Rounds, & Armstrong, 2009).

*Gender Differences in Math Achievement, Math Self-Concept, and Realistic and Investigative Interests*

Gender differences in math achievement have often been used to explain gendered career choices in the STEM domains (e.g., Hyde, Fennema, Ryan, Frost, & Hopp, 1990; Reilly, Neumann, & Andrews, 2015). Historically, there has been a pattern of young men outperforming young women in math achievement (e.g., Hyde, Fennema, & Lamon, 1990). However, more recent research has provided mixed evidence: Some studies have suggested no or only slight differences in math achievement between young women and men in high school (e.g., Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Voyer & Voyer, 2014), whereas others have indicated that such differences still exist and that the magnitude of the differences between young men and women varies between countries and according to the educational requirements of the system (e.g., Else-Quest, Hyde, & Linn, 2010; Reilly et al., 2015). For German samples, previous research has consistently indicated that young men still perform better in math in high school than young women (e.g., Else-Quest et al., 2010; Nagy et al., 2008).

Regarding math self-concept, previous research has shown that—after achievement is controlled for—boys tend to report higher math self-concept than girls even in primary school, and such gender differences remain constant across high school (e.g., Marsh & Yeung, 1998; Nagy et al., 2008).

With respect to realistic and investigative interests, previous research has consistently shown that men score higher on both interest dimensions than women (e.g., Lippa, 1998; Su et al., 2009).

*Relations between Achievement, Self-Concept, and Vocational Interests*

Academic achievement, the self-evaluation of academic achievement (i.e., self-concept), and interests have been found to be interrelated, which means that, in general, people are interested in and feel competent in domains they are good at. The relations between these constructs have been described in different theoretical frameworks, such as Eccles et al. (1983) expectancy-value theory and Lent, Brown, and Hackett's (1994) social cognitive career theory. According to these theories, prior achievement influences an individual's evaluation of his or her achievement (e.g., self-concept), as well as his or her interests in the same domain. A person's interests are furthermore influenced by his or her perception of competence, and both self-concept and interests are believed to predict later achievement. The rationale behind these relations is that individuals who have positive previous achievement-related experiences in one domain will feel more competent and will develop interests in the same domain. Furthermore,

if they feel competent and are interested, they will engage more frequently and intensely in tasks and activities related to that domain, and thereby, they will show high levels of persistence and effort. In the end, this leads to better performance in the same domain (Wigfield, Tonks, & Klauda, 2009).

There is a lot of empirical support for such relations between achievement, self-concept, and interests. With respect to the relation between achievement and self-concept, several studies have indicated that achievement and self-concept are positively correlated (e.g., Chen, Yeh, Hwang, & Lin, 2013), and bidirectional relations have been found, indicating that students' prior achievement influences their self-concept and that their self-concept influence their later achievement (for a review, see Marsh, 2007). Furthermore, there is evidence that self-concept predicts changes in interests (Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Wigfield et al., 1997) and that interests and achievement are also interrelated. Thereby, correlation-based research has shown positive relations between achievement and interests for various conceptualizations of interest, such as individual interest (see Schiefele, Krapp, & Wintler, 1992) or task values (see e.g., Updegraff, Eccles, Barber, & O'Brien, 1996), but also for vocational interests, where positive correlations between math achievement and realistic as well as investigative interests have been found (Ackerman et al., 1997; Rolfhus & Ackerman, 1996). Furthermore, self-concept has been found to predict later interests (e.g. Denissen, Zarrett, & Eccles, 2007; Marsh et al., 2005), and a reciprocal relation has been found between interests and achievement (e.g., Denissen et al., 2007; Jansen, Lüdtke, & Schroeders, 2016). However, prior studies have so far focused on subject-specific conceptualizations of interest, and less is known about directional relations between these constructs and realistic and investigative interests.

#### *Effects of Course Level on Achievement, Self-Concept, and Vocational Interests*

Students' achievement, self-concept, and vocational interests have been linked to their enrollment in advanced and basic courses in high school (e.g., Köller, Baumert, & Schnabel, 2001; Marsh, 2005). The effects of high school coursework on achievement, self-concept, and interests have been explained by variability in the benefits for and constraints on students taking basic and advanced courses (e.g., Köller, 2001; Marsh, 2005). In Germany, as in most school systems in developed countries, students in upper secondary school self-select into basic and advanced math courses, which differ in terms of curricular content and level as well as in class composition (Schnabel, Alfeld, Eccles, Köller, & Baumert, 2002). These differences between advanced and basic coursework have been found to lead to differential effects on students'

achievement, self-concept, and interests, after students' previous performance was controlled for (e.g., Köller, et al. 2001; Trautwein, Köller, Lüdtke, & Baumert, 2005). Regarding students' academic achievement, course level and achievement have been found to be positively associated; students in advanced courses have typically shown higher achievement at the end of high school than those in basic courses, even after students' prior achievement was taken into account (e.g., Gamoran & Mare, 1989; Köller et al., 2001).

Effects of course level on self-concept and vocational interests are less clear. Regarding self-concept, positive associations have been found between a student's own achievement and his or her self-concept in the same domain, as described in the previous section (Marsh, 1986; Marsh et al., 2014). Thus, students showed higher self-concept in advanced courses than in basic courses in general (Chmielewski, Dumont, & Trautwein, 2013). However, students tend to compare their own achievement with the perceived achievement of their classmates and consequently judge their own achievement as relatively lower when they are surrounded by students with higher achievement. Therefore, students in advanced courses have shown a lower self-concept than students with comparable achievement in basic courses (Chmielewski et al., 2013; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006).

With respect to vocational interests, research has shown that students in advanced and basic courses differ in their vocational interests because their course choices are based on their vocational interests (Nagy & Husemann, 2010; Patrick, Care, & Ainley, 2011). However, it is less clear if or how course level might also predict vocational interests. First, a positive association has been identified between achievement and vocational interests as described above, on which basis one might speculate that course level in math might positively influence realistic and investigative interests (Ackerman & Heggestad, 1997; Anthony & Armstrong, 2010). Second, initial findings have indicated effects of the average level of class achievement on students' vocational interests. Cambria, Brandt, Nagengast, and Trautwein (2016) investigated 10th graders' achievement in several domains and their vocational interests. They found that achievement in math was positively associated with realistic and investigative interests and that students with the same individual math achievement level had higher realistic and investigative interests when they were in a class with a higher mean level of achievement.

To sum up, math achievement, math self-concept, and vocational interests are central predictors of mathematically intensive STEM career choices, and these predictors explain gendered career choices in these fields. The findings regarding gender differences in math achievement have been inconsistent, but a considerable amount of research has shown that young men demonstrate higher math self-concept and STEM-related vocational interests than

young women. Furthermore, the existing literature indicates that students' achievement and self-concept in math as well as their STEM-related interests are closely related to high school coursework.

### *The Present Study*

In the present study, we examined the effects of a reform in upper secondary high school on gender differences in central predictors of STEM career choices and students' choice of STEM university subjects by reanalyzing representative data from 9,545 German students. Math high school coursework has been found to be closely linked to achievement, self-concept, and interests in the STEM fields (Nagy et al., 2008; Updegraff et al., 1996), all of which are central predictors of STEM career choices (Ma & Johnson, 2008; Nagy et al., 2008). A lower percentage of young women than men had chosen advanced math courses before the reform took place, but this difference was completely eliminated by the reform because the reform required all students to take advanced math courses. Thus, we expected effects of the reform on gender differences in STEM-related outcomes.

There is ample evidence of such effects of high school coursework on achievement, self-concept, and interests, but previous research has not addressed how gender differences in math achievement, self-concept, and interests as key predictors of STEM career choices may be influenced by requiring all students to enroll in advanced courses in math. The present study takes a major step toward filling this gap by investigating such an educational policy and its effects on women's participation in the STEM fields. We examined how changes in high school coursework are related to gender differences in predictors of STEM career choices and students' subjects of study at university after school. To do so, we evaluated effects of a school reform that was introduced in 2002 in one of the largest German states. The reform included the abolition of different math courses. Before the reform, students had been allowed to take math as either an advanced or a basic course. After the reform, all students had to take an obligatory advanced-level math course (Ministry of Education and Cultural Affairs, Youth and Sport Baden-Württemberg, 2002).

Because high school course level tends to predict students' achievement and self-concept, and because young women were less likely than young men to choose advanced courses in math before the reform, we expected that the effects of the reform on these outcomes would differ between the young women and men in the current study. As positive effects of course level on students' achievement have been documented, we hypothesized that gender differences in math achievement would be smaller after the reform (when all young men and

women took advanced math courses) compared with before the reform (when more young men than young women had taken advanced math courses). Here, we assume that the smaller gender differences in achievement expected after the reform would be based on the higher achievement of young women after the reform compared with before. Regarding gender differences in math self-concept, we hypothesized that gender differences would be larger after the reform than before. This proposition was based on the finding that course level tends to have negative effects on a student's self-concept, and there was a higher percentage of young men than young women in advanced courses before the reform, whereas all students took advanced courses after the reform. We therefore expected that young women's self-concept would be lower after the reform than before on average, which would lead to greater gender differences in math self-concept. So far, there is less work on effects of high school coursework on vocational interests, and it is therefore not clear whether and how the reform might be related to gender differences in realistic and investigative interests. However, if we were to find similar effects of course level on STEM-related vocational interests as on self-concept and subject-specific interest, we would tentatively expect larger gender differences in realistic and investigative interests after the reform than before.

Because we expected differential reform effects on central predictors of STEM career choices (math achievement, math self-concept, realistic and investigative interests), we did not specify what the effects on the actual choice of STEM university subjects would be.

## **Method**

### ***The Reform of Upper Secondary School in the German School System***

Before the reform of upper secondary school education, students in most German states self-selected their courses and were given the choice between math as an advanced course (about five hours per week) or a basic course (about three hours per week). In total, each student was required to select two advanced courses and typically six basic courses in different subjects. The individual combination of advanced and basic courses represented an individual profile for each student for all of their upper secondary school trajectories, and students were not able to choose different courses each semester. Beginning in 2002, most German states enacted reforms of their higher secondary education systems and implemented a course program. This program can be characterized by a reduction in the number of options in favor of a higher subject-related average amount of time allocated across all students to specific compulsory core subjects (e.g., German, mathematics, and foreign language). In most states, students were no longer able to self-select into different courses from that point in time on but were instead required to take a

total of five courses from specific fields (e.g., math, foreign language, science) for a similar amount of time (4 hr per week). Besides these compulsory courses, students had to participate in other courses for a reduced number of hours (2 hr per week; e.g., arts, science, or social studies; Köller, Watermann, Trautwein, & Lüdtke, 2004; Trautwein, Neumann, Nagy, Lüdtke, & Maaz, 2010). To sum up, the two major changes of the reform were (a) an increase in the number of courses that had to be chosen for final examinations in upper secondary school on an advanced course level and (b) written exams in the first four of these courses (instead of the first three).

### *Description of Study and Sample*

Data were drawn from the study “Transformation of the Secondary School System and Academic Careers” (TOSCA; Köller et al., 2004; Trautwein et al., 2010). The TOSCA study was designed to assess a representative sample of students in the last 4 months of their final year of upper secondary school in one German state (Baden-Württemberg). The data from the first waves of TOSCA 2002 and TOSCA 2006 are representative for all students in the final year of upper secondary school in the state of Baden-Württemberg. We considered data from  $N = 149$  schools in the first wave of the first cohort (TOSCA 2002;  $N = 4,730$ ; 54.5% female) as well as data from  $N = 146$  schools in the first wave of the second cohort (TOSCA 2006;  $N = 4,715$ ; 54.1% female). Over the course of the reform, another school type (biotechnological Gymnasium) was introduced. Robustness checks revealed no differences in results when students from this type of school were included versus not included. In our sample, roughly 60% of the students were enrolled in a general higher secondary school, and 40% were in a vocational upper secondary school. The time between the start of the course and our measurement was approximately 1.5 years. The measurement took place right at the end of the course. Data collection was executed by trained research assistants who visited every class and lasted for approximately one day per school. The first cohort contains data from students who chose basic and advanced courses in upper secondary high school, whereas the second cohort consists of data from students who all took the obligatory advanced math courses. The data from the two cohorts were drawn from the same schools. In both cohorts, a second assessment took place 2 years after the first measurement point via questionnaires that were sent to the participants. Overall, 80% of all students agreed to participate in the first wave of TOSCA 2002, and 82% of all students agreed to participate in the first wave of TOSCA 2006. At the second assessment, which followed 2 years after the first assessments for TOSCA 2002 and TOSCA

2006, respectively, information was obtained about students' field of study at university from  $N = 1,741$  students from TOSCA 2002 and  $N = 2,157$  from TOSCA 2006 (see Figure 1).

Measures	2002	2004	2006	2008
Math achievement Math self-concept Vocational interests	TOSCA 2002 Wave 1 $N = 4,730$		TOSCA 2006 Wave 1 $N = 4,715$	
Field of study at university		TOSCA 2002 Wave 2 $N = 2,318$		TOSCA 2006 Wave 2 $N = 2,852$

Figure 1. Schematic illustration of the study's timeline. All data in Wave 1 were collected at the end of upper secondary school.

### *Instruments*

#### *Math achievement*

The *Advanced mathematics test* was based on items from the Third International Mathematics and Science Study (TIMSS; Mullis et al., 1998). According to Mullis et al. (1998), the advanced mathematics test takes into account "current thinking and priorities in the field of mathematics" (p. 284). The advanced mathematics test contained a total of 68 items from the areas of (a) Numbers, Equations, and Functions, (b) Analysis, (c) Geometry, (d) Propositional Logic and Proofs, as well as (e) Probability and Statistics. Most of the items were related to the first area and directly tested competencies from upper secondary school. Approximately two thirds of all of the items were multiple-choice questions, whereas the other items were administered in an open-ended format. A multimatrix design was used to administer the items; therefore, the students did not work on all 68 items but on a subset of items in one of four booklets that contained six different item clusters that were rotated systematically. In order to be able to compare the two different TOSCA 2002 and TOSCA 2006 cohorts, items were scaled by applying item response theory (IRT; Rasch model) to account for the multimatrix design and to test for differential item functioning. As reported by Nagy, Neumann, Trautwein, & Lüdtke (2010), we used five completed data sets with *plausible values (PVs)*, which were estimated in Mplus 5.2. These PVs were based on multiply imputed data, which was imputed previously with NORM (Schafer, 1997). As reported by Nagy et al. (2010), the psychometric properties of the test are good (PV reliability TOSCA 2002: .88; PV reliability TOSCA 2006: .90).



*Mathematics self-concept*

Mathematics self-concept was measured with four items from the Self Description Questionnaire III (SDQ III; Marsh & O'Neill, 1984; Marsh & Shavelson, 1985; Marsh, 1992), using the German translation by Schwanzer, Trautwein, Lüdtke, and Sydow (2005). The translated items focused on the evaluation of cognitive aspects (e.g., "I was always good in mathematics," e.g., Marsh, Trautwein, Lüdtke, Baumert, & Köller, 2007). The scale showed high internal consistency in both samples, TOSCA 2002 (Cronbach's  $\alpha = .89$ ) and TOSCA 2006 (Cronbach's  $\alpha = .90$ ).

*Vocational interests*

Vocational interests were assessed with the Revised General Interest Structure Test (AIST-R; Bergmann & Eder, 2005), which is based on Holland's (1997) RIASEC model. This instrument categorizes students with regard to six different dimensions of interest, namely, realistic (R), investigative (I), artistic (A), social (S), enterprising (E), and conventional (C) interests by using a total of 60 items (six 10-item scales). Students were asked to rate how interested they were in the described activities on a 5-point Likert scale ranging from 1 (*not at all*) to 5 (*very much*). An example item of realistic interests is "Working with machines or technical devices" and "Doing physically challenging work," whereas investigative interests were assessed with items such as "Dealing with unexplored things" and "Working in an experimental laboratory." The realistic and investigative facets, which were of specific interest in the present context, showed high internal consistencies (realistic interests—TOSCA 2002: Cronbach's  $\alpha = .86$ ; TOSCA 2006: Cronbach's  $\alpha = .87$ ; investigative interests—TOSCA 2002 and 2006: Cronbach's  $\alpha$ s = .85).

*Field of study at university*

The field of study at university was assessed for each cohort 2 years after they graduated from high school. Students were able to report their subject of study or a combination of study subjects. Students' data were coded according to the official classification system of the Federal Statistical Office, the Fachserie 11 (Federal statistical office, 2008). In the current study, we used information about the field of study and computed one variable for which mathematics, engineering, computer science, and physics were coded as STEM subjects only if they were indicated as the first subject of study. In addition, we also specified various alternative codings where only the first, the first two, or all three subject indications were used to calculate the dependent variable and included biology, chemistry, or both as STEM subjects. The general

pattern of results was identical across all these different analyses. Furthermore, we did not find any significant differences in STEM-related course change or student withdrawal patterns when comparing the first and second assessments between TOSCA 2002 and TOSCA 2006. The results were based on analyses in which mathematics, engineering, computer science, and physics were coded as STEM subjects.

### *Covariates*

We controlled for the influence of several variables described below.

*School types.* Because students from different school types (e.g., vocational higher secondary schools and general higher secondary schools) usually differ in cognitive and noncognitive aspects (Trautwein et al., 2010), we included a dummy variable to be able to distinguish between vocational and general higher secondary schools.<sup>1</sup>

*Socioeconomic background.* Socioeconomic background was measured with information about the highest level of occupation in the family (of either the father or mother) and coded in accordance with the International Standard Classification of Occupations (ISCO-88). The ISCO scores were in turn converted into International Socio-Economic Index of Occupational Status (ISEI) 88 scores (Ganzeboom, De Graaf, & Treiman, 1992; Ganzeboom & Treiman, 1996). The highest ISEI value between the two parents was used to characterize the socioeconomic background of the students.

*Number of books available in the home.* The number of books available in the home was measured on a 7-point scale ranging from zero books available to more than 500 books available. This variable has been shown to be a good indicator of a family's cultural capital (e.g., Evans, Kelley, Sikora, & Treiman, 2010).

*Age.* The age of the students at the time of the assessment was calculated on the basis of information about students' year and month of birth.

*Immigration background.* Students with at least one parent born outside of Germany were coded as students with an immigration background.

### *Statistical Analyses*

In order to test for reform effects, we specified multiple regression models involving the TOSCA study survey weights and tested gender as a moderator of the effect of the reform on the different STEM-related outcomes. The models contained the variables gender and cohort as

---

<sup>1</sup> Due to the different vocational school types that were considered in the TOSCA studies, we also specified models with dummy-coded variables for every type of vocational school as additional robustness checks. The results did not differ meaningfully.

well as socioeconomic background (HISEI), cultural capital (number of books), immigration background, type of school (general Gymnasium vs. type of vocational Gymnasium), and age as covariates. We controlled for these covariates to eliminate the influence of these potential confounders and to increase the precision of our estimation. In addition, we added the Cohort  $\times$  Gender interaction in order to examine whether the reform had differential effects on young women and men. Because students from different types of schools usually differ in their cognitive and behavioral outcomes (Trautwein et al., 2010), we also controlled for this differential impact by including the three-way interaction between Cohort  $\times$  School Type  $\times$  Gender as well as the interaction between School Type  $\times$  Cohort.

We also specified a multivariate model with a Wald test for the interaction effects and controlled for the false discovery rate of all parameter estimates in each multiple regression afterwards by applying the Benjamini-Hochberg adjustment (Benjamini & Hochberg, 1995).

We additionally investigated students' actual field of study at university 2 years after they completed high school. Of special interest in the current analysis were potential differences with regard to whether or not students chose a STEM-related field of study. We therefore specified models to predict field of study in STEM versus other fields of study in multiple logistic regressions.

We used the statistical software R (R Development Core Team, 2014) and the survey package (Lumley, 2014) to inspect the data. The final models were specified in Mplus 7.4 (Muthén & Muthén, 2012). All models took into consideration survey weights to obtain representative results for students in upper secondary schools in Baden-Württemberg.

In order to report meaningfully interpretable coefficients, we present fully standardized coefficients, meaning that both the dependent and continuous independent variables were standardized. We also present partially standardized coefficients, meaning that only the dependent variable was standardized (also referred to as Cohen's *d*; Cohen, 1988). Continuous variables were centered. The partially standardized coefficients might be especially useful for interpreting effects of dichotomous variables. With regard to the fully standardized solution, the interaction terms were standardized before we included them in the regression models. In order to explore and interpret possible interaction effects, we additionally estimated simple main effects between the two cohorts for young women and men and school types for statistically significant three-way interactions by using the model constraint option in Mplus 7.4. Estimating simple main effects to interpret interactions is also recommended by Jaccard, Wan, and Turrisi (1990). Furthermore, we also calculated structure coefficients (e.g., Courville & Thompson, 2001) to gain further insights into the dynamics of our data. Structure coefficients

indicate the proportion of the multiple correlation that can be accounted for by the first-order correlation. When multicollinearity is high, the beta weights might be relatively small. However, structure coefficients are able to indicate this more precisely.

### *Effect sizes*

Regarding the interpretation of effect sizes and on the basis of a literature review, as suggested by Henson (2006), we argue that effect sizes of  $d > 0.05$  should be considered practically relevant. As can be seen in the literature, this seems to be the average amount of growth that can be expected from a half to 1 year of schooling (e.g., Hill, Bloom, Black, & Lipsey, 2008; Low, Yoon, Roberts, & Rounds, 2005; Low, 2009; Nagy et al., 2010; Wagner, Rose, Dicke, Neumann, & Trautwein, 2014). However, as stated in Henson (2006), benchmarks should be used cautiously.

### *Cluster structure*

Students from the same class or school cannot be treated as independent observations because they are more similar to each other than they are to students from other classes or schools. Not considering this cluster structure leads to overestimated standard errors (Snijders & Bosker, 2012). To address the clustered data structure (students were nested within classes), standard errors were adjusted by applying a design-based correction as implemented in Mplus (Muthén & Muthén, 1998-2012), which automatically takes the multilevel structure into account and makes use of a sandwich estimator (see e.g., Asparouhov, 2005; Muthén & Satorra, 1995). Here, we followed McNeish, Stapleton, and Silverman's (2016) recommendations as they pointed out that alternative design-based methods (or population-averaged methods) can be more intuitive and do not rely on assumptions that are inherent in the specification of random effects in hierarchical linear modeling. Design-based methods allow the researcher to adjust the standard errors of estimates and fit statistics for the nested structure of the data and have been shown to perform well in various different nested data settings (e.g., Stapleton, Yang, & Hancock, 2016).

### *Missing values*

Missing values are a common problem in the social sciences, and several approaches have been implemented to account for missing values in a meaningful way (e.g., Enders, 2010; Graham, 2009). There is a growing consensus that approaches such as multiple imputation (MI) or full information maximum likelihood (FIML) estimation are superior to traditional methods (e.g., complete case analysis or pairwise deletion). For all outcomes except math achievement

and all independent variables, missing values were addressed with full information maximum likelihood in Mplus 7.4 (Muthén & Muthén, 2012). There were no missing values on the math achievement tests as we used plausible values that were generated for every student and the primary analysis of the TOSCA study (Nagy et al., 2010).

### Results

In the first step, the two cohorts were compared with respect to possible differences in the covariates (Table 1). Overall, these pre-existing differences between the two cohorts seemed to be of small practical relevance. Differences were found only for age ( $d = -0.22, p < .001$ ), largely due to the TOSCA 2006 assessment taking place a little bit later in the school year because of an organizational issue. However, because this difference applied equally to young women and men, it should not have had any effect on the results. Furthermore, we controlled for age in all analyses. In addition, a difference in the number of books available in the home ( $d = -0.06, p = .021$ ) was significant, whereas differences on all other variables (including gender) were not significant.

Next, we compared the lengths of time (in hours per week) allocated to mathematics by gender between the two cohorts before and after the reform. Table 2 shows a difference in the average amount of time allocated to math for both young men (3.5 min per week) and young women (19.7 min per week) and an average increase in the total sample after the reform (12.2 min). As expected, the average amount of time allocated to mathematics increased more for young women than for young men as shown by a significant Gender  $\times$  Cohort interaction ( $B = 16.20, p < .001$ ).

#### *Test of Advanced Mathematics Achievement*

We hypothesized that the gender difference in math achievement in favor of young men would be smaller after the reform that introduced the obligatory advanced mathematics course for both young men and women. To test our prediction, we used multiple regression analyses to explore a possible difference between the two cohorts in advanced mathematics achievement (Table 3).

The Cohort  $\times$  Gender interaction was statistically significant ( $d = 0.14, p = .025, 95\%$  CI [0.01, 0.26]). In line with our hypothesis, the interaction indicated a smaller difference between young women and men after the reform than before (see Figure 2). This was mainly due to a higher average level of young women's achievement after the reform ( $d = 0.14, p = .002, 95\%$  CI [0.05, 0.22]), whereas young men's achievement did not differ before and after the reform ( $d = 0.00, p = .988, 95\%$  CI [-0.11, 0.11]). The Cohort  $\times$  School Type interaction ( $d$

= 0.08,  $p = .255$ , 95% CI [-0.08,0.22]) was not statistically significant, but the Cohort  $\times$  Gender  $\times$  School Type interaction had a significant regression weight ( $d = -0.19$ ,  $p = .029$ , 95% CI [-0.35,0.02]), indicating that the effects of the reform differed between the different school types. Our results indicate a three-way interaction between Cohort  $\times$  Gender  $\times$  School Type. Exploring this interaction revealed statistically significant differences for young women, but not for young men, before versus after the reform for general gymnasiums but not for vocational gymnasiums, in favor of the cohort that was measured after the reform. However, for young men, the effect of the reform was not statistically significantly different between vocational gymnasiums and general gymnasiums.

Table 1

*Descriptive Statistics for the Two Cohorts*

Variable	TOSCA 2002	TOSCA 2006	Effect size	$p$
Gender (% female)	54.0%	53.1%	0.98	.679
Immigration (% immigrants)	20.0%	20.8%	1.08	.237
HISEI	59.16 (15.57)	58.49 (15.73)	-0.04	.120
Books	5.64 (1.22)	5.57 (1.23)	-0.06	.021
Age	19.56 (0.79)	19.40 (0.65)	-0.22	< .001
Math achievement	50.10 (9.82)	51.07 (9.42)	0.10	.002
Math self-concept	2.76 (0.81)	2.70 (0.85)	-0.08	.003
Realistic vocational interests	2.08 (0.74)	2.24 (0.80)	0.20	< .001
Investigative vocational interests	2.60 (0.83)	2.64 (0.81)	0.04	.138

*Note.* Weighted results. For dichotomous dependent variables, logistic regression was used to test the differences. For continuous dependent variables, linear regression was used. HISEI = highest international socioeconomic index. Effect sizes: for dichotomous dependent variables, odds ratios (*ORs*) are displayed; for continuous dependent variables, Cohen's  $d$  (Cohen, 1988) is displayed.

Table 2

*Time Allocated to Mathematics Before and After the Reform*

	Enrollees in advanced math courses (TOSCA 2002)	Average allocated time (TOSCA 2002)	Average allocated time (TOSCA 2006)	Increase (TOSCA 2006)	<i>p</i>
Young men	44.7%	3.92 hr (177 min)	4 hr (180 min)	3.49 min	.007
Young women	27.9%	3.56 hr (160 min)	4 hr (180 min)	19.68 min	< .001
Total	35.5%	3.73 hr (167 min)	4 hr (180 min)	12.17 min	< .001

*Note.* Results for TOSCA 2002 are based on self-reported course choice. The analyses took into consideration the survey weights and clustered structure of the data. One lesson lasted for 45 min. In TOSCA 2006, the average time allocated by young men and women was equal because of the mandatory advanced course.

Table 3

*Predicting Advanced Mathematics Achievement: Results from Multiple Regressions Models*

Predictor	<i>B</i>	<i>p</i>	<i>SE</i>	<i>d</i> <sup>a</sup>
Cohort (T2 = 1)	.00	.988	0.06	0.00
Gender (f = 1)	-.58	< .001	0.04	-0.58
HISEI	.00	.912	0.01	0.00
Books	.07	< .001	0.01	0.06
Immigration (=1)	-.16	< .001	0.03	-0.16
Age	-.18	< .001	0.02	-0.24
School type (VS =1)	-.61	< .001	0.06	-0.61
Cohort × Gender	.14	.025	0.06	0.14
Cohort × School Type	.09	.255	0.07	0.08
Cohort × Gender × School Type	-.19	.029	0.09	-0.19
R <sup>2</sup>		.23		

*Note.* All values are fully standardized. Continuous variables are centered. T2 = TOSCA 2006; HISEI = highest international socioeconomic index; VS = vocational school.

<sup>a</sup>The dependent variable is standardized.

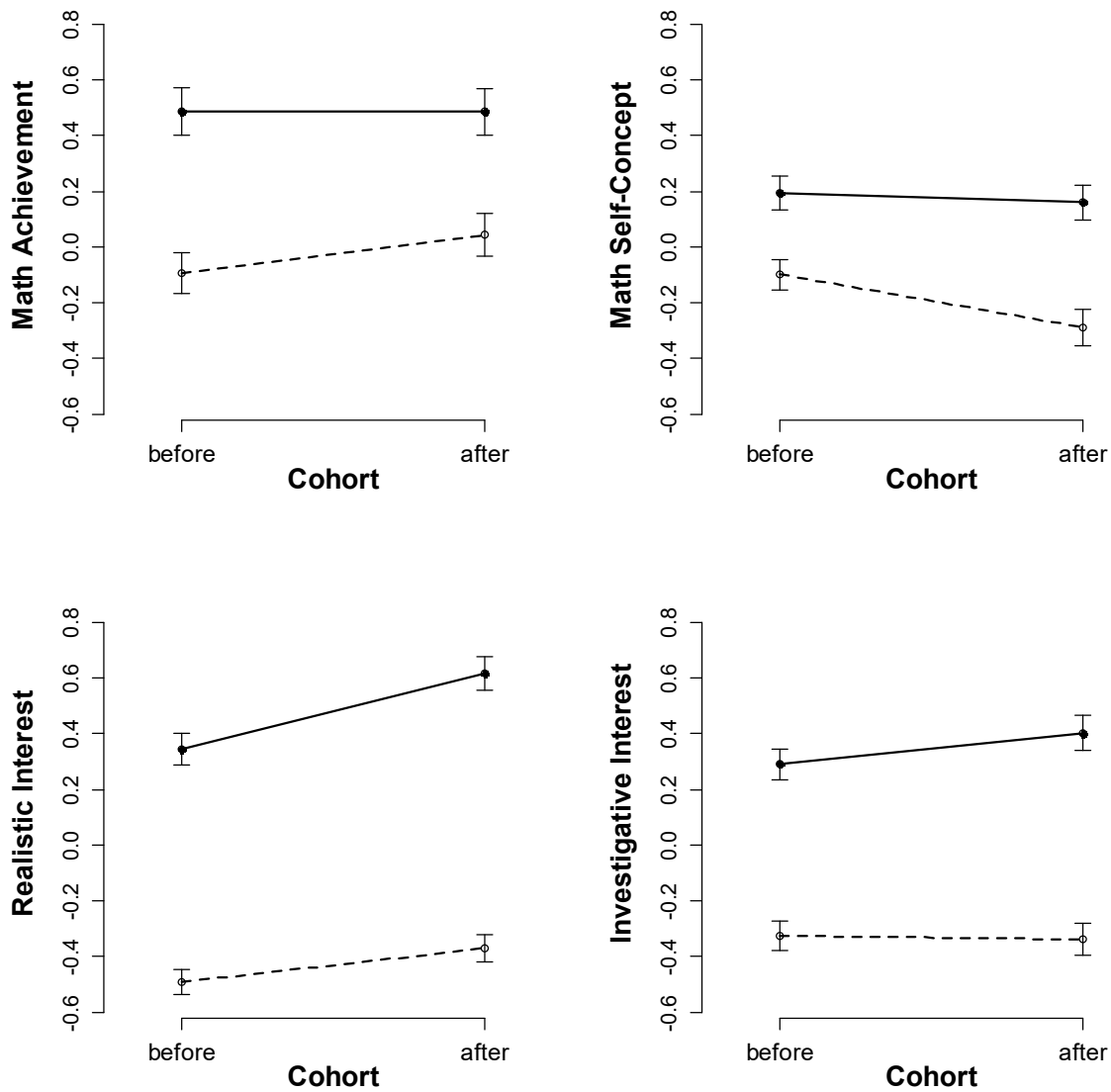


Figure 2. Plots of the moderating effect of gender on the relation between reform and math achievement, math self-concept, realistic interests, and investigative interests with 95% confidence intervals. The dependent variables are presented in standard deviation units. Note that gender differences after the reform were statistically significant for achievement ( $d = -0.44$ ,  $p < .001$ ), math self-concept ( $d = -.045$ ,  $p < .001$ ), realistic interests ( $d = -0.99$ ,  $p < .001$ ), and investigative interests ( $d = -0.74$ ,  $p < .001$ ). Gender differences from before the reform are reported in Tables 3-6.



### *Math Self-Concept*

With regard to math self-concept, we expected a larger gender difference after the reform. In line with our expectations, and as shown in Table 4, the moderating effect of gender on the relation between cohort and self-concept was statistically significant ( $d = -0.16, p = .007, 95\% \text{ CI } [-0.27, -.04]$ ). The larger gender difference after the reform was the result of a statistically significantly lower average math self-concept for young women after the reform ( $d = -0.19, p < .001, 95\% \text{ CI } [-0.27, -0.11]$ ) compared with before the reform. For young men, math self-concept did not differ significantly before versus after the reform ( $d = 0.05, p = .433, 95\% \text{ CI } [-0.18, 0.08]$ ). The other two interaction effects, Cohort  $\times$  School Type ( $d = -0.03, p = .619, 95\% \text{ CI } [-0.16, 0.09]$ ) and Cohort  $\times$  Gender  $\times$  School Type ( $d = 0.11, p = .157, 95\% \text{ CI } [-0.04, 0.27]$ ), were both not statistically significant.

Table 4

#### *Predicting Advanced Math Self-Concept: Results from Multiple Regressions Models*

Predictor	<i>B</i>	<i>p</i>	<i>SE</i>	<i>d</i> <sup>a</sup>
Cohort (T2 = 1)	-.04	.433	0.04	-0.04
Gender (f = 1)	-.29	< .001	0.03	-0.29
HISEI	.01	.370	0.01	0.00
Books	.05	< .001	0.01	0.04
Immigration (=1)	.00	.925	0.03	0.00
Age	-.13	< .001	0.02	-0.18
School type (VS =1)	.06	.131	0.04	0.06
Cohort $\times$ Gender	-.16	< .001	0.06	-0.16
Cohort $\times$ School Type	-.03	.619	0.06	-0.03
Cohort $\times$ Gender $\times$ School Type	.11	.157	0.08	0.11
R <sup>2</sup>		.05		

*Note.* All values are fully standardized. Continuous variables are centered. T2 = TOSCA 2006; HISEI = highest international socioeconomic index; VS = vocational upper secondary school.

<sup>a</sup>The dependent variable is standardized.

### *Realistic and Investigative Vocational Interests*

According to our hypotheses, we expected larger gender differences in realistic and investigative interests after the reform. As reported in Table 5, we found a significant and negative interaction between cohort and gender in predicting realistic vocational interests ( $d = -0.15$ ,  $p = .007$ , 95% CI [-0.26, -0.04]), thus indicating a larger gender difference after the reform than before. This larger gender difference resulted from a significantly higher mean score for young men ( $d = 0.27$ ,  $p < .001$ , 95% [0.19, 0.35]) and a smaller, albeit also significantly higher mean score for young women ( $d = 0.12$ ,  $p < .001$ , 95% [0.05, 0.19]) after the reform (see Figure 2).

Table 5

#### *Predicting Realistic Vocational Interests: Results from Multiple Regressions Models*

Predictor	<i>B</i>	<i>p</i>	<i>SE</i>	<i>d</i> <sup>a</sup>
Cohort (T2 = 1)	.27	< .001	0.04	0.27
Gender (f = 1)	-.84	< .001	0.03	-0.84
HISEI	-.04	.004	0.01	0.00
Books	.04	.001	0.01	0.03
Immigration (= 1)	-.08	.002	0.03	-0.08
Age	-.03	.013	0.01	-0.05
School type (VS = 1)	.09	.099	0.06	0.09
Cohort × Gender	-.15	.007	0.06	-0.15
Cohort × School Type	.00	.932	0.07	0.01
Cohort × Gender × School Type	.00	.948	0.09	0.01
R <sup>2</sup>		.22		

*Note.* All values are fully standardized. Continuous variables are centered. T2 = TOSCA 2006; HISEI = highest international socioeconomic index; VS = vocational upper secondary school.

<sup>a</sup> The dependent variable is standardized.

In addition to realistic vocational interests, we tested for a gender difference in investigative interests (Table 6). Taking a closer look at our results, we found a significant interaction effect (see Figure 2), indicating a larger gender difference in investigative vocational interests after the reform ( $d = -0.12$ ,  $p = .019$ , 95% CI [-0.23, -0.02]). No significant difference between before and after the reform was found for young women in investigative interests ( $d =$

-0.01,  $p = .773$ , 95% CI [-0.09, 0.07]), but young men showed, on average, a higher level of interest after the reform ( $d = 0.12$ ,  $p = .01$ , 95% CI [0.03, 0.21]). For both outcomes, the Cohort  $\times$  School Type interaction and the Cohort  $\times$  Gender  $\times$  School Type interaction were not statistically significant (see Table 6).

The results for the multivariate approach were similar to the results for the univariate approach: The Wald test for the interaction effect was statistically significant,  $\chi^2(12) = 55.06$ ,  $p < .001$ . Furthermore, even after the Benjamini-Hochberg corrections, all interaction effects remained statistically significant in the multivariate and univariate approaches. Overall, we found that the structure coefficients supported our results regarding multiple linear regression models and the interpretation of the relevance of the Cohort  $\times$  Gender interaction for all outcome variables (see Table 7).

Table 6

*Predicting Investigative Vocational Interests: Results from Multiple Regressions Models*

Predictor	<i>B</i>	<i>p</i>	<i>SE</i>	<i>d</i> <sup>a</sup>
Cohort (T2 = 1)	.11	.01	0.04	0.11
Gender (f = 1)	-.62	< .001	0.03	-0.62
HISEI	.00	.745	0.01	0.00
Books	.11	< .001	0.01	0.09
Immigration (= 1)	.01	.668	0.03	0.01
Age	-.03	.045	0.01	-0.04
School type (VS = 1)	.07	.142	0.05	0.07
Cohort $\times$ Gender	-.12	.019	0.05	-0.12
Cohort $\times$ School Type	-.05	.371	0.06	-0.05
Cohort $\times$ Gender $\times$ School Type	.02	.770	0.08	0.02
R <sup>2</sup>		.12		

*Note.* All values are fully standardized. Continuous variables are centered. T2 = TOSCA 2006; HISEI = highest international socioeconomic index; VS = vocational upper secondary school.

<sup>a</sup>The dependent variable is standardized.

Table 7

*Structure Coefficients for Multiple Linear Regression Models*

Predictor	Advanced mathematics	Math self-concept	Realistic interests	Investigative interests
Cohort (T2 = 1)	0.11	-0.18	0.21	0.05
Gender (f = 1)	-0.54	-0.74	-0.96	-0.95
HISEI	0.28	0.22	-0.00	0.15
Books	0.32	0.25	0.03	0.28
Immigration (= 1)	-0.25	-0.14	-0.09	-0.07
Age	-0.49	-0.50	-0.02	-0.03
School type (VS = 1)	-0.70	-0.02	0.08	0.00
Cohort × Gender	-0.23	-0.61	-0.47	-0.55
Cohort × School Type	-0.42	-0.07	0.14	0.00
Cohort × Gender × School Type	-0.47	-0.25	-0.20	-0.28

*Note.* The table displays structure coefficients (e.g., Courville & Thompson, 2001) for each predictor of all four multiple linear regression models. T2 = TOSCA 2006; HISEI = highest international socioeconomic index; VS = vocational upper secondary school.

***Field of Study at University***

Whether or not the upper secondary school reform had an effect on university subject choices was handled as an open research question. Therefore, we did not formulate an explicit hypothesis with regard to this construct. The results presented here are based on an analysis that considered only students who did not intend to become teachers.<sup>2</sup> As reported in Table 8, none of the additional interaction effects were statistically significant. Thus, a potential shift, which would go along with an increase in women enrolling in STEM subjects at university was not found in our data set (Cohort × Gender:  $OR = 1.01$ ,  $p = .838$ , 95% CI [0.86, 1.21]). We further tested for potential differences between students who provided information about their university subject and those who did not. Results revealed that women ( $OR = 0.73$ ,  $p < .001$ )

<sup>2</sup> The pattern of gender differences in the literature varies with respect to different professions within the STEM fields. Whereas a larger percentage of young men than women tend to choose mathematically intensive STEM subjects, gender differences are much less pronounced with regard to STEM teaching professions (Watt, Richardson, & Devos, 2013). To meet this objective, we excluded teaching students from our analysis. However, robustness checks did not reveal any substantial difference between the results of these two groups of students. Furthermore, although men tended to start their studies a bit later (e.g., due to mandatory community or military services), we did not find significant gender differences before and after the reform regarding students who attended university and those who did not.

and students from vocational schools ( $OR = 0.54, p < .001$ ) as well as older students ( $B = -.20, p < .001$ ) were less likely to report their subject, whereas students with a higher HISEI ( $B = .28, p < .001$ ), more books at home ( $B = .33, p < .001$ ), and higher cognitive abilities ( $B = .28, p < .001$ ) reported their subject more often. We controlled for these variables in all analyses. It is important to note that these differences did not differ significantly between the two cohorts, as shown by the Wald test,  $\chi^2(7) = 7.75, p = .356$ .

Table 8

*Predicting Field of Study at University: Results from Multiple Logistic Regressions Models*

Predictor	OR	CI	<i>p</i>
Cohort (T2 = 1)	0.97	0.85 1.11	.702
Gender (f = 1)	0.37	0.32 0.42	< .001
HISEI	0.90	0.82 0.98	.022
Books	0.90	0.82 0.99	.037
Immigration (= 1)	0.97	0.88 1.07	.538
Age	0.83	0.75 0.92	< .001
School type (VS = 1)	1.08	0.90 1.31	.411
Cohort × Gender	1.02	0.86 1.21	.838
Cohort × School Type	1.01	0.86 1.20	.871
Cohort × Gender × School Type	1.01	0.87 1.16	.948
Pseudo-R <sup>2</sup>		.24	

*Note.* The table displays standardized results where mathematics, engineering, computer science, and physics were coded as STEM subjects. Odds ratios significantly larger than 1 indicate a higher likelihood of studying STEM subjects. *OR* = odds ratio; T2 = TOSCA 2006; HISEI = highest international socioeconomic index; VS = vocational upper secondary school.

## Discussion

In the current study, we examined effects of a higher secondary school education reform on STEM-related outcomes in a large and representative sample. The reform is of high theoretical and practical interest because it abolished a prior imbalance between young men and women in taking advanced math courses. High school coursework in math has been shown to be related to STEM career choices as well as to math achievement, math self-concept, and vocational interests, all of which are important predictors of STEM career choices. Therefore, we expected that the effects of the reform on these outcomes would differ by gender. Overall, the results supported most of our predictions. First, there were significant gender differences in all outcomes before the reform, with higher scores for young men than for young women. Second, we found differential effects of the reform for young women and men in all outcomes except field of study at university. However, the direction of the effects differed: The gender difference in math achievement was smaller after the reform, but gender differences in math self-concept and STEM-related vocational interests were even larger after the reform than before. However, the larger gender difference after the reform in math self-concept was based on young women's lower scores, whereas young men's scores did not differ. Also, the greater differences in vocational interests were due to young men's higher interests after the reform, whereas young women's interests were only slightly higher (realistic) or did not differ (investigative). Third, we found no overall effect of the reform on gender differences in the choice of STEM subjects at university.

### *Differential Reform Effects for Young Men and Women*

The effects of the reform on math achievement are in accordance with previous research that reported positive effects of course level on achievement, which can be attributed to more demanding curricula, more teaching time, and larger weights from grades in advanced courses with respect to their contribution to final GPA (e.g., Brunello & Checchi, 2007; Gamoran & Mare, 1989; Hanushek & Wössmann, 2006; Kelly, 2004; Lucas, 2001). Presumably because a larger proportion of young men than women had chosen advanced courses in math before the reform, but all students took the same math course after the reform, young women were able to come closer to young men's math achievement, although there was still a significant gender difference after the reform. In addition, there was a difference in teaching time between the courses before versus after the reform, with more lessons taught per week in the advanced course (five lessons) than in the basic course (three lessons). Although meta-analyses do not suggest a clear pattern with regard to the effects of extended learning time on achievement,

most studies have shown zero to small positive effects (e.g., Patall, Cooper, & Allen, 2010; Scheerens & Hendriks, 2014). Thus, the difference in teaching time might provide a possible explanation for the differential effects of the reform on young women's and men's math achievement. However, we cannot explicitly test for or disentangle the effects of instructional time or course level on our results at this point.

Against this background, the larger gender differences after the reform with respect to math self-concept and STEM-related interests might come as a surprise at first glance. A change in reference group provides a good explanation for the larger gender difference in math self-concept after the reform: It is a common finding that social comparisons are central for the development of students' self-concept. In evaluating their own abilities, individuals refer not only to their own prior achievement in a domain, but also to the level of achievement they perceive in their surroundings (e.g., Marsh, 2005; Niepel, Brunner, & Preckel, 2014; Trautwein et al., 2006). As discussed above, students' achievement differs between advanced and basic courses; thus, both courses provide different frames of reference for social comparisons. Higher course levels are usually associated with negative effects on students' evaluations of their own abilities after individual ability is controlled for (Marsh, 2005; Trautwein et al., 2006). Before the reform, young women tended to choose basic courses in math where they were surrounded by an (on average) a weaker reference group, compared with students in advanced courses. Therefore, they perceived their own math ability in comparison with other, on average, lower achieving classmates. After the reform, all students were instructed at the same course level. Consequently, after the reform, young women could compare their own achievement with the achievement of all other students in their class, which included students with relatively lower achievement but also those with relatively higher achievement. It is therefore likely that the reason why young women's evaluation of their own math abilities was somewhat lower was due to the, on average, higher achieving reference group. There was no significant difference in young men's self-concept after the reform, which can be explained by the proportions of young men in advanced and basic courses before the reform, as they participated in advanced and basic courses in almost equal parts before the reform. According to the literature described above, it is therefore likely that possible effects of course level on young men's self-concept cancelled each other out. These explanations are further supported by the fact that young women's math self-concept in basic courses before the reform was statistically lower, compared with young women's self-concept after the reform ( $d = -0.12$ ,  $p < .001$ ), whereas the reverse was true for young women in advanced courses before the reform ( $d = 0.83$ ,  $p < .001$ ). Furthermore, the difference between young men and women in basic courses was not

statistically significant ( $d = -0.07$ ,  $p = .086$ ), whereas the gender gap for advanced course students was statistically significant, favoring young men ( $d = -0.13$ ,  $p = .001$ ).

In our study, we found larger gender difference after the reform in realistic and investigative interests as well, but in contrast to math self-concept, the greater differences were based on young men's higher levels of interests after the reform, whereas young women showed only slightly higher interests (realistic) or even similar scores (investigative) after the reform. There is a gap in research on how vocational interests might be related to course level. However, as reported in the Introduction, previous research has indicated positive relations between individual levels of math achievement and realistic and investigative interests (Ackerman & Heggestad, 1997; Anthony & Armstrong, 2010). Furthermore, previous research has shown negative effects of the mean level of achievement on domain-specific levels of interest (Köller, Trautwein, Lüdtke, & Baumert, 2006; Schurtz, Pfof, Nagengast, Artelt, 2014; Trautwein et al., 2006) and initial findings with respect to vocational interests. These findings indicate that there might be positive effects of the mean level of math achievement on realistic and investigative vocational interests (Cambria et al., 2016). However, as these findings provide only initial indications on how vocational interests might be related to class level, they enable us to discuss our findings only on a speculative basis. Thereby, one could argue that there might be a positive association between class level and students' realistic and investigative interests, but this association differs by gender, with larger associations for young men than for young women. Previous research on vocational interests has indeed indicated differential associations between ability and vocational interests, although such findings have so far been limited to general cognitive ability and have not been applied to math (e.g., Reeve & Heggestad, 2004). However, more research is needed to explore the relation between class level and vocational interests for young women as well as for young men.

Although we found differential effects of the reform on central predictors of STEM career choices, we found no difference in gender ratios in the numbers of students who chose to study STEM university subjects. There are two aspects to consider when interpreting the absence of effects of the reform on gender differences in STEM university subject choices. First, we found opposite effects of the reform on gender differences in four important predictors of STEM career choices: Whereas differences in math achievement were eliminated, differences in math self-concept and both interest facets were larger. Consequently, it is possible that the effects of the reform on the predictors cancelled each other out, with the consequence that no effect on the choice of STEM subjects remained. Second, choosing a university subject is a complex process that involves numerous factors (see Schoon & Eccles,



2014). The reform influenced students' upper secondary high school coursework, but it did not directly affect other structural factors or the wider context they grew up in, such as their family structure, the role models they perceived, or their stereotypical views of STEM professions.

### *Practical Implications*

Our study adds to the increasing number of studies that have found intended as well as unintended effects of educational reforms. In fact, educational policy reforms do not necessarily improve educational outcomes but can instead result in numerous unintended consequences. In addition, the aspects of the reforms most likely interact differently with different student characteristics, even if such aspects are well-structured and carefully planned (Gross, Booker, & Goldhaber, 2009). For instance, studies by Gross et al. (2009), Domina, McEachin, Penner, and Penner (2015), and Lee and Reeves (2012) showed that school reforms could have differential effects for minority students (e.g., African American and Hispanic students) or could vary for specific school districts. The results show that school reforms can have differential effects on several outcomes, and such outcomes can even differ for particular subgroups such as young women and men; not every well-intentioned reform will reach all goals, and some might even backfire.

Unintended consequences of reforms can be attributed to, amongst other factors, the complex nature of establishing and especially of implementing reforms (e.g., McLaughlin, 1987; Young & Lewis, 2015) in the education sector as a "loosely coupled system" (Porter, Fusarelli, & Fusarelli, 2015, p. 114). Conversely, with regard to the current study, one might argue that the higher achievement and realistic interests that came with this reform came at a price—a lower math self-concept for young women—which had to be expected given the change in reference group.

Although high school coursework is central to young people's career choices, and although we found differential effects of the reform on central predictors of STEM career choices for young men and women, we did not find effects of the reform on gender differences in the choice of STEM university subjects, which indicates that one single reform might not significantly influence students' career choices. In the complex context that young people grow up in, there is a cumulative process of multiple experiences that shape young people's academic attitudes and behavior, such as career choice (cf. Schoon & Eccles, 2014). Influencing gender differences in high school course selection by restricting choice options might be one way to balance some gender differences in the STEM context, namely, gender differences in math achievement. However, reforming course choice options does not necessarily impact any of the

reasons for why young women are less likely to choose advanced math courses than young men (e.g., gender stereotypes, different expectancies of parents, teachers, peers; cf. Schoon & Eccles, 2014; Wigfield & Eccles, 2000). Such high school reforms might therefore be “too little too late” to increase gender equity in the STEM fields in a meaningful and sustainable way. Furthermore, although course-taking gaps in other countries have narrowed in recent decades (e.g., Domina & Saldana, 2012; Osborne & Dillon, 2008), subsequent changes in STEM career plans do not seem to be of considerable size (Jerrim & Schoon, 2014).

### ***Limitations and Further Research***

The current study demonstrates that intensifying school curricula and providing equal access to advanced courses “does not necessarily level the [educational] playing field” with regard to all important outcomes (Domina & Saldana, 2012, p. 688). Although our investigation was based on a strong data set, some limitations should be kept in mind when interpreting the results. First, our results were limited to the domain of math. Math is a key domain within the STEM fields (Ma & Johnson, 2008; Sells, 1980), and math achievement, self-concept, and interests are very important for math-intensive STEM career choices (e.g., Parker et al., 2012; Schoon & Eccles, 2014). Nevertheless, other STEM domains such as physics or chemistry are also meaningful for later math-intensive STEM career choices (e.g., Hazari, Sonnert, Sadler, & Shanahan, 2010), and gender differences in such high school courses are often even larger than in math (e.g., NSF, 2015). Evaluating the effects of a reform on central STEM outcomes in these domains might therefore provide additional information about effects on important predictors of math-intensive STEM career choices.

Second, the current study was based on cross-sectional data. According to Shadish, Cook, and Campbell (2002), quasi-experiments lack “random assignment of units to conditions” (p. 104), which may lead to selection bias. We attempted to address these challenges by using a lagged cohort control design that should have led to relatively small selection differences between cohorts (drawn from the same schools). We additionally checked for potential differences between cohorts and used covariates to control for these. Third, besides these methodological issues, there are other possible reasons for the results that we found. Our results may be explained by the multidimensional structure of the reform. As stated by Malen and Knapp (1997), “policy takes many forms, performs many functions, and begets many effects,” which is why “it is difficult to get a fix on the boundaries, let alone the ‘workings’ of a policy or a set of policies” (p. 419). In our case, as mentioned, not only did time vary between the groups before and after the reform, but the reference groups and course levels also varied.

Therefore, the effects of the reform cannot be directly attributed to one specific aspect or mechanism of the reform in a causal manner but must be interpreted from within the multilayered framework of the entire policy reform.

However, as society is constantly changing, it would be reasonable to expect main and interaction effects that indicate the increased participation of young women in STEM classes because they are now as able to do so as young men. However, the results of our study instead indicate the opposite pattern. Regarding this point, it is also important to mention that society's growing interest and all resulting efforts had already increased in the beginning of this century and not just between these two cohorts in particular (National-State-Commission for Educational Planning and Scientific Promotion, 2002; NSF, 2000). In addition, we checked closely whether any other educational reforms had been implemented between the two cohorts, but this was not the case.

Further research should address the question of whether effects, such as the drop in self-concept, can be found in different subsamples. This refers to questions such as whether such effects can be found for all young women or only the subsample of those who would have chosen basic courses if they had been allowed to, and whether similar effects can be found for young men who would have chosen basic courses if they had been allowed to. Fourth, our results are limited to the issue of gender differences in STEM career choices at the end of secondary education, and more research is needed to explore the complex pattern of gender differences in the STEM fields throughout students' educational careers. In our study, we focused on important predictors of STEM career choices as well as students' choice of university major in the STEM fields. Therefore, our results provide insights into various effects of the reform on central STEM outcomes. However, regarding the issue of gender differences in the STEM area, not only do women tend to choose such majors less frequently than their male counterparts, but women also drop out of university at higher rates (Ackerman, Kanfer, & Beier, 2013; Perez-Felkner, McDonald, & Schneider, 2014). Considering social comparison processes, one could possibly argue that women entering the STEM fields are likely to experience such comparison processes during their studies, where they need to deal with other high-achieving students. Experiencing such comparison processes at an earlier point in high school might therefore make women less likely to pursue such careers and—consequently—less vulnerable to dropping out of STEM fields during college. Furthermore, prior work on the development of interest suggests that interest takes time to develop (see Hidi & Renninger, 2006) and that such a change in upper secondary high school coursework as investigated in the present study might be less related to students' vocational interests than to their achievement

and self-concept or that such effects might take longer. In this study, we investigated effects of changes in coursework requirements on students' interests 1.5 years after they started taking these high school courses. It might be the case that such a time period is insufficient to fully study effects on interest developments and that effects would be different or more pronounced if more time could have elapsed between when the students began taking these high school courses and the measurement point. Further research spanning a longer time frame is needed to test such propositions as well as to develop more potent remedies for the gender differences that still exist.

### *Conclusion*

The present study was aimed at taking a closer look at effects of high school coursework on gender differences in math-intensive STEM fields. To this end, we investigated effects of a statewide educational reform in Germany with a large representative sample. The reform required all students to take advanced courses in math and eliminated the prior imbalance between young men and women in choosing such courses. Our results showed that it is crucial to take multiple aspects into consideration in order to obtain insights into possible differential effects of changes in coursework requirements. Although requiring all students to take advanced math courses appears to be adequate for eliminating gender differences in math achievement, it seems that young women were not aware of this: Young men and women's achievement differed less after the reform, but young women showed an even lower self-concept compared with young men than had been there before the reform. With respect to realistic and investigative interests, although young women showed no or only slightly higher interests after the reform, the interests of young men were substantially higher after the reform. Mechanisms that ensure that all students will benefit in comparable ways from such school reforms and impede negative side effects, such as those found for young women's self-concept, should be a primary focus of future research.

### References

- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*(2), 219–245. doi:10.1037/0033-2909.121.2.219
- Ackerman, P. L., Kanfer, R., & Beier, M. E. (2013). Trait complex, cognitive ability, and domain knowledge predictors of baccalaureate success, STEM persistence, and gender differences. *Journal of Educational Psychology*, *105*(3), 911–927. doi:10.1037/a0032338
- Alarcon, G. M., & Edwards, J. M. (2013). Ability and motivation: Assessing individual factors that contribute to university retention. *Journal of Educational Psychology*, *105*(1), 129–137. doi:10.1037/a0028496
- Anthony, S. F., & Armstrong, P. I. (2010). Individuals and environments: Linking ability and skill ratings with interests. *Journal of Counseling Psychology*, *57*(1), 36–51. doi:10.1037/a0018067
- Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Structural Equation Modeling*, *12*, 411–434. [https://doi.org/10.1207/s15328007sem1203\\_4](https://doi.org/10.1207/s15328007sem1203_4)
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, *57*(1), 289–300. <https://doi.org/10.2307/2346101>
- Bergmann, C., & Eder, F. (2005). *AIST-R: Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R) - Revision*. [General Interest Structure Test and Environmental Structure Test - Revision]. Göttingen: Beltz.
- 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>
- Brunello, G., & Checchi, D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy*, *22*(52), 782–861. <https://doi.org/10.1111/j.1468-0327.2007.00189.x>
- Cambria, J., Brandt, H., Nagengast, B., & Trautwein (2016). Vocational interests. The impact of class achievement and gender. *Manuscript in preparation*.
- Camp, A. G., Gilleland, D., Pearson, C., & Putten, J. V. (2009). Women's path into science and engineering majors: A structural equation model. *Educational Research and Evaluation*, *15*(1), 63–77. <https://doi.org/10.1080/13803610802591725>
- Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current*

- Directions in Psychological Science*, 19(5), 275–279.  
<https://doi.org/10.1177/0963721410383241>
- Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, 135(2), 218–261. <https://doi.org/10.1037/a0014412>
- Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., & Lin, S. S. J. (2013). The relationship between academic self-concept and achievement: A multicohort–multioccasion study. *Learning and Individual Differences*, 23, 172–178. <https://doi.org/10.1016/j.lindif.2012.07.021>
- Chmielewski, A. K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal*, 50(5), 925–957.  
<https://doi.org/10.3102/0002831213489843>
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc. <https://doi.org/10.1234/12345678>
- Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles:  $\beta$  is not enough. *Educational and Psychological Measurement*, 61(2), 229–248. <https://doi.org/10.1177/0013164401612006>
- Denissen, J. J. A., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development*, 78(2), 430–447. <https://doi.org/10.1111/j.1467-8624.2007.01007.x>
- Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). Aiming high and falling short: California's eighth-grade algebra-for-all effort. *Educational Evaluation and Policy Analysis*, 37(3), 275–295. <https://doi.org/10.3102/0162373714543685>
- Domina, T., & Saldana, J. (2012). Does raising the bar level the playing field? Mathematics curricular intensification and inequality in American high schools, 1982-2004. *American Educational Research Journal*, 49(4), 685–708.  
<https://doi.org/10.3102/0002831211426347>
- Eccles, J. S., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. C. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–121). San Francisco, CA: W. H. Freeman & Co.
- 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>
- Else-Quest, N. M., Mineo, C. C., & Higgins, A. (2013). Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychology of Women Quarterly*, 37(3), 293–309. <https://doi.org/10.1177/0361684313480694>
- Enders, C. K. (2010). *Applied missing data analysis. Methodology in the social sciences*. New York, NY: Guilford Press.
- Evans, M. D. R., Kelley, J., Sikora, J., & Treiman, D. J. (2010). Family scholarly culture and educational success: Books and schooling in 27 nations. *Research in Social Stratification and Mobility*, 28(2), 171–197. <https://doi.org/10.1016/j.rssm.2010.01.002>
- Federal statistical office (Ed.). (2008). *Bildung und Kultur - Studierende an Hochschulen Wintersemester 2007/2008 [Education and culture - Students at universities in wintersemester 2007/2008]. Fachserie 11, Reihe 4.1*. Wiesbaden: Statistisches Bundesamt.
- Gamoran, A., & Mare, R. D. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality? *American Journal of Sociology*, 94(5), 1146. <https://doi.org/10.1086/229114>
- Gamoran, A., Porter, A. C., Smithson, J., & White, P. A. (1997). Upgrading high school mathematics instruction: Improving learning opportunities for low-achieving, low-income youth. *Educational Evaluation and Policy Analysis*, 19(4), 325–338.  
<https://doi.org/10.3102/01623737019004325>
- Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.  
[https://doi.org/10.1016/0049-089X\(92\)90017-B](https://doi.org/10.1016/0049-089X(92)90017-B)
- Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research*, 25(3), 201–239. <https://doi.org/10.1006/ssre.1996.0010>
- Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology*, 107(3), 663–677.  
<https://doi.org/10.1037/edu0000003>
- Gottfredson, L. S. (1986). Occupational aptitude patterns map: Development and implications for a theory of job aptitude requirements. *Journal of Vocational Behavior*, 29(2), 254–291. [https://doi.org/http://dx.doi.org/10.1016/0001-8791\(86\)90008-4](https://doi.org/http://dx.doi.org/10.1016/0001-8791(86)90008-4)

- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*, 549–576.  
<https://doi.org/10.1146/annurev.psych.58.110405.085530>
- Gross, B., Booker, T. K., & Goldhaber, D. (2009). Boosting student achievement: The effect of comprehensive school reform on student achievement. *Educational Evaluation and Policy Analysis*, *31*(2), 111–126. <https://doi.org/10.3102/0162373709333886>
- Hanushek, E. A., & Wössmann, L. (2006). Does educational tracking affect performance and inequality? Differences- in-differences evidence across countries. *The Economic Journal*, *116*(510), C63–C76. <https://doi.org/10.1111/j.1468-0297.2006.01076.x>
- Hazari, Z., Sonnert, G., Sadler, P. M., & Shanahan, M.-C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, *47*(8), 978–1003.  
<https://doi.org/10.1002/tea.20363>
- Henson, R. K. (2006). Effect-size measures and meta-analytic thinking in counseling psychology research. *The Counseling Psychologist*, *34*(5), 601–629.  
<https://doi.org/10.1177/0011000005283558>
- Hidi, S., & Ainley, M. (2002). Interest and adolescence. In F. Pajares & T. C. Urdan (Eds.), *Academic motivation of adolescence* (pp. 247–275). Greenwich, Conn: Information Age Pub.
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127. [https://doi.org/10.1207/s15326985ep4102\\_4](https://doi.org/10.1207/s15326985ep4102_4)
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives*, *2*(3), 172–177.  
<https://doi.org/10.1111/j.1750-8606.2008.00061.x>
- Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, *6*(1), 35–45. <https://doi.org/10.1037/h0040767>
- Holland, J. L. (1966). *The psychology of vocational choice: A theory of personality types and model environments*. Waltham, MA: Blaisdell.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, Fla.: Psychological Assessment Resources.
- Huang, J. L., & Pearce, M. (2013). The other side of the coin: Vocational interests, interest differentiation and annual income at the occupation level of analysis. *Journal of Vocational Behavior*, *83*(3), 315–326. <https://doi.org/10.1016/j.jvb.2013.06.003>
- Humphreys, L. G., & Yao, G. (2002). Prediction of graduate major from cognitive and self-



- report test scores obtained during the high school years. *Psychological Reports*, *90*(1), 3–30. <https://doi.org/10.2466/PRO.90.1.3-30>
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*(2), 139–155. <https://doi.org/10.1037/0033-2909.107.2.139>
- Hyde, J. S., Fennema, E., Ryan, M., Frost, L. A., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect. *Psychology of Women Quarterly*, *14*, 299–324. <https://doi.org/10.1111/j.1471-6402.1990.tb00022.x>
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Diversity: Gender similarities characterize math performance. *Science*, *321*(5888), 494–495. <https://doi.org/10.1126/science.1160364>
- Jaccard, J., Wan, C. K., & Turrisi, R. (1990). The detection and interpretation of interaction effects between continuous variables in multiple regression. *Multivariate Behavioral Research*, *25*(4), 467–478. [http://dx.doi.org/10.1207/s15327906mbr2504\\_4](http://dx.doi.org/10.1207/s15327906mbr2504_4)
- Jansen, M., Lüdtke, O., & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology*, *46*, 116–127. <https://doi.org/10.1016/j.cedpsych.2016.05.004>
- Jansen, M., Scherer, R., & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology*, *41*, 13–24. <https://doi.org/10.1016/j.cedpsych.2014.11.002>
- Jerrim, J., & Schoon, I. (2014). Do teenagers want to become scientists? A comparison of gender differences in attitudes toward science, career expectations, and academic skills across 29 countries. In I. Schoon & J. S. Eccles (Eds.), *Gender differences in aspirations and attainment* (pp. 203–223). Cambridge, U.K.: Cambridge University Press.
- Kelly, S. (2004). Are teachers tracked? On what basis and with what consequences. *Social Psychology of Education*, *7*, 55–72. <https://doi.org/10.1023/B:SPOE.0000010673.78910.f1>
- Köller, O., Baumert, J., & Schnabel, K. U. (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>
- Köller, O., Trautwein, U., Lüdtke, O., & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe [On the

- interplay of academic achievement, self-concept, and interest in upper secondary schools]. *Zeitschrift Für Pädagogische Psychologie*, 20(1), 27–39.  
<https://doi.org/10.1024/1010-0652.20.12.27>
- Köller, O., Watermann, R., Trautwein, U., & Lüdtke, O. (2004). *Wege zur Hochschulreife in Baden-Württemberg: TOSCA - eine Untersuchung an allgemein bildenden und beruflichen Gymnasien [Ways to qualification for university entrance. TOSCA - A study at traditional and vocational Gymnasium schools]*. Opladen: Leske + Budrich.
- Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education*, 14, 23–40.  
<https://doi.org/10.1007/BF03173109>
- Lee, J., & Reeves, T. (2012). Revisiting the Impact of NCLB High-Stakes School Accountability, Capacity, and Resources: State NAEP 1990-2009 Reading and Math Achievement Gaps and Trends. *Educational Evaluation and Policy Analysis*, 34(2), 209–231. <https://doi.org/10.3102/0162373711431604>
- Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45, 79–122. <http://dx.doi.org/10.1006/jvbe.1994.1027>
- Lippa, R. (1998). Gender-related individual differences and the structure of vocational interests: the importance of the people-things dimension. *Journal of Personality and Social Psychology*, 74(4), 996–1009. <http://dx.doi.org/10.1037/0022-3514.74.4.996>
- Low, K. S. (2009). *Patterns of mean-level changes in vocational interests: A quantitative review of longitudinal studies*. Unpublished doctoral dissertation of the University of Illinois at Urbana-Champaign.
- Low, K. S., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: A quantitative review of longitudinal studies. *Psychological Bulletin*, 131(5), 713–737.  
<https://doi.org/10.1037/0033-2909.131.5.713>
- Lubinski, D., & Benbow, C. P. (2006). Study of mathematically precocious youth after 35 years. *Special Section: Doing Psychological Science*, 1(4), 316–345.  
<https://doi.org/10.1111/j.1745-6916.2006.00019.x>
- Lucas, S. R. (2001). Effectively maintained inequality: Education transitions, track mobility, and social background effects. *American Journal of Sociology*, 106(6), 1642–1690.  
<https://doi.org/10.1086/321300>
- Lumley, T. (2014). “*Survey: Analysis of complex survey samples*”. *R package version 3.30*.

- Ma, X., & Johnson, W. (2008). Mathematics as the critical filter: Curricular effects on gendered career choices. In H. M. G. Watt & J. S. Eccles (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 55–83). Washington, DC: American Psychological Association.
- Malen, B., & Knapp, M. (1997). Rethinking the multiple perspectives approach to education policy analysis: Implications for policy-practice connections. *Journal of Education Policy, 12*(5), 419–445. <https://doi.org/10.1080/0268093970120509>
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*(1), 129–149. <https://doi.org/10.3102/00028312023001129>
- Marsh, H. W. (1992). *Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: A test manual and a research monograph*. Sydney, New South Wales, Australia: Macarthur.
- Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *Zeitschrift Für Pädagogische Psychologie, 19*(3), 119–129. <https://doi.org/10.1024/1010-0652.19.3.119>
- Marsh, H. W. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, UK: British Psychological Society.
- Marsh, H. W., Abduljabbar, A. S., Morin, A. J. S., Parker, P. D., Abdelfattah, F., Nagengast, B., & Abu-Hilal, M. M. (2015). The big-fish-little-pond effect: Generalizability of social comparison processes over two age cohorts from Western, Asian, and Middle Eastern Islamic countries. *Journal of Educational Psychology, 107*(1), 258–271. <https://doi.org/10.1037/a0037485>
- Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., Nagengast, B., ... Abu-Hilal, M. M. (2014). The internal/external frame of reference model of self-concept and achievement relations: Age-cohort and cross-cultural differences. *American Educational Research Journal, 52*(1), 168–202. <https://doi.org/10.3102/0002831214549453>
- Marsh, H. W., & O'Neill, R. (1984). Self Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21*(2), 153–174. <https://doi.org/0.1111/j.1745-3984.1984.tb00227.x>
- Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107–123. [https://doi.org/10.1207/s15326985ep2003\\_1](https://doi.org/10.1207/s15326985ep2003_1)

- Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, *44*(3), 631–669.  
<https://doi.org/10.3102/0002831207306728>
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. <http://dx.doi.org/10.1111/j.1467-8624.2005.00853.x>
- 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>
- McLaughlin, M. W. (1987). Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, *9*(2), 171–178.  
<https://doi.org/10.3102/01623737009002171>
- McNeish, D., Stapleton, L. M., & Silverman, R. D. (2016). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*. *Advance Online Publication*.  
<https://doi.org/10.1037/met0000078>
- Ministry of Education and Cultural Affairs, Youth and Sport Baden-Württemberg. (2002). *Die neue gymnasiale Oberstufe in Baden-Württemberg. Infodienst Schule spezial [The new upper secondary school in Baden-Württemberg]*. Stuttgart: Ministerium für Kultus Jugend und Sport Baden-Württemberg.
- Mullis, I. V. S., Martin, M. O., Beaton, A., Gonzalez, E., Kelly, D., & Smith, T. (1998). *Mathematics and science achievement in the final year of secondary school: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: TIMSS International Study Center, Boston College.
- Muthén, B. O., & Muthén, L. K. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, 267–316. <http://dx.doi.org/10.2307/271070>
- Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., & Eccles, J. S. (2008). Gendered high school course selection as a precursor of gendered careers: The mediating role of self-concept and intrinsic value. In J. S. Eccles & H. M. G. Watt (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural*

- influences* (pp. 115–143). Washington, DC, US: American Psychological Association.
- Nagy, G., & Husemann, N. (2010). Berufliche Interessen vor und nach dem Übergang in die gymnasiale Oberstufe [Vocational interests before and after the transition to Gymnasium upper secondary school]. In W. Bos, E. Klieme, & O. Köller (Eds.), *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert* [Learning opportunities and development of competencies at school. Commemorative volume for Jürgen Baumert] (pp. 59–84). Münster, New York, München, Berlin: Waxmann.
- Nagy, G., Neumann, M., Trautwein, U., & Lüdtke, O. (2010). Voruniversitäre Mathematikleistungen vor und nach der Neuordnung der gymnasialen Oberstufe in Baden-Württemberg [Advanced math before and after the rearrangement of upper secondary school in Baden- Württemberg]. In U. Trautwein, M. Neumann, G. Nagy, O. Lüdtke, & K. Maaz (Eds.), *Schulleistungen von Abiturienten. Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand [School achievement of upper secondary school students. The rearranging upper secondary school on the trial]*. (pp. 147–180). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation, 12*(4), 323–345.  
<https://doi.org/10.1080/13803610600765687>
- 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 Science Foundation (NSF). (2000). *Summary Report on the Impact Study of the National Science Foundation's Program for Women and Girls*. Retrieved from <https://www.nsf.gov/pubs/2001/nsf0127/nsf0127.pdf>
- National Science Foundation (NSF). (2013). *Women, minorities, and persons with disabilities in science and engineering*. Retrieved from [http://www.nsf.gov/statistics/wmpd/2013/pdf/nsf13304\\_full.pdf](http://www.nsf.gov/statistics/wmpd/2013/pdf/nsf13304_full.pdf)
- National Science Foundation (NSF). (2015). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2015 Digest. Special Report NSF*.  
<https://doi.org/Special Report NSF 11-309>
- Niepel, C., Brunner, M., & Preckel, F. (2014). The longitudinal interplay of students'

- academic self-concepts and achievements within and across domains: Replicating and extending the reciprocal internal/external frame of reference model. *Journal of Educational Psychology*, *106*(4), 1170–1191. <https://doi.org/10.1037/a0036307>
- Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science*, *7*(4), 384–403. <https://doi.org/10.1177/1745691612449021>
- Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.
- OECD. (2010). *OECD information technology outlook 2010*. Paris: OECD Publishing. [https://doi.org/10.1787/it\\_outlook-2010-en](https://doi.org/10.1787/it_outlook-2010-en)
- Osborne, J., & Dillon, J. (2008). *Science education in europe: Critical reflections. A report to the Nuffield Foundation*. Retrieved from [http://efepereth.wdfiles.com/local--files/science-education/Sci\\_Ed\\_in\\_Europe\\_Report\\_Final.pdf](http://efepereth.wdfiles.com/local--files/science-education/Sci_Ed_in_Europe_Report_Final.pdf)
- Parker, P. D., 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* (pp. 224–246). Cambridge, U.K.: Cambridge University Press.
- Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: A multicontext study. *Developmental Psychology*, *48*(6), 1629–1642. <https://doi.org/10.1037/a0029167>
- Pässler, K., Beinicke, A., & Hell, B. (2014). Gender-related differential validity and differential prediction in interest inventories. *Journal of Career Assessment*, *22*(1), 138–152. <https://doi.org/10.1177/1069072713492934>
- Patall, E.A., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: A systematic review of research (1985-2009). *Review of Educational Research*, *80*(3), 401–436. <https://doi.org/10.3102/0034654310377086>
- Patrick, L., Care, E., & Ainley, M. (2011). The relationship between vocational interests, self-efficacy, and achievement in the prediction of educational pathways. *Journal of Career Assessment*, *19*, 61–74. <https://doi.org/10.1177/1069072710382615>
- Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology*, *106*(1), 315–329. <https://doi.org/10.1037/a0034027>

- Perez-Felkner, L., McDonald, S.-K., & Schneider, B. (2014). What happens to high-achieving females after high school? Gender and persistence on the postsecondary STEM pipeline. In I. Schoon & J. S. Eccles (Eds.), *Gender differences in aspirations and attainment* (pp. 285–320). Cambridge, U.K.: Cambridge University Press.
- Porter, R. E., Fusarelli, L. D., & Fusarelli, B. C. (2015). Implementing the common core: How educators interpret curriculum reform. *Educational Policy*, *29*(1), 111–139. <https://doi.org/10.1177/0895904814559248>
- R Development Core Team. (2014). *R*. Vienna, Austria: R foundation for statistical computing.
- Reeve, C. L., & Heggstad, E. D. (2004). Differential relations between general cognitive ability and interest-vocation fit. *Journal of Occupational and Organizational Psychology*, *77*, 385–402. <https://doi.org/10.1348/0963179041752673>
- Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Educational Progress assessments. *Journal of Educational Psychology*, *107*(3), 645–662. <https://doi.org/10.1037/edu0000012>
- Rolfhus, E. L., & Ackerman, P. L. (1996). Self-report knowledge: At the crossroads of ability, interest, and personality. *Journal of Educational Psychology*, *88*(1), 174–188. <https://doi.org/10.1037/0022-0663.88.1.174>
- Rounds, J., & Su, R. (2014). The nature and power of interests. *Current Directions in Psychological Science*, *23*(2), 98–103. <https://doi.org/10.1177/0963721414522812>
- Sáinz, M., & Eccles, J. S. (2012). Self-concept of computer and math ability: Gender implications across time and within ICT studies. *Journal of Vocational Behavior*, *80*(2), 486–499. <http://dx.doi.org/10.1016/j.jvb.2011.08.005>
- Schäfer, J. L. (1997). *Analysis of incomplete multivariate data: Monographs on statistics and applied probability: Vol. 72*. New York, NY: Chapman & Hall.
- Scheerens, J., & Hendriks, M. (2014). State of the art of time effectiveness. In J. Scheerens (Ed.), *SpringerBriefs in Education. Effectiveness of Time Investments in Education* (pp. 7–29). Cham, Germany: Springer International Publishing. [http://dx.doi.org/10.1007/978-3-319-00924-7\\_2](http://dx.doi.org/10.1007/978-3-319-00924-7_2) Schiefele,
- Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In S. Hidi (Ed.), *The role of interest in learning and development* (pp. 183–212). Book Section, Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

- Schnabel, K. U., Alfeld, C., Eccles, J. S., Köller, O., & Baumert, J. (2002). Parental influence on students' educational choices in the United States and Germany: Different ramifications—same effect? *Journal of Vocational Behavior*, *60*(2), 178–198.  
<https://doi.org/10.1006/jvbe.2001.1863>
- Schoon, I., & Eccles, J. S. (Eds.). (2014). *Gender differences in aspiration and attainment: A life course perspective*. Cambridge, U.K.: Cambridge University Press.
- Schoon, I., Ross, A., & Martin, P. (2007). Science related careers: Aspirations and outcomes in two British cohort studies. *Equal Opportunities International*, *26*(2), 129–143.  
<https://doi.org/10.1108/02610150710732203>
- Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction*, *34*, 32–41.  
<https://doi.org/10.1016/j.learninstruc.2014.08.001>
- Schwanzer, A. D., Trautwein, U., Lüdtke, O., & Sydow, H. (2005). Entwicklung eines Instruments zur Erfassung des Selbstkonzepts junger Erwachsener [Development of a new instrument to assess self-concept of young adults]. *Diagnostica*, *51*(4), 183–194.  
<https://doi.org/10.1026/0012-1924.51.4.183>
- Sells, L. W. (1980). Mathematics: The invisible filter. *Engineering Education*, *70*, 340–341.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-experimental designs for generalized causal inferences*. Berkeley: Houghton Mifflin.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Los Angeles, CA: Sage.
- Stapleton, L. M., Yang, J. S., & Hancock, G. R. (2016). Construct meaning in multilevel settings. *Journal of Educational and Behavioral Statistics*, *41*, 481–520.  
<https://doi.org/10.3102/1076998616646200>
- Strong, E. K. J. (1943). *Vocational interests of men and women*. Stanford: Stanford University Press.
- 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*, 1–20. <https://doi.org/10.3389/fpsyg.2015.00189>
- Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin*, *135*(6), 859–884.  
<https://doi.org/10.1037/a0017364>
- Trautwein, U., Köller, O., Lüdtke, O., & Baumert, J. (2005). Student tracking and the



- powerful effects of opt-in courses on self-concept: Reflected-glory effects do exist after all. In H. W. Marsh, R. Craven, & D. M. McInerney (Eds.), *New frontiers for self research* (pp. 307–327). Greenwich, CT: Information Age Press.
- Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology, 98*(4), 788–806. <https://doi.org/10.1037/0022-0663.98.4.788>
- Trautwein, U., Neumann, M., Nagy, G., Lüdtke, O., & Maaz, K. (Eds.). (2010). *Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand [School achievement of upper secondary school students. The rearranging upper secondary school on the trial]*. Wiesbaden: VS Verl. für Sozialwissenschaften.
- Updegraff, K. A., Eccles, J. S., Barber, B. L., & O'Brien, K. M. (1996). Course enrollment as self-regulatory behavior: Who takes optional high school math courses? *Learning and Individual Differences, 8*, 239–259. [https://doi.org/10.1016/S1041-6080\(96\)90016-3](https://doi.org/10.1016/S1041-6080(96)90016-3)
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174–1204. <https://doi.org/10.1037/a0036620>
- Wagner, W., Rose, N., Dicke, A.-L., Neumann, M., & Trautwein, U. (2014). Alle alles lehren – Schulleistungen in Englisch, Mathematik und den Naturwissenschaften vor und nach der Neuordnung der gymnasialen Oberstufe in Sachsen [Teaching everyone everything: Student achievement in English, mathematics and the natural sciences before and after Saxony's upper secondary school reform]. *Zeitschrift Für Erziehungswissenschaft, 17*(2), 345–369. <https://doi.org/10.1007/s11618-014-0492-7>
- Watt, H. M. G., & Eccles, J. S. (Eds.). (2008). *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences*. Washington, DC: American Psychological Association. <https://doi.org/10.1037/11706-000>
- Watt, H. M. G., Eccles, J. S., & Durik, A. M. (2006). The leaky mathematics pipeline for girls. *Equal Opportunities International, 25*(8), 642–659. <https://doi.org/10.1108/02610150610719119>
- Watt, H. M. G., Richardson, R. W., & Devos, C. (2013). (How) Does gender matter in the choice of a STEM teaching career and later teaching behaviours? *International Journal of Gender, Science and Technology, 5*(3), 187–206. Retrieved from <http://genderandset.open.ac.uk/index.php/genderandset/article/viewArticle/331>
- Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes.

*Developmental Review*, 30(1), 1–35. <https://doi.org/10.1016/j.dr.2009.12.001>

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., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children’s competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, 89(3), 451–469. <https://doi.org/10.1037/0022-0663.89.3.451>

Wigfield, A., Tonks, S., & Klauda, S. T. (2009). Expectancy-value theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 55–57). New York, NY: Routledge.

Young, T., & Lewis, W. D. (2015). Educational policy implementation revisited. *Educational Policy*, 29(1), 3–17. <https://doi.org/10.1177/0895904815568936>

## Study 3:

### It Takes Two:

# Expectancy-Value Constructs and Vocational Interests Predict STEM Careers Differently and Differ Between Men and Women

Wille, E., Cambria, J., Stoll, G., Nagengast, B., & Trautwein, U. (2017). It takes two: Expectancy-value constructs and vocational interests predict STEM careers differently and differ between men and women. *Manuscript submitted for publication*.

This research was partially funded by the state government of Baden-Württemberg.

**The following manuscript has not been accepted and published yet. This version of the manuscript has been submitted to *Developmental Psychology* on August 16, 2017.**

***Developmental Psychology* will have the copyright of the final version of the article, if it will be accepted. The version displayed here may not exactly replicate the final version published in the journal. It is not the copy of record.**

### Abstract

Eccles et al.'s (1983) expectancy-value theory and Holland's (1997) theory of vocational interests are two of the most prominent theories for investigating STEM career choices. Constructs from both models have been highly predictive in explaining STEM outcomes for men and women. However, their relative predictive validity is so far unclear. Aiming to fill this gap, we reanalyzed longitudinal data from 4,984 students in Germany at the end of high school and 2 years later after the transition to university. We investigated different STEM outcomes (i.e., math achievement, choice of different STEM majors at university) by applying multiple regression analyses and multiple logistic regression analyses. Results showed that both expectancy-value constructs and vocational interests predicted these outcomes. Their relative predictive power, however, depended on the outcome: Whereas expectancy-value constructs were better predictors of math achievement, vocational interests were better predictors of the choice of a STEM major at university. In addition, we investigated the role of gender in such predictions: Gender differences at the mean levels of expectancy-value constructs and vocational interests partly explained gender differences in math achievement and the choice of a STEM major. Furthermore, multiple-group analyses revealed differences in the relative power of vocational interests in predicting the outcomes. Results therefore suggest that expectancy-value constructs and vocational interests contribute differently to young men's and women's STEM career pathways. Implications include the careful consideration of the constructs with respect to the outcomes under investigation in future research and with respect to outcomes targeted in interventions.

*Keywords:* expectancies of success, task values, vocational interests, math achievement, STEM university majors, gender

## **It Takes Two: Expectancy-Value Constructs and Vocational Interests Predict STEM Careers Differently and Differ Between Men and Women**

Majoring in science, technology, engineering, or mathematics (STEM) at university provides high-status career options. People working in STEM earn on average almost 30% more than people working outside of STEM, and working in STEM occupations is also associated with lower unemployment rates compared with other occupations (Langdon, McKittrick, Beede, Khan, & Doms, 2011; Noonan, 2017). Given such career prospects, it makes sense to ask why more students do not choose STEM majors, and why young women in particular tend to choose majors outside of STEM.

The present study contributes to answering these questions. It focuses on two “interest theories” that are both known to be associated with the choice of university major but that surprisingly have not been brought together until now: Eccles et al.’s (1983) expectancy-value theory and Holland’s (1997) theory of vocational interests. Both theories have been shown to be highly predictive in explaining career choices on the basis of motivation and interests and have also been frequently applied to explain gender differences in such choices (Schoon & Eccles, 2014; Su, Rounds, & Armstrong, 2009). However, they emphasize different aspects: whereas expectancy-value theory focuses on expectancies and subjective task values related to specific educational domains such as mathematics, Holland’s model emphasizes more general levels of interest related to different fields of professions, such as investigative and realistic interests, which describe activities related to the STEM fields (Eccles & Wigfield, 2002; Holland, 1997). Yet, little is known about the interrelations of expectancies and subjective task values with vocational interests as well as their relative value in predicting different STEM careers for men and women because the constructs have thus far been investigated separately.

Using a large longitudinal data set of German pre-university students, we aimed to contribute to filling this research gap. First, we examined whether the constructs could predict math achievement, which is one of the most central cognitive predictors of STEM career choices (Parker, Nagy, Trautwein, & Lüdtke, 2014; Sells, 1980). Next, we looked at the power of the expectancy-value constructs and vocational interests to predict the choice of a STEM major at university. Because STEM majors include a wide range of subjects, we investigated different subgroups of majors; that is, we examined STEM university majors compared with majors outside of STEM and mathematically intensive STEM majors compared with STEM majors in the area of life sciences. Drawing on the existing literature, we expected that both expectancy-value constructs and vocational interests would predict the indicators of STEM careers. Given the different levels of domain-specificity of the constructs, however, we

expected that the predictive validity would depend on the level of specificity of the criterion. Thus, we propose that both are necessary to better understand the complex pattern of STEM career choices. As a third step, we investigated the role of gender in such predictions in order to obtain more insights into potential mechanisms that lead to women's underrepresentation in mathematically intensive STEM fields. On the one hand, we examined whether the gender differences in expectancy-value constructs and vocational interests could explain gender differences in STEM careers, and on the other hand, we explored whether associations between expectancy-value constructs and vocational interests with STEM careers would differ by gender.

***Eccles et al.'s Expectancy-Value Theory and Holland's Theory of Vocational Interests:  
Similarities and Differences***

Eccles et al.'s (1983) expectancy-value theory and Holland's (1997) theory of vocational personalities and work environments are among the most widely used frameworks for investigating young people's career choices in developmental, work, and vocational psychology, and research based on these theories has contributed considerably to the understanding of STEM career choices for men and women. Both theories explain career choices based on interest-related constructs and emphasize the motivational function of their constructs, which are conceptualized as dispositions for specific activities. They propose that these dispositions, in turn, guide individuals' educational choices and choices of career environments (Holland, 1997; Wigfield & Eccles, 2000). However, there are conceptual differences between the theories.

Expectancy-value theory includes the influences of different domain-specific components (i.e., expectancies for success and subjective task values) as predictors of academic decision making that furthermore lead to specific educational outcomes. Expectancies for success are defined as individuals' beliefs about personal success in a future task (Eccles & Wigfield, 2002). Subjective task values include four components: intrinsic value, attainment value, utility value, and cost (Eccles, 2005). Intrinsic value is defined as the enjoyment a person experiences when performing the activity or the person's subjective interests in a domain (Wigfield & Cambria, 2010). Intrinsic value is thus similar to other motivational constructs such as intrinsic motivation defined by Deci and Ryan (1985) or interest as defined by Schiefele (1999) or Renninger and Hidi (2011). Attainment value indicates how personally important it is for a person to do well on a task, and utility value indicates how well a task relates to the individual's feelings of how useful an activity is for his or her subjective goals, such as

educational or career goals (Eccles, 2005). Finally, cost describes negative factors of values by reflecting negative emotions (e.g., performance anxiety), the effort that is necessary to succeed in the task, and opportunity costs (i.e., losing other opportunities as a consequence of choosing one option instead of others; Eccles & Wigfield, 2002).

In expectancy-value theory, a comprehensive model of different relations between expectancies and subjective task values with outcomes as well as potential influences on the previous ones are proposed. Both expectancies and values, however, are assumed to be the most proximal predictors of later achievement and educational choices (Wigfield et al., 2015). In this sense, expectancies and subjective task values are considered highly domain-specific and have traditionally been examined by aligning domain-specific motivation constructs with domain-specific outcomes (e.g., Schoon & Eccles, 2014).

Unlike expectancy-value theory, which focuses on several components within one domain, Holland's model of vocational interests focuses on interests in different domains to explain a person's preference for specific careers. The theory distinguishes between six different vocational interests (i.e., realistic, investigative, artistic, social, enterprising, and conventional interests), which are assumed to represent preferences for activities across the entire range of different occupations (Holland, 1959, 1997). With respect to the STEM fields, investigative interests are particularly important as they refer to preferences for activities in science and research, including intellectual, analytical, logical, and creative investigations. People with investigative interests strive for the examination and understanding of biological and cultural phenomena by engaging in systematic observation and research (Holland, 1997). Realistic interests are also associated with STEM fields, albeit not with STEM fields in general but rather with disciplines in the natural sciences and engineering. Realistic interests describe preferences for typical work activities in mechanical and technical areas. People with realistic interests prefer systematic approaches in work and like to work with materials, tools, and machines (Holland, 1997). In addition, for STEM disciplines in the field of life sciences (e.g., medicine), social interests are also important because social interests reflect preferences for social activities that include helping other people (Holland, 1997).

Holland proposed that individuals strive to find environments that correspond with their vocational interests, leading to career choices that are based on these interests (Holland, 1997). Each of the interests represents specific attitudes, skills, and preferences for certain activities or situations in which these activities are located (Holland, 1997). Thus, comparable to expectancy-value theory, Holland (1997) also assumed that interests correspond with specific

values, goals, ability beliefs, and abilities. However, such associations are less explicitly formulated in Holland's model than in expectancy-value theory.

Another difference between Eccles et al.'s and Holland's models is the level of specificity of the constructs. Expectancies and subjective task values are traditionally considered at a domain-specific level (e.g., expectancy and value beliefs in mathematics, which are related to STEM careers; Wigfield et al., 2015). By contrast, vocational interests typically refer to tasks and activities on a more general level and capture a broad set of activities and preferences that can be associated with typical activities in various occupational fields of professions (e.g., realistic and investigative interests for the STEM area; Holland, 1997).

### ***Expectancies, Task Values, Vocational Interests, and STEM Career Choices***

A wide range of research has shown the high predictive validity of expectancy-value constructs and vocational interests, respectively, for different career pathways including STEM careers (e.g., Humphreys & Yao, 2002; Schoon & Eccles, 2014). Expectancy-value constructs as well as vocational interests are associated with choice of an university major (e.g., Humphreys & Yao, 2002; Lauermann, Tsai, & Eccles, 2017) and also with math achievement, which is one of the most central cognitive predictors of STEM career choices and also acts as a gate keeper because STEM careers often require specific math prerequisites (Ackerman & Heggestad, 1997; Parker et al., 2014; e.g., Sells, 1980). However, based on the different levels of domain-specificity, the theories have traditionally been applied with different areas of focus.

As expectancy-value theory is traditionally used in research in school contexts, previous research has shown that domain-specific expectancies and subjective task values are strong predictors of performance in the respective domain (Cole, Bergin, & Whittaker, 2008; Denissen, Zarrett, & Eccles, 2007; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005). Concerning choices, ample research has shown that expectancies and subjective task values are closely linked to course selection in high school (e.g., Nagy et al., 2008) as well as the aspiration to choose (e.g., Lauermann et al., 2017; Parker et al., 2014) and the actual choice of a university major (e.g., Parker et al., 2012, 2014; Wang & Kenny, 2014) across domains. For example, math expectancy and value beliefs are closely linked to the choice of a mathematically intensive STEM university major (e.g., mathematics or engineering), and science expectancy and value beliefs are linked to the choice of a science university major (Parker et al., 2014; Riegler-Crumb, Moore, & Ramos-Wada, 2011; Wang & Kenny, 2014). In comparing expectancies and subjective task values, previous research has provided evidence that subjective task values are



more closely linked to students' choices, whereas expectancies are more strongly related to students' performance (e.g., Meece, Wigfield, & Eccles, 1990).

Vocational interests have mostly been applied to vocational choices and have been used to predict career choices in a more comprehensive manner, such as choice of major across the entire range of different subjects (e.g., Humphreys & Yao, 2002; Su & Rounds, 2015). Research involving vocational interests has therefore also tended to focus on how vocational interests are related to intelligence rather than to domain-specific math abilities (see Ackerman & Heggestad, 1997). However, some research has indicated that realistic and investigative interests are positively correlated with math achievement, whereas artistic and social interests are negatively correlated with math achievement, although these negative associations have not been found consistently (Ackerman & Heggestad, 1997; Warwas, Nagy, Watermann, & Hasselhorn, 2009). With respect to career choices, research has shown strong relations between vocational interests and students' choice of university major in different fields of study such as engineering, science, humanities, and social sciences (Humphreys & Yao, 2002; Päßler & Hell, 2012). For the STEM fields, positive associations between realistic and investigative interests and STEM majors in general have been reported (Nagy, 2006; Rolfhus & Ackerman, 1996; Roloff Henoch, Klusmann, Lüdtke, & Trautwein, 2015). Specifically comparing different STEM fields, a study by Pässler and Hell (2012) indicated that realistic interests were positively related to the choice of mathematically intensive STEM fields using life science majors as the reference group, whereas investigative and social interests were negatively related to such majors.

Despite the importance of expectancies, subjective task values, and vocational interests for math achievement and STEM university majors, the constructs have usually been investigated separately so far. Some research has examined relations between vocational interests and domain-specific expectancies. In a study by Volodina and Nagy (2016), for instance, math self-concept, which is empirically highly related to expectancy beliefs (e.g., Eccles & Wigfield, 1995), was found to be positively correlated with realistic ( $r = .32$ ) and investigative interests ( $r = .24$ ). Furthermore, they found similar correlations for artistic interests with English and German self-concepts. However, to the best of our knowledge, interrelations of subjective task values and vocational interests have not been studied thus far. Yet, in order to understand the complex process of career pathways, it is essential to know the extent to which important predictors are related to each other and how they are relatively associated with different STEM outcomes.

### *The Role of Gender in STEM Career Choices*

Women are less likely to choose mathematically intensive STEM majors at university such as physics or engineering than their male counterparts (National Science Board, 2016). In explaining such gender differences, research based on expectancy-value theory and Holland's theory of vocational interests has shown that women's underrepresentation in mathematically intensive STEM fields can be linked to gender differences in expectancy-value constructs and vocational interests, respectively (Schoon & Eccles, 2014; Su & Rounds, 2015).

Regarding expectancies and subjective task values, early emerging stereotypical patterns of gender differences have been found with higher scores for boys and men in mathematics, whereas girls and women usually report higher expectancies and values in verbal domains (for recent reviews, see e.g., Meece, Glienke, & Burg, 2006; Wigfield et al., 2015). Higher beliefs in mathematics have been found for boys' and men's expectancy beliefs and related constructs compared with the beliefs of girls and women even when achievement was controlled for (e.g., Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Marsh & Yeung, 1998). Furthermore, higher math values for boys and men than for girls and women have been found for math task values in general and for intrinsic value (e.g., Frenzel, Goetz, Pekrun, & Watt, 2010; Watt, 2004), whereas gender differences with respect to attainment value, utility value, and cost have been found to be somewhat different for specific subfacets (Gaspard, Dicke, Flunger, Schreier, et al., 2015; Steinmayr & Spinath, 2010; Watt, 2004). Gender differences have also been reported for vocational interests (e.g., Päßler & Hell, 2012; Su et al., 2009). A meta-analysis by Su, Rounds, and Armstrong (2009) of studies involving 81 different samples showed large advantages for men with respect to STEM-related interests (i.e., realistic and investigative interests). Furthermore, they found higher artistic, social, and conventional interests for women than for men.

In addition to gender differences in the mean levels of expectancies, task values, and vocational interests, there are also findings that suggest differences between men and women in the relative importance of these constructs for STEM outcomes (e.g., Larson, Wu, Bailey, Borgen, & Gasser, 2010; Päßler & Hell, 2012; Watt et al., 2012). The existing literature on the moderating role of gender on relations between the constructs and STEM outcomes, however, is limited and has provided mixed evidence. With respect to expectancy-value constructs, some researchers have found slightly different associations with STEM careers between women and men (Korhonen, Tapola, Linnanmäki, & Aunio, 2016; Watt et al., 2012, 2016). In general, however, these associations seem to be comparable for men and women (e.g., Guo, Parker, Marsh, & Morin, 2015; Marsh & Craven, 2006; Sáinz & Eccles, 2012; Simpkins, Fredricks, &

Eccles, 2012; Tynkkynen, Tolvanen, & Salmela-Aro, 2012). Research with respect to the relative importance of vocational interests for STEM outcomes is scarce, but there is some support for the moderating role of gender in the relation between vocational interests and career choices (Larson et al., 2010; Päßler, Beinicke, & Hell, 2014; Päßler & Hell, 2012). Päßler and Hell (2012), for instance, found that realistic interests more strongly predicted the choice of an engineering major for women than for men, whereas social interests predicted the choice of a science major only for women.

### *The Present Study*

In the present study, we aimed to extend previous research on the roles of expectancies, subjective task values, and vocational interests in STEM career choices and on gender differences in these fields. Expectancy-value constructs and vocational interests are both highly predictive of STEM careers, and both have often been used to explain gender differences in these fields. Somewhat surprisingly, the constructs have usually been studied separately, and their relative predictive validity is so far unclear. To understand the complex process of career pathways, however, it is necessary to know how expectancy-value constructs and vocational interests are related to each other and their relative power in predicting STEM careers. In an effort to fill this research gap, we reanalyzed longitudinal data of 4,984 students who were in their final year in high school at the first measurement point and at university at the second measurement point 2 years later. We compared the validity of expectancy-value constructs and vocational interests in predicting different indicators of STEM careers (i.e., math achievement as a central cognitive predictor of STEM career choices and the choice of STEM university majors). For the latter, we investigated different classifications of STEM majors because STEM majors include a wide range of different subjects: We compared STEM majors in general with majors outside of STEM, and we compared mathematically intensive STEM majors with STEM majors in the area of life sciences. Due to the different levels of domain-specificity of expectancy-value constructs and vocational interests, we expected the predictive validity of the constructs to be related to the level of specificity of the criterion investigated. We expected that expectancy-value constructs and vocational interests would both predict math achievement (Hypothesis 1a). Furthermore, we expected the expectancy-value constructs to be stronger predictors of math achievement than we expected vocational interests to be (Hypothesis 1b). Concerning the individual predictors, we expected math expectancy and value beliefs as well as realistic and investigative interests to predict math achievement (Hypothesis 1c), but again, we expected the math expectancy-value constructs (particularly expectancies) to be stronger

predictors of math achievement compared with realistic and investigative interests (Hypothesis 1d).

Regarding the choice of STEM university majors, we expected that expectancy-value constructs and vocational interests would both predict the choice of a STEM university major and the choice of a mathematically intensive STEM major (Hypothesis 2a). Furthermore, we expected higher predictive validity for vocational interests than for expectancy and value beliefs as a set of predictors (Hypothesis 2b). Concerning the individual predictors, we expected that math values and STEM-related vocational interests (i.e., especially realistic and investigative interests) would each predict the choice of a STEM university major and a mathematically intensive STEM major (Hypothesis 2c). For the choice of a mathematically intensive STEM major compared with a STEM major in the area of life sciences, we also expected that social interest would be a predictor but a negative one (also Hypothesis 2c).

Because there are existing gender differences in mathematically intensive STEM fields, we subsequently examined the role of gender in the prediction of STEM careers along with two questions: First, how do gender differences in expectancy-value constructs and vocational interests explain gender differences in math achievement and the choice of a STEM university major? Second, does the relative predictive power of expectancy-value constructs and vocational interests vary between men and women? Concerning the first question, we expected that gender differences in expectancy-value constructs and vocational interests would partially explain gender differences in the STEM careers (Hypothesis 3). Due to a lack of previous research, we explored whether the constructs were comparably able to explain gender differences in the constructs and how they jointly contributed to gender differences in STEM careers. Concerning the second question, our research was exploratory because only a few studies have looked at the different roles of these constructs for men and women in the STEM context, and various classifications of STEM outcomes have been used.

## **Method**

### ***Sample***

Data were drawn from the ongoing, longitudinal, multicohort study “Transformation of the Secondary School System and Academic Careers” (TOSCA), hosted by the University of Tübingen (see Köller, Watermann, Trautwein, & Lüdtke, 2004; Trautwein, Neumann, Nagy, Lüdtke, & Maaz, 2010). The TOSCA study is a strong data set for exploring the process of transitioning from high school to university as it consists of students from university-track schools in their final year in high school as well as several years afterward in a large German

state (Baden-Württemberg). The school-leaving certificate from university-track schools entitles students to attend university, and usually at least 80% of students in university-track schools enter higher education. For the present study, the second TOSCA cohort was used because subjective task values were administered only in this cohort. In this cohort, at the first measurement point in spring 2006, trained research assistants collected data in 157 academic-track secondary schools. At the second measurement point, which was 2 years later, questionnaires were sent to participants who had agreed to continue to participate in the study; this led to a response rate of 57% at Time 2. Students participated in the study on a voluntary basis.

Concerning the first measurement point, we considered data from students who participated in the assessment of the achievement test and the questionnaire ( $N = 4,984$ ). For this time point, the mean age of participants was 19.4 years ( $SD = 0.69$ ), and in accordance with the usual composition of academic track schools in Germany, there were slightly more girls in this sample (55.4% female); 58% of the students attended a traditional secondary school (Gymnasium), and 42% attended a vocational secondary school.

Data from this sample were used to investigate whether the constructs of interest could predict math achievement, which was also assessed at the first measurement point. In the analysis of the different STEM majors, students were included only if they stated that they had chosen one of the investigated majors at the second measurement point ( $N = 2,032$  and  $N = 833$  students). The sample sizes for the different analyses and percentages of women in these samples are reported in Table 1.

Table 1

*Sample Sizes for all Dependent Variables*

Dependent Variable	<i>N</i>	% Women
Math achievement	4,984	55.4
Choice of a STEM major <sup>a</sup>	2,032	57.8
Choice of a math intensive STEM major <sup>b</sup>	833	42.5

<sup>a</sup>Mathematics, natural sciences, and engineering were coded as STEM subjects, whereas linguistics and cultural studies, law, business, and social sciences were included in the reference group. <sup>b</sup>Mathematics, physics, computer science, and engineering were coded as math intensive STEM subjects, whereas biology and medicine were included in the reference group.

### *Instruments*

We used a set of independent variables (i.e., self-concepts as indicators of expectancy beliefs and subjective task values for mathematics and English, vocational interests), dependent variables (i.e., math achievement and choice of university major), and control variables (i.e., gender, socioeconomic background, immigration status, type of school, English achievement), all of which have been shown to be related to career choices (Parker et al., 2014; Wang, Degol, & Ye, 2015). All independent and control variables were assessed at the first measurement point, which was the final year of high school. Math achievement was also assessed at the first measurement point, whereas choice of university major was assessed 2 years later at the second measurement point.

#### *Expectancy beliefs*

Mathematics and English self-concept scales were used to operationalize expectancy beliefs. Although expectancy-value theory conceptually differentiates between expectancies for success (referring to upcoming tasks) and competence beliefs such as self-concept (referring to individuals' evaluation of their current competence on the basis of their interpretation of previous competence; Wigfield & Cambria, 2010), the two concepts are empirically highly correlated (e.g., Eccles & Wigfield, 1995). Therefore, they have often been used interchangeably (e.g., Nagengast et al., 2011; Trautwein, Lüdtke, Schnyder, & Niggli, 2006).

All self-concept items are presented in the Appendix. Math self-concept was measured with four items from the Self Description Questionnaire III (Marsh, 1992; SDQ III; Marsh & O'Neill, 1984). We used the German translation by Schwanzer, Trautwein, Lüdtke, and Sydow (2005). Because English as a foreign language was not included in the SDQ III, four items for English self-concept were created on the basis of the SDQ III math and verbal self-concept scales as well as a well-known German self-concept inventory (see Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006). All items were measured with a 4-point Likert scale ranging from 1 (*completely disagree*) to 4 (*completely agree*). The scales showed high internal consistency (Cronbach's  $\alpha \geq .94$ ).

#### *Subjective task values*

Math and English value scales were assessed on the basis of Eccles et al.'s expectancy-value theory (Wigfield & Eccles, 2000) with most items adapted from previous large-scale studies (see Marsh et al., 2005; Trautwein et al., 2006; Wigfield & Eccles, 2000). All items are reported in the Appendix. The wording was usually the same for mathematics and English such

that only the subject name changed; only three items differed marginally in their wording. The items covered all four conceptual dimensions of task values (intrinsic, attainment, and utility value and cost). The four dimensions form empirically distinct factors as determined by factor analyses, and other studies have employed them separately (e.g., Trautwein et al., 2012). In the present study, however, we were interested in the predictive validity of task values as a set rather than in differential relations of specific dimensions. For reasons of simplicity, we therefore combined the intrinsic, attainment, and utility value items into one scale, as they were highly interrelated. The final combined value scale consisted of 12 items for each domain. Cost was included separately in the analyses and was assessed with two items. As robustness checks, we re-ran all analyses with separate value scales as well. These analyses yielded a similar pattern of results, and therefore, we report the analyses with the combined value scale. All items were measured with a 4-point Likert scale ranging from 1 (*completely disagree*) to 4 (*completely agree*). All scales showed good internal consistency (Cronbach's  $\alpha > .80$ ).

#### *Vocational interests*

Vocational interests were assessed with the Revised General Interest Structure test (GIST-R; Bergmann & Eder, 2005), which is based on Holland's model. Participants were asked to rate how much they were interested in specific activities on a 5-point Likert scale, ranging from 1 (*not interested at all*) to 5 (*very interested*). The test consists of 60 items overall, that is, 10 items from each of the six dimensions of vocational interests: realistic (e.g., "Working with machines or technical devices"), investigative (e.g., "Dealing with unexplored things"), artistic (e.g., "Writing stories or reports"), social (e.g., "Teaching or educating others"), enterprising (e.g., "Guiding a group at work"), and conventional (e.g., "Carrying out work that requires precision and persistence"). All scales showed high internal consistency (Cronbach's  $\alpha > .85$ ).

There are different analytical approaches in research on vocational interests and their predictive validity for performance and choices. Many researchers have focused on the fit between a person's vocational interests and (assumed) characteristics of career environments such as STEM majors (see Brown & Gore Paul A., 1994; Nye, Su, Rounds, & Drasgow, 2017). However, others have used individuals' vocational interests as predictors of performance and choices without considering environmental characteristics in the analyses (see Stoll & Trautwein, 2017; Van Iddekinge, Roth, Putka, & Lanivich, 2011). In this paper, we followed the latter approach by using individuals' vocational interests as predictors rather than the fit to be able to estimate the predictive validity of specific interests that are assumed to be important

for the choice of STEM majors. In addition, this approach enhances the comparability of the results for the expectancy-value constructs and vocational interests.

### *Math achievement*

Math achievement was assessed with an advanced math test that was based on items from the Third International Mathematics and Science Study (TIMSS; Mullis, Martin, & Foy, 2008). The advanced math test contained 68 items from the fields of (a) Numbers, Equations and Functions, (b) Analysis, (c) Geometry, (d) Propositional logic and Proofs, and (e) Probability and Statistics. A multimatrix design was applied to decide how to administer the items. This design allowed us to estimate scores for each student on every dimension, although students responded only to a subset of items in one of four test manuals. Item response theory was used to scale the items to account for the multimatrix design.

### *University major*

Students' choice of university major was assessed with an open-ended format, and the answers were then coded according to the official classification system of the Federal Statistical Office (Fachserie 11, Federal Statistical Office, 2008). In this classification system, majors were classified on different taxonomic levels: Ten subject groups are differentiated at the first level (e.g., linguistics and cultural studies, mathematics and natural sciences, engineering), 60 study areas at the second level (e.g., mathematics or biology), and all specific subjects at the third level. Two dummy variables were created on the basis of this classification: (a) the choice of STEM subject (mathematics, natural sciences, medicine, engineering) versus other fields of study (linguistics and cultural studies, law, business, and social sciences) based on subject groups, and (b) the choice of mathematically intensive STEM subjects (mathematics, physics, computer science, engineering) versus life sciences (biology, medicine) based on study areas. In Germany, students can choose to study some majors as a single-subject program or as a two-subject or even three-subject program (e.g., mathematics, physics, biology). Other majors, however, can be studied only as a single-subject program (e.g., medicine). For the presented analyses, we used information about students' first and the second subjects. In addition, we specified alternative codings, where only the first subject was used to calculate the dummy variables. The general pattern of the results was similar for all of these analyses.<sup>1</sup>

---

<sup>1</sup> In addition, we used the aspiration to choose a certain university major stated at the end of high school as a criterion variable instead of students' choice 2 years after graduation. In general, the results followed the same pattern, although there were tendencies toward more predictor variables being significantly associated with the aspiration to choose (mathematically intensive) STEM majors and slightly stronger associations between the



### *Covariates*

*English achievement.* We controlled for English achievement using a shortened version of the Test of English as a Foreign Language (TOEFL). The test consisted of 70 items.

*Type of school.* We considered type of school in all of our analyses using a dummy-coded variable (general vs. vocational Gymnasium) because academic outcomes usually differ across students from different types (Trautwein et al., 2010).

*Socioeconomic background.* Socioeconomic background was measured with information about parents' highest occupation, which was coded in accordance with the International Standard Classification of Occupation (ISCO-88). The ISCO values were in turn converted into values from the International Socio-Economic Index of Occupational Status (ISEI) 88 (Ganzeboom & Treiman, 1996). The highest ISEI (i.e., HISEI) value of the students' parents was used to characterize the students' socioeconomic background.

*Immigration background.* Students with one or two parents who were not born in Germany were coded as students with an immigration background.

### *Statistical Analyses*

We analyzed the data by applying a series of multiple regressions for math achievement. To predict students' university major (dichotomous variable), we ran a series of logistic regression analyses. For each criterion variable, four different models were calculated: Model I contained the covariates gender, type of school (dummy-coded, vocational Gymnasium = 1), socioeconomic background (HISEI), immigration background (dummy-coded, immigration background = 1), English achievement, and the dependent variables university major and math achievement. We controlled for these covariates to eliminate the influence of these potential confounders and to increase the precision of our estimation. Model II consisted of subjective task values and self-concepts for mathematics and English. Model III contained the six vocational interests. The covariates from Model I were also included in Models II and III. We calculated Model IV as a full model including subjective task values, self-concepts, vocational interests, and the covariates in order to explore the relative predictive power of these variables in predicting the criterion variables.

For the multiple regression analyses, the dependent and all continuous independent variables were standardized, and thus the coefficients were fully standardized. All reported

---

predictor variables with the aspiration to choose a certain major compared with the actual choice of major. Nevertheless, we present results on the actual choice of university major in this paper, as this actually marks an individual's entry into a STEM career, but the results for the aspiration to choose a certain major are available in the supplemental material.

coefficients for the logistic regressions were in logit form, which is the natural logarithm of the odds of studying a (mathematically intensive) STEM major in our study (controlling for all other variables in the model; Cohen, Cohen, West, & Aiken, 2003). In order to report meaningfully interpretable coefficients, we standardized all continuous independent variables, and thus the logistic regression coefficients were standardized as well. In logistic regressions, estimates are affected by the degree of unobserved heterogeneity in the model. Therefore, they cannot be compared across different models or different samples. Standardization of the estimates is one possible way to make coefficients in different models more comparable (Mood, 2010). We furthermore report odds ratios (*ORs*), which are assumed to be easier to interpret than slope coefficients (e.g., Hosmer & Lemeshow, 2000; Long, 1997). In our study, the *ORs* indicate the change in the ratio of the odds of studying a (mathematically intensive) STEM major for a one-unit change in the predictor variable by the odds of studying a major included in the reference group for a one-unit change in the predictor variable, adjusted for all other predictors in the model. *ORs* higher than 1 indicate higher odds for the criterion, and values lower than 1 indicate lower odds for the criterion (e.g., Hosmer & Lemeshow, 2000; Long, 1997).

As model fit indices, we provide R-square values for the multiple regressions and Nagelkerke's pseudo-R-square for the logistic regressions. In logistic regression analyses, there is no statistic that is equivalent to the R-squared estimate in multiple regression analyses for calculating the variance explained by a model. However, R-square analogues, such as pseudo-R-squared estimates can be used to compare models, especially when the compared models consist of different predictors but use the same sample and the same dependent variable. This was the case in our study with respect to the different models for each criterion variable (Long, 1997; Menard, 2000). Nagelkerke's pseudo-R-squared estimates are related to the R-square values from OLS regression and indicate the improvement in the model compared with the null model. Comparable to the R-square values from OLS regression, possible values range from 0 to 1, with 1 indicating a perfect prediction of the criterion (Menard, 2000).

We used *F* tests to test differences in explained variance between the models for each criterion variable for the multiple regression analyses. For the logistic regression analyses, we applied likelihood ratio tests.

In order to investigate differences in the predictive power of the independent variables between women and men, we computed multiple-group regression analyses with gender as the grouping variable. Wald tests were used to test for significant differences in the individual parameters between men and women.

All analysis were computed in Mplus 7.3 (Muthén & Muthén, 2012).

#### *Clustered structure*

The data had a clustered structure with students nested in schools (i.e., students from the same school are usually more similar to each other than they are to students from other schools). Because the clustered structure of the data violates the assumption of independence of observations, it has to be considered so that the standard errors will not be underestimated (Snijders & Bosker, 2012). We addressed the nested data structure by estimating robust standard errors (Muthén & Muthén, 2012).

#### *Missing data*

There were different patterns of missingness in the data we used. At the first measurement point, task values were assessed only in a subsample of  $N = 2,030$  students (planned missing data design; Enders, 2010). Furthermore, some data were missing due to incomplete responses. Response rates therefore ranged from 0% to 60% (see Table 2). To deal with the missing data, we applied the full information maximum likelihood (FIML) approach implemented in Mplus 7.3 (Muthén & Muthén, 2012). This procedure, which is based on maximum likelihood estimation, is highly preferable to most other procedures that can be applied to handle the missing data problem (Schafer & Graham, 2002). The FIML approach uses information from all of the variables included in the model to estimate the model parameters (Schafer & Graham, 2002). In order to increase the efficacy of FIML, we included variables that were not considered in the respective model as auxiliary variables (Enders, 2010). As robustness checks, we also ran the analyses using only the subsample for which task values were assessed at the first measurement point. The results followed the same pattern and are presented in the supplemental materials.

At the second measurement point, some data were missing due to study attrition and nonresponses to single items (57%, see sample description). In the analyses involving data from the second measurement point, we used only the subsample of students who had indicated their choice of major (i.e., STEM majors or majors from the reference group). To deal with missing data in the predictor variables, we also applied the FIML approach as implemented in Mplus 7.3 (Muthén & Muthén, 2012). We again included variables that were not considered as predictors in the respective model as auxiliary variables to increase the efficacy of FIML (Enders, 2010).

## Results

### *Descriptive Analyses*

The means and standard deviations for all scales separated by gender are presented in Table 2. We found significant gender differences for all STEM-related variables. Men scored higher on the combined value and self-concept variables and lower on cost in mathematics. Gender differences favoring men were also found for realistic and investigative interests. Conversely, women scored higher on social interests than men. Furthermore, gender differences favoring men were also found for math achievement and the choice of a (mathematically intensive) STEM major.

The correlations of all study variables are presented in Table 3. Concerning associations between expectancy-value constructs in mathematics and vocational interests, we found that math self-concept and task values had moderate positive correlations with realistic and investigative interests. The strongest correlations were found between the combined math value and realistic or investigative interest variables, whereas math cost and self-concept had somewhat weaker correlations with these interest dimensions.

### *Prediction of Math Achievement*

In the first step, we investigated the predictive power of the expectancy-value constructs and vocational interests in predicting math achievement by applying a series of multiple regression analyses. The results are reported in Table 4. In line with Hypothesis 1a, each set of predictors reliably predicted math achievement while the covariates were controlled for (see Models II and III). Nevertheless, the  $R^2$  values for the different models showed that the predictive power differed for the constructs as proposed in Hypothesis 1b: The set of predictors based on expectancy-value theory was more predictive than the set of vocational interests (see the  $R^2$  values for Models II and III). In addition, the difference between the  $R^2$  values for Models II and IV was smaller than the difference between those from Models III and IV, indicating that vocational interests added only a little predictive power over and above the expectancy-value constructs, whereas vice versa, the expectancy-value constructs added predictive power over vocational interests.  $F$  tests, however, yielded significant differences between Models IV and II,  $F(90, 189) = 6.32, p < .05$ , and Models IV and III,  $F(90, 189) = 271.68, p < .05$ .<sup>2</sup>

---

<sup>2</sup>We additionally checked whether this pattern held when Models II and IV were estimated without the self-concept variables, that is, Model II consisting of values and covariates and Model IV consisting of values, vocational interests, and covariates. It held, although the differences were smaller ( $R^2_{\text{Model II}} = .430$ ;  $R^2_{\text{Model III}} = .304$ ;  $R^2_{\text{Model IV}} = .432$ ).

Table 2

*Descriptive Statistics for all Study Variables Separated by Gender*

Variable	N	Response rates in %	Men		Women		$d^a$	$d$	
			M	SD	M	SD		95% CI	
School type (% VG)	4984	0.0	39.40		44.20		-0.10	-0.20	0.01
HISEI	4854	2.6	59.10	0.61	57.44	0.55	0.11	0.04	0.17
Immigration (% immigrants)	4915	1.4	20.50		21.90		-0.04	-0.08	0.02
English achievement	4977	0.1	517.36	2.38	501.53	2.39	0.30	0.23	0.36
Realistic	4673	6.2	2.64	0.03	1.90	0.01	0.93	0.87	0.99
Investigative	4671	6.3	2.97	0.02	2.41	0.02	0.68	0.63	0.74
Artistic	4669	6.3	2.55	0.02	3.11	0.02	-0.64	-0.69	-0.59
Social	4671	6.3	2.63	0.02	3.34	0.02	-0.82	-0.87	-0.77
Enterprising	4669	6.3	3.07	0.02	3.14	0.02	-0.09	-0.15	-0.04
Conventional	4668	6.3	2.54	0.02	2.43	0.02	0.15	0.10	0.21
Math value	1990	60.1	2.55	0.03	2.27	0.02	0.39	0.31	0.47
Math cost	1973	60.4	2.56	0.03	2.85	0.03	-0.31	-0.39	-0.24
English value	2020	59.5	2.69	0.02	2.75	0.02	-0.09	-0.17	-0.01
English cost	1976	60.4	2.35	0.03	2.35	0.03	0.01	-0.08	0.09
Math self-concept	4865	2.4	2.84	0.02	2.55	0.02	0.35	0.29	0.41
English self-concept	4870	2.3	2.92	0.02	2.87	0.02	0.07	0.01	0.12
Math achievement	4984	0.0	0.19	0.04	-0.22	0.04	0.45	0.39	0.51
STEM major <sup>a</sup> (% of students in STEM majors)	2032	57.0	58.30		31.30		0.55	0.47	0.62
Math-intensive STEM major <sup>b</sup> (% of students in math-intensive STEM majors)	833	57.0	83.80		51.70		0.70	0.59	0.81

*Note.* VG = vocational Gymnasium; HISEI = highest international socioeconomic index.

<sup>a</sup>Mathematics, natural sciences, and engineering were coded as STEM subjects, whereas linguistics and cultural studies, law, business, and social sciences were included in the reference group. <sup>b</sup>Mathematics, physics, computer science, and engineering were coded as math-intensive STEM subjects, whereas biology and medicine were included in the reference group.

Table 3

*Correlations between all Study Variables*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 Gender	-																			
2 School type	.05***	-																		
3 HISEI	-.05***	-.24***	-																	
4 Immigration	.02	.04**	-.21***	-																
5 Engl. AC	-.15***	-.44***	.22***	-.06***	-															
6 Realistic	-.46***	.00	.02	-.04**	.01	-														
7 Investigative	-.34***	.01	.06**	-.03	.12**	.63***	-													
8 Artistic	.32***	-.05***	.05***	.06***	.11***	-.04*	.06***	-												
9 Social	.41***	-.03	.03	-.01	-.08***	-.15***	-.09***	.44***	-											
10 Enterprising	.05*	-.02	.07**	-.02	.00	.05***	.05***	.26***	.48***	-										
11 Conventional	-.08***	.05***	-.03*	.00	-.06***	.33***	.25***	.06***	.16***	.55***	-									
12 Math value	-.19***	.00	-.02	.00	-.01	.36***	.38***	-.25***	-.13***	-.02	.23***	-								
13 Math cost	.16***	.03	-.06**	.04	-.06**	-.21***	-.22***	.23***	.15***	.10***	-.04	-.46***	-							
14 Engl. value	.05*	-.02	.03	.04	.35***	-.01	.04	.24***	.09***	.23***	.14***	-.08***	.21***	-						
15 Engl. cost	.00	.11***	-.11***	.00	-.44***	.13***	.07**	-.15***	.03	.01	.12***	.20***	-.07**	-.29***	-					
16 Math SC	-.17***	-.03	.06**	-.04**	.10***	.27***	.31***	-.24***	-.16***	-.03	.13***	.69***	-.66***	-.16***	.12***	-				
17 Engl. SC	-.03*	-.11***	.11***	.03*	.56***	-.12***	-.03*	.24***	.04*	.13***	-.02	-.24***	.17***	.61***	-.65***	-.17***	-			
18 Math AC	-.22***	-.34***	.15***	-.11***	.41***	.21***	.27***	-.14***	-.16***	-.06***	.03	.41***	-.41***	-.02	-.07***	.53***	.03	-		
19 STEM <sup>a</sup>	-.27***	.01	.04	-.03	.04	.40***	.45***	-.24***	-.24***	-.25***	-.08***	.41***	-.27***	-.08*	.09**	.35***	-.16***	.30***	-	
20 Math-STEM <sup>b</sup>	-.35***	.03	-.14**	-.03	-.09**	.34***	-.04	-.20***	-.23***	.02	.15***	.32***	-.17**	-.13*	.20***	.19***	-.16***	.06	.08*	-

*Note.* HISEI = highest international socioeconomic index; Engl. = English; AC = achievement; SC = self-concept.

<sup>a</sup>Mathematics, natural sciences and engineering were coded as STEM subjects, with linguistics, cultural studies, law, business, and social sciences included in the reference group. <sup>b</sup>Mathematics, physics, informatics, and engineering were coded as math-intensive STEM subjects, with biology and medicine included in the reference group.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

Table 4

*Predicting Mathematics Achievement: Results of Multiple Regression Analyses*

Predictor	Model I				Model II				Model III				Model IV			
	$\beta$	SE	p	95% CI	$\beta$	SE	p	95% CI	$\beta$	SE	p	95% CI	$\beta$	SE	p	95% CI
Gender (f = 1)	-0.23	0.14	<.001	[-0.27, -0.19]	-0.11	0.13	<.001	[-0.14, -0.07]	-0.01	0.15	.710	[-0.05, 0.03]	-0.07	0.14	.001	[-0.11, -0.04]
School type (VG = 1)	-0.31	0.11	<.001	[-0.40, -0.24]	-0.30	0.09	<.001	[-0.38, -0.24]	-0.33	0.10	<.001	[-0.42, -0.26]	-0.31	0.09	<.001	[-0.39, -0.25]
HISEI	0.01	0.01	.428	[-0.01, 0.03]	0.00	0.01	.973	[-0.02, 0.02]	0.02	0.01	.185	[0.00, 0.04]	0.00	0.01	.974	[-0.02, 0.02]
Immigration (= 1)	-0.18	0.14	<.001	[-0.25, -0.13]	-0.15	0.11	<.001	[-0.20, -0.10]	-0.15	0.13	<.001	[-0.21, -0.09]	-0.15	0.11	<.001	[-0.20, -0.10]
Engl. AC	0.30	0.02	<.001	[0.27, 0.32]	0.32	0.02	<.001	[0.29, 0.34]	0.31	0.02	<.001	[0.28, 0.33]	0.30	0.02	<.001	[0.27, 0.33]
Realistic					0.06	0.02	<.001	[0.03, 0.09]	0.06	0.02	<.001	[0.03, 0.09]	0.00	0.02	.945	[-0.03, 0.03]
Investigative					0.19	0.02	<.001	[0.17, 0.22]	0.19	0.02	<.001	[0.17, 0.22]	0.07	0.02	<.001	[0.04, 0.09]
Artistic					-0.16	0.02	<.001	[-0.19, -0.13]	-0.16	0.02	<.001	[-0.19, -0.13]	-0.02	0.02	.305	[-0.04, 0.01]
Social					-0.02	0.02	.201	[-0.05, 0.01]	-0.02	0.02	.201	[-0.05, 0.01]	-0.01	0.02	.479	[-0.04, 0.01]
Enterprising					-0.04	0.02	.017	[-0.07, -0.01]	-0.04	0.02	.017	[-0.07, -0.01]	-0.02	0.02	.122	[-0.05, 0.00]
Conventional					0.02	0.02	.180	[-0.01, 0.05]	0.02	0.02	.180	[-0.01, 0.05]	-0.02	0.02	.284	[-0.04, 0.01]
Math value					0.10	0.02	<.001	[0.06, 0.14]					0.08	0.02	.001	[0.04, 0.12]
Math cost					-0.08	0.02	<.001	[-0.11, -0.04]					-0.07	0.02	.001	[-0.10, -0.04]
Engl. value					-0.01	0.02	.754	[-0.05, 0.03]					0.00	0.02	.952	[-0.04, 0.04]
Engl. cost					-0.04	0.02	.054	[-0.08, -0.01]					-0.04	0.02	.048	[-0.08, -0.01]
Math SC					0.34	0.02	<.001	[0.31, 0.38]					0.34	0.02	<.001	[0.31, 0.38]
Engl. SC					-0.10	0.02	<.001	[-0.14, -0.06]					-0.09	0.02	<.001	[-0.13, -0.05]
R <sup>2</sup>			0.239				.472				.304				.476	

*Note.* The table displays standardized results. CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; Engl. = English; AC = achievement; SC = self-concept.

The predictive power of the individual predictors was also as proposed in Hypotheses 1c and 1d. Math achievement was positively predicted by math self-concept, the combined math values, realistic interests, and investigative interests and negatively predicted by math cost when the sets of predictors were investigated separately, always when the covariates were controlled for (see Models II and III). Concerning the expectancy-value constructs, the strongest predictor was math self-concept, whereas math values were a weaker predictor of the outcome. For vocational interests, investigative interests were a stronger predictor of math achievement than realistic interests were. When all predictors were included in one model, math self-concept and task values were positive predictors of math achievement, similar to what was found in Model II (see Model IV). Math self-concept was the strongest predictor of math achievement also in this model. The prediction of math achievement from vocational interests, however, was different when self-concept and task values were included (Model IV) compared with the model without expectancy-value constructs (Model III): Realistic interests did not significantly predict math achievement in the full model, and investigative interests were less predictive than in Model III but were comparable in size with math values.

### ***Prediction of STEM Majors***

Next, we investigated the predictive power of expectancy-value constructs and vocational interests for the choice of a university STEM major. Here, we looked at different classifications of STEM majors because STEM majors include a wide range of different majors. We again applied a series of regressions, but because of the dichotomous nature of the criterion variables, we used multiple logistic regression analyses.

#### *STEM Majors versus Other Majors*

The results for the choice of a STEM university major are displayed in Table 5. In line with Hypothesis 2a, expectancy-value constructs and vocational interests both predicted the choice of a STEM major as a set of predictors while the covariates were controlled for (see Models II and III). As proposed in Hypothesis 2b, the predictive validity of vocational interests was higher than for expectancy-value constructs (see Pseudo- $R^2$  values for Models II and III). When included simultaneously in one model, the predictive validity was even higher, and the likelihood ratio tests indicated significant improvement in the full model compared with the models consisting of only expectancy-value constructs,  $\chi^2(18) = 492.91, p < .001$ , or vocational interests,  $\chi^2(18) = 44.11, p < .001$ . The constructs therefore added predictive power over and above each other.



Table 5

## Predicting Choice of STEM University Major: Results of Logistic Regression Analyses

Predictor	Model I			Model II			Model III			Model IV										
	$\beta$	OR	SE	$\beta$	OR	SE	$\beta$	OR	SE	$\beta$	OR	SE								
Gender (f = 1)	-0.30	0.74	0.20	<.001	[0.70, 0.79]	-0.25	0.78	0.20	<.001	[0.73, 0.83]	0.07	1.07	0.23	.061	[1.01, 1.14]	0.02	1.02	0.22	.544	[0.96, 1.08]
School type (VG = 1)	0.16	1.18	0.19	.013	[1.06, 1.28]	0.12	1.12	0.20	.080	[1.01, 1.23]	0.03	1.03	0.13	.547	[0.95, 1.1]	0.01	1.01	0.14	.852	[0.92, 1.08]
HISEI	0.04	1.04	0.03	.110	[1.00, 1.09]	0.05	1.05	0.03	.054	[1.01, 1.49]	0.04	1.04	0.03	.082	[1.00, 1.09]	0.06	1.06	0.03	.025	[1.02, 1.11]
Immigration (= 1)	0.02	1.02	0.22	.677	[0.93, 1.12]	0.01	1.01	0.23	.828	[0.91, 1.11]	0.05	1.05	0.20	.322	[0.97, 1.14]	0.04	1.04	0.20	.456	[0.95, 1.12]
Math AC	0.37	1.44	0.03	<.001	[1.37, 1.51]	0.15	1.16	0.04	<.001	[1.09, 1.23]	0.19	1.21	0.03	<.001	[1.16, 1.26]	0.09	1.09	0.03	.004	[1.04, 1.14]
Engl. AC	-0.10	0.91	0.03	<.001	[0.87, 0.95]	0.05	1.05	0.04	.153	[0.99, 1.12]	-0.07	0.94	0.03	.016	[0.90, 0.98]	-0.02	0.98	0.03	.637	[0.93, 1.04]
Realistic						0.25	1.28	0.03	<.001	[1.22, 1.35]	0.21	1.23	0.03	<.001	[1.17, 1.30]					
Investigative						0.39	1.47	0.03	<.001	[1.40, 1.54]	0.34	1.40	0.03	<.001	[1.33, 1.48]					
Artistic						-0.22	0.81	0.03	<.001	[0.77, 0.84]	-0.16	0.85	0.03	<.001	[0.82, 0.89]					
Social						0.04	1.04	0.03	.228	[0.99, 1.09]	0.04	1.04	0.03	.242	[0.99, 1.09]					
Enterprising						-0.15	0.86	0.03	<.001	[0.82, 0.90]	-0.15	0.86	0.03	<.001	[0.82, 0.91]					
Conventional						-0.20	0.82	0.03	<.001	[0.78, 0.86]	-0.24	0.79	0.03	<.001	[0.75, 0.83]					
Math value						0.31	1.36	0.05	<.001	[1.25, 1.48]						0.17	1.19	0.05	.001	[1.09, 1.29]
Math cost						0.02	1.02	0.05	.673	[0.94, 1.12]						0.04	1.04	0.05	.474	[0.95, 1.13]
Engl. value						0.02	1.02	0.05	.679	[0.94, 1.11]						0.07	1.07	0.05	.140	[0.99, 1.15]
Engl. cost						-0.03	0.97	0.05	.519	[0.89, 1.05]						-0.04	0.96	0.05	.390	[0.89, 1.04]
Math SC						0.05	1.05	0.06	.342	[0.96, 1.15]						0.08	1.09	0.05	.120	[1.00, 1.19]
Engl. SC						-0.16	0.85	0.05	.002	[0.78, 0.93]						-0.12	0.89	0.05	.019	[0.82, 0.97]
Pseudo-R <sup>2</sup>									.216					.334					.348	

Note. The table displays standardized results. Mathematics, natural sciences, and engineering were coded as STEM subjects. Linguistics, cultural studies, law, business, and social sciences were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of studying a STEM subject, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; AC = achievement; Engl. = English; SC = self-concept.

Concerning the individual predictors, the results were in line with Hypothesis 2d. The combined math values, realistic interests, and investigative interests positively predicted the choice of STEM major while the covariates were controlled for (see Models II and III). Comparable predictions from these variables to the choice of STEM majors were also found in the full model in which all predictors were included (see Model IV). In this model, investigative interests were the strongest predictor of the choice of a STEM major, whereas realistic interests were a somewhat weaker predictor. Math value was the weakest predictor out of the three predictors but still significantly predicted the choice of a STEM major.

#### *Mathematically Intensive STEM Majors versus Life Science STEM Majors*

The results for the choice of a mathematically intensive STEM major are presented in Table 6. As predicted by Hypothesis 2a, expectancy-value constructs and vocational interests both predicted the choice of a STEM major as a set of predictors while the covariates were controlled for (see Models II and III). As proposed in Hypothesis 2b, vocational interests were a more powerful predictor compared with the expectancy-value constructs. Again, the sets of predictors added predictive validity over and above each other when included together in one model. The pseudo- $R^2$  for the full model was the highest out of those models, and the likelihood ratio tests yielded a significant improvement for this model compared with the models that included the expectancy-value constructs,  $\chi^2(18) = 140.45, p < .001$ , or vocational interests alone,  $\chi^2(18) = 37.29, p < .01$ , always when the covariates were controlled for.

We also found support for our hypotheses concerning the individual predictors as proposed in Hypothesis 2c. Math values positively predicted the choice of a mathematically intensive STEM major (see Model II). The same was true for realistic interests, whereas investigative and social interests negatively predicted the criterion variable (see Model III). The same pattern was also found when the constructs were included together in one model (see Model IV). Here, investigative and realistic interests were the strongest predictors, although in different directions. However, math values and social interests also significantly predicted the choice of a mathematically intensive STEM major.

#### *The Role of Gender*

Finally, we investigated the role of gender in the prediction of math achievement and the choice of a STEM university major. Here, we first investigated how expectancy-value constructs and vocational interests explained gender differences in indicators of STEM careers. As shown in Tables 4, 5, and 6, gender was one of the strongest covariates for math achievement and the choice of a STEM major and a mathematically intensive STEM major (see Model I).

Table 6

*Predicting Choice of Mathematically intensive STEM University Major: Results of Logistic Regression Analyses*

Predictor	Model I				Model II				Model III				Model IV							
	$\beta$	OR	SE	p	OR 95% CI	$\beta$	OR	SE	p	OR 95% CI	$\beta$	OR	SE	p	OR 95% CI					
Gender (f = 1)	-0.62	0.54	0.22	<.001	[0.48, 0.59]	-0.58	0.56	0.24	<.001	[0.5, 0.63]	-0.38	0.69	0.26	<.001	[0.61, 0.76]	-0.38	0.68	0.30	<.001	[0.60, 0.77]
School type (VG = 1)	-0.05	0.95	0.28	.606	[0.78, 1.11]	-0.09	0.92	0.30	.425	[0.74, 1.09]	0.03	1.03	0.25	.762	[0.87, 1.17]	-0.01	0.99	0.26	.943	[0.83, 1.14]
HISEI	-0.15	0.86	0.05	.001	[0.80, 0.93]	-0.10	0.90	0.05	.024	[0.84, 0.97]	-0.13	0.88	0.04	.001	[0.82, 0.94]	-0.09	0.91	0.04	.029	[0.85, 0.98]
Immigration (= 1)	-0.11	0.90	0.33	.268	[0.74, 1.05]	-0.16	0.85	0.35	.125	[0.7, 1.01]	-0.13	0.88	0.30	.152	[0.74, 1.02]	-0.17	0.85	0.31	.071	[0.71, 0.99]
Math AC	0.08	1.09	0.05	.099	[1.00, 1.18]	-0.09	0.91	0.06	.128	[0.83, 1.01]	0.10	1.11	0.05	.040	[1.02, 1.20]	-0.04	0.96	0.05	.444	[0.88, 1.05]
Engl. AC	-0.20	0.82	0.05	<.001	[0.76, 0.88]	-0.10	0.90	0.07	.129	[0.81, 1.01]	-0.14	0.87	0.04	.001	[0.81, 0.93]	-0.04	0.96	0.06	.501	[0.87, 1.06]
Realistic						0.47	1.60	0.05	<.001	[1.49, 1.73]	0.42	1.52	0.05	<.001	[1.40, 1.65]					
Investigative						-0.47	0.63	0.05	<.001	[0.58, 0.68]	-0.53	0.59	0.06	<.001	[0.54, 0.65]					
Artistic						-0.04	0.96	0.05	.363	[0.89, 1.03]	0.03	1.03	0.05	.592	[0.95, 1.12]					
Social						-0.21	0.81	0.05	<.001	[0.75, 0.88]	-0.26	0.77	0.06	<.001	[0.70, 0.85]					
Enterprising						0.08	1.08	0.06	.173	[0.98, 1.18]	0.11	1.12	0.06	.062	[1.01, 1.24]					
Conventional						0.10	1.11	0.06	.070	[1.01, 1.21]	0.06	1.06	0.06	.321	[0.96, 1.17]					
Math value						0.33	1.39	0.10	.001	[1.18, 1.62]						0.29	1.33	0.09	.002	[1.14, 1.55]
Math cost						-0.04	0.97	0.10	.714	[0.83, 1.13]						-0.06	0.94	0.08	.464	[0.83, 1.08]
Engl. value						0.02	1.02	0.08	.802	[0.89, 1.17]						-0.07	0.93	0.09	.393	[0.81, 1.07]
Engl. cost						0.12	1.12	0.09	.223	[0.96, 1.31]						0.18	1.20	0.09	.037	[1.04, 1.39]
Math SC						-0.02	0.98	0.09	.780	[0.85, 1.12]						-0.01	0.99	0.09	.886	[0.85, 1.14]
Engl. SC						0.02	1.02	0.09	.867	[0.87, 1.18]						0.12	1.13	0.08	.161	[0.98, 1.29]
Pseudo-R <sup>2</sup>									.149					.290					.322	

*Note.* The table displays standardized results. Mathematics, physics, computer science, and engineering were coded as mathematically intensive STEM subjects. Biology and medicine were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of studying a mathematically intensive STEM subject, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; AC = achievement; Engl. = English; SC = self-concept.

For math achievement, women showed lower math achievement than men when the other covariates were controlled for. Furthermore, women were less likely than men to choose STEM majors and mathematically intensive STEM majors when the covariates were controlled for. As proposed in Hypothesis 3, the expectancy-value constructs and vocational interests both partly explained the gender differences in the criterion variables. For math achievement, the expectancy-value constructs partly explained the gender differences in math achievement when the covariates were controlled for, although women still showed lower math achievement than men. However, this was not the case for vocational interests, indicating that gender differences in vocational interests were able to fully explain gender differences in math achievement when the other covariates were controlled for (see Table 4, Models II and III). However, when the expectancy-value constructs and vocational interests were included together in the model, women again showed lower math achievement than men (see Table 4, Model IV).

With respect to university majors, we found a similar pattern for the choice of a STEM major as we did for math achievement. Whereas the expectancy-value constructs only partly explained the gender differences in such choices (see Table 5, Model II), vocational interests fully explained the gender differences (see Table 5, Model III), always when the covariates were controlled for. Also in the full model, gender was not a significant predictor. Thus, the expectancy-value constructs and vocational interests together explained the different choices made by the different genders (see Table 5, Model IV).

Concerning the choice of a mathematically intensive STEM major, however, gender differences could not be explained by the expectancy-value constructs or vocational interests separately or by the two sets together (see Models II, III, and IV in Table 6). Nevertheless, there were differences in the relative predictive power of gender when investigated along with the expectancy-value constructs and vocational interests, indicating that gender was less predictive of such choices when vocational interests were included as predictors: Gender was still one of the strongest predictors or the strongest predictor when self-concepts and values were included (see Table 6, Model II), whereas gender was a relatively weaker predictor than realistic and investigative interests when vocational interests were added to the covariates (see Table 6, Model III).

We also explored whether the constructs predicted the criterion variables differently for men and women by applying multiple-group analyses (see Tables 7-12). We tested differences in the predictors between men and women with Wald tests. Concerning math achievement, we found no differences between men and women in the predictive power of math self-concept and task values in predicting the criterion while the covariates were controlled for. The combined

variable for English values, however, was negatively predictive for men but not for women; in Model II:  $\chi^2(1) = 4.58, p = .032$ ; in Model IV:  $\chi^2(1) = 7.82, p = .005$ . Furthermore, the relative predictive power of investigative interests differed by gender, with higher predictive power for men than for women,  $\chi^2(1) = 14.70, p < .001$ , while the covariates were controlled for but not when the covariates and the expectancy-value constructs were controlled for.

With respect to the choice of university major, we found no significant differences in the relative predictive power of the math expectancy-value constructs for men and women. Conversely, we found some differences in the role of vocational interests. For the choice of a STEM major, social interests significantly predicted such majors only for men but not for women when the covariates were controlled for,  $\chi^2(1) = 6.28, p = .012$ , and when the covariates and the expectancy-value constructs were controlled for,  $\chi^2(1) = 4.49, p = .034$ . In addition, we found differences for enterprising interests, as these interests were stronger negative predictors of the choice of a STEM major for men than for women,  $\chi^2(1) = 5.17, p = .023$ , while the covariates were controlled for. For the choice of a mathematically intensive STEM major, investigative interests were a more powerful negative predictor for women than for men when the covariates were controlled for,  $\chi^2(1) = 6.56, p = .010$ . In addition, conventional interests were predictive of the choice of a mathematically intensive STEM major only for women when the covariates were controlled for,  $\chi^2(1) = 3.95, p = .05$ , whereas social interests were a stronger negative predictor of these choices for men than for women when the covariates and the expectancy-value constructs were controlled for,  $\chi^2(1) = 4.96, p = .026$ .

Table 7

## Predicting Mathematics Achievement for Men: Results of Multiple-Group Multiple Regression Analyses

Predictor	Model I			Model II			Model III			Model IV										
	$\beta$	SE	p	$\beta$	SE	p	$\beta$	SE	p	$\beta$	SE	p	95% CI							
School type (VG = 1)	-0.25	0.16	<.001	-0.37	-0.15]		-0.23	0.12	<.001	[-0.33, -0.15] <sup>a</sup>		-0.27	0.14	<.001	[-0.39, -0.18]		-0.24	0.12	<.001	[-0.34, -0.16] <sup>a</sup>
HISEI	0.03	0.02	.088	[0.00, 0.06]		0.02	0.02	.374	[-0.01, 0.05]		0.03	0.02	.065	[0.00, 0.06]		0.01	0.02	.479	[-0.02, 0.04]	
Immigration (= 1)	-0.12	0.18	.011	[-0.21, -0.04] <sup>a</sup>		-0.08	0.14	.043	[-0.15, -0.01] <sup>a</sup>		-0.09	0.17	.046	[-0.17, -0.01] <sup>a</sup>		-0.08	0.14	.040	[-0.15, -0.01] <sup>a</sup>	
Engl. AC	0.30	0.02	<.001	[0.27, 0.34]		0.32	0.02	<.001	[0.28, 0.36]		0.29	0.02	<.001	[0.26, 0.33]		0.30	0.03	<.001	[0.26, 0.34]	
Realistic											0.02	0.02	.323	[-0.02, 0.06]		-0.01	0.02	.722	[-0.05, 0.03]	
Investigative											0.25	0.02	<.001	[0.21, 0.28] <sup>a</sup>		0.08	0.02	<.001	[0.05, 0.12]	
Artistic											-0.16	0.02	<.001	[-0.20, -0.12]		-0.01	0.02	.547	[-0.05, 0.02]	
Social											-0.01	0.03	.802	[-0.05, 0.04]		-0.01	0.02	.771	[-0.04, 0.03]	
Enterprising											-0.03	0.03	.285	[-0.08, 0.02]		0.00	0.02	.926	[-0.04, 0.04]	
Conventional											0.01	0.03	.851	[-0.04, 0.05]		-0.01	0.02	.812	[-0.05, 0.03]	
Math value											0.12	0.04	.002	[0.06, 0.19]		0.09	0.04	.026	[0.02, 0.16]	
Math cost											-0.07	0.03	.029	[-0.12, -0.02]		-0.07	0.03	.036	[-0.12, -0.02]	
Engl value											-0.06	0.03	.054	[-0.12, -0.01] <sup>a</sup>		-0.08	0.03	.028	[-0.13, -0.02] <sup>a</sup>	
Engl cost											-0.03	0.03	.363	[-0.08, 0.02]		-0.04	0.03	.233	[-0.09, 0.02]	
Math SC											0.36	0.04	<.001	[0.30, 0.42]		0.36	0.04	<.001	[0.31, 0.42]	
Engl SC											-0.09	0.03	.008	[-0.14, -0.03]		-0.08	0.03	.014	[-0.13, -0.03]	

Note. The table displays standardized results. CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; Engl. = English; AC = achievement; SC = self-concept.

<sup>a</sup>Significantly different parameter estimates between men and women at  $p < .05$ .

Table 8

*Predicting Mathematics Achievement for Women: Results of Multiple-Group Multiple Regression Analyses*

Predictor	Model I				Model II				Model III				Model IV			
	$\beta$	SE	p	95% CI	$\beta$	SE	p	95% CI	$\beta$	SE	p	95% CI	$\beta$	SE	p	95% CI
School type (VG = 1)	-0.36	0.12	<.001	[-0.46, -0.28]	-0.36	0.11	<.001	[-0.45, -0.28] <sup>a</sup>	-0.38	0.11	<.001	[-0.48, -0.30]	-0.37	0.11	<.001	[-0.47, -0.30] <sup>a</sup>
HISEI	-0.01	0.02	.761	[-0.04, 0.03]	-0.01	0.02	.481	[-0.04, 0.02]	0.01	0.02	.785	[-0.02, 0.03]	-0.01	0.02	.616	[-0.03, 0.02]
Immigration (= 1)	-0.24	0.18	<.001	[-0.32, -0.17] <sup>a</sup>	-0.22	0.15	<.001	[-0.28, -0.15] <sup>a</sup>	-0.21	0.18	<.001	[-0.29, -0.14] <sup>a</sup>	-0.21	0.15	<.001	[-0.28, -0.15] <sup>a</sup>
Engl. AC	0.30	0.02	<.001	[0.27, 0.33]	0.33	0.02	<.001	[0.29, 0.37]	0.32	0.02	<.001	[0.29, 0.36]	0.32	0.02	<.001	[0.28, 0.36]
Realistic					0.08	0.02	<.001	[0.05, 0.12]	0.08	0.02	<.001	[0.05, 0.12]	0.01	0.02	.674	[-0.03, 0.04]
Investigative					0.14	0.02	<.001	[0.10, 0.17] <sup>a</sup>	0.14	0.02	<.001	[0.10, 0.17] <sup>a</sup>	0.05	0.02	.025	[0.01, 0.08]
Artistic					-0.16	0.02	<.001	[-0.19, -0.13]	-0.16	0.02	<.001	[-0.19, -0.13]	-0.03	0.02	.112	[-0.06, 0.00]
Social					-0.03	0.02	.159	[-0.07, 0.01]	-0.03	0.02	.159	[-0.07, 0.01]	-0.01	0.02	.631	[-0.04, 0.02]
Enterprising					-0.06	0.03	.019	[-0.10, -0.02]	-0.06	0.03	.019	[-0.10, -0.02]	-0.05	0.02	.021	[-0.09, -0.01]
Conventional					0.04	0.02	.129	[0.00, 0.08]	0.04	0.02	.129	[0.00, 0.08]	-0.01	0.02	.526	[-0.05, 0.02]
Math value					0.09	0.03	.006	[0.03, 0.14]	0.09	0.03	.006	[0.03, 0.14]	0.07	0.03	.019	[0.02, 0.13]
Math cost					-0.08	0.03	.004	[-0.13, -0.04]	-0.08	0.03	.004	[-0.13, -0.04]	-0.07	0.03	.021	[-0.11, -0.02]
Engl value					0.04	0.03	.285	[-0.02, 0.09] <sup>a</sup>	0.04	0.03	.285	[-0.02, 0.09] <sup>a</sup>	0.06	0.03	.074	[0.01, 0.12] <sup>a</sup>
Engl cost					-0.05	0.03	.141	[-0.10, 0.01]	-0.05	0.03	.141	[-0.10, 0.01]	-0.04	0.03	.179	[-0.10, 0.01]
Math SC					0.32	0.03	<.001	[0.28, 0.37]	0.32	0.03	<.001	[0.28, 0.37]	0.32	0.03	<.001	[0.28, 0.37]
Engl SC					-0.12	0.04	.001	[-0.18, -0.06]	-0.12	0.04	.001	[-0.18, -0.06]	-0.12	0.04	.001	[-0.18, -0.06]

*Note.* The table displays standardized results. CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; Engl. = English; AC = achievement; SC = self-concept.

<sup>a</sup>Significantly different parameter estimates between men and women at  $p < .05$ .

Table 9

## Predicting Choice of STEM University Major for Men: Results of Multiple-Group Logistic Regression Analyses

Predictor	Model I			Model II			Model III			Model IV										
	$\beta$	OR	SE	$\beta$	OR	SE	$\beta$	OR	SE	$\beta$	OR	SE	$p$	OR	95% CI					
School type (VG = 1)	0.13	1.13	0.28	.209	[0.96, 1.30]	0.16	1.18	0.28	.111	[0.99, 1.35]	0.01	1.01	0.20	.873	[0.89, 1.12]	0.03	1.03	0.20	.666	[0.90, 1.15]
HISEI	-0.04	0.96	0.05	.338	[0.89, 1.03] <sup>a</sup>	-0.02	0.98	0.05	.685	[0.91, 1.06] <sup>a</sup>	-0.04	0.96	0.04	.379	[0.90, 1.03] <sup>a</sup>	-0.02	0.98	0.05	.586	[0.91, 1.05] <sup>a</sup>
Immigration (= 1)	0.00	1.00	0.28	.971	[0.85, 1.16]	0.03	1.03	0.30	.790	[0.86, 1.19]	0.06	1.06	0.25	.455	[0.92, 1.20]	0.07	1.07	0.25	.397	[0.93, 1.21]
Math AC	0.34	1.40	0.05	<.000	[1.30, 1.51]	0.15	1.16	0.05	.004	[1.07, 1.26]	0.15	1.16	0.04	<.000	[1.09, 1.24]	0.09	1.10	0.04	.032	[1.02, 1.18]
Engl. AC	-0.06	0.94	0.04	.153	[0.87, 1.01]	0.13	1.14	0.06	.018	[1.04, 1.25]	-0.03	0.97	0.04	.460	[0.91, 1.04]	0.02	1.02	0.05	.631	[0.94, 1.11]
Realistic						0.27	1.31	0.04	<.000	[1.23, 1.40]	0.27	1.31	0.04	<.000	[1.23, 1.40]	0.25	1.28	0.04	<.000	[1.20, 1.37]
Investigative						0.41	1.51	0.04	<.000	[1.40, 1.62]	0.41	1.51	0.04	<.000	[1.40, 1.62]	0.38	1.46	0.05	<.000	[1.35, 1.59]
Artistic						-0.23	0.80	0.04	<.000	[0.74, 0.85]	-0.23	0.80	0.04	<.000	[0.74, 0.85]	-0.19	0.83	0.04	<.000	[0.77, 0.89]
Social						0.13	1.14	0.05	.010	[1.05, 1.23] <sup>a</sup>	0.13	1.14	0.05	.010	[1.05, 1.23] <sup>a</sup>	0.12	1.13	0.05	.019	[1.04, 1.23] <sup>a</sup>
Enterprising						-0.24	0.79	0.05	<.000	[0.73, 0.86] <sup>a</sup>	-0.24	0.79	0.05	<.000	[0.73, 0.86] <sup>a</sup>	-0.22	0.80	0.05	<.000	[0.74, 0.87]
Conventional						-0.24	0.79	0.05	<.000	[0.73, 0.85]	-0.24	0.79	0.05	<.000	[0.73, 0.85]	-0.26	0.77	0.05	<.000	[0.71, 0.84]
Math value						0.37	1.45	0.09	<.000	[1.25, 1.69]	0.37	1.45	0.09	<.000	[1.25, 1.69]	0.15	1.16	0.08	.066	[1.02, 1.32]
Math cost						-0.01	0.99	0.08	.873	[0.86, 1.13]	-0.01	0.99	0.08	.873	[0.86, 1.13]	-0.02	0.98	0.08	.843	[0.86, 1.13]
Engl. value						-0.02	0.98	0.07	.760	[0.87, 1.10] <sup>a</sup>	-0.02	0.98	0.07	.760	[0.87, 1.10] <sup>a</sup>	0.06	1.06	0.08	.488	[0.93, 1.21]
Engl. cost						-0.02	0.98	0.09	.832	[0.85, 1.13]	-0.02	0.98	0.09	.832	[0.85, 1.13]	-0.10	0.91	0.08	.204	[0.80, 1.03]
Math SC						-0.09	0.92	0.08	.285	[0.80, 1.05]	-0.09	0.92	0.08	.285	[0.80, 1.05]	-0.03	0.97	0.08	.732	[0.85, 1.12]
Engl. SC						-0.20	0.82	0.08	.014	[0.71, 0.94]	-0.20	0.82	0.08	.014	[0.71, 0.94]	-0.18	0.84	0.08	.022	[0.74, 0.95]

*Note.* The table displays standardized results. Mathematics, natural sciences, and engineering were coded as STEM subjects. Linguistics, cultural studies, law, business, and social sciences were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of studying STEM subjects, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; AC = achievement; Engl. = English; SC = self-concept.

<sup>a</sup>Significantly different parameter estimates between men and women at  $p < .05$ .



Table 10

*Predicting Choice of STEM University Major for Women: Results of Multiple-Group Logistic Regression Analyses*

Predictor	Model I				Model II				Model III				Model IV							
	$\beta$	OR	SE	p	$\beta$	OR	SE	p	$\beta$	OR	SE	p	$\beta$	OR	SE	p	OR	95% CI		
School type (VG = 1)	0.20	1.22	0.20	.005	[1.10, 1.33]	0.10	1.10	0.22	.198	[0.97, 1.23]	0.03	1.03	0.16	.556	[0.94, 1.12]	-0.03	0.98	0.17	.686	[0.87, 1.07]
HISEI	0.12	1.12	0.04	.002	[1.06, 1.19] <sup>a</sup>	0.12	1.13	0.04	.001	[1.06, 1.20] <sup>a</sup>	0.11	1.12	0.04	.002	[1.05, 1.18] <sup>a</sup>	0.12	1.13	0.04	.001	[1.07, 1.19] <sup>a</sup>
Immigration (= 1)	0.04	1.04	0.32	.626	[0.90, 1.18]	0.00	1.00	0.31	.950	[0.87, 1.14]	0.04	1.04	0.29	.591	[0.91, 1.17]	0.01	1.01	0.30	.880	[0.88, 1.14]
Math AC	0.41	1.50	0.04	<.000	[1.41, 1.60]	0.17	1.19	0.05	.001	[1.09, 1.30]	0.23	1.26	0.04	<.000	[1.18, 1.33]	0.10	1.10	0.04	.029	[1.02, 1.18]
Engl. AC	-0.13	0.87	0.04	.001	[0.82, 0.93]	-0.01	1.00	0.05	.918	[0.91, 1.08]	-0.10	0.91	0.04	.014	[0.85, 0.97]	-0.05	0.95	0.05	.311	[0.87, 1.03]
Realistic						0.17	1.18	0.04	<.000	[1.10, 1.26]	0.12	1.13	0.04	.005	[1.05, 1.21]	0.12	1.13	0.04	.005	[1.05, 1.21]
Investigative						0.35	1.42	0.04	<.000	[1.34, 1.51]	0.31	1.36	0.04	<.000	[1.28, 1.45]	0.31	1.36	0.04	<.000	[1.28, 1.45]
Artistic						-0.22	0.81	0.04	<.000	[0.76, 0.86]	-0.15	0.86	0.04	<.000	[0.80, 0.92]	-0.15	0.86	0.04	<.000	[0.80, 0.92]
Social						-0.03	0.97	0.04	.427	[0.91, 1.03] <sup>a</sup>	-0.03	0.97	0.04	.427	[0.91, 1.03] <sup>a</sup>	-0.02	0.98	0.04	.613	[0.92, 1.05] <sup>a</sup>
Enterprising						-0.11	0.90	0.04	.008	[0.84, 0.96] <sup>a</sup>	-0.11	0.90	0.04	.008	[0.84, 0.96] <sup>a</sup>	-0.13	0.88	0.04	.003	[0.82, 0.95]
Conventional						-0.19	0.83	0.04	<.000	[0.78, 0.89]	-0.19	0.83	0.04	<.000	[0.78, 0.89]	-0.23	0.80	0.04	<.000	[0.74, 0.85]
Math value						0.24	1.27	0.07	<.000	[1.14, 1.41]	0.24	1.27	0.07	<.000	[1.14, 1.41]	0.17	1.19	0.07	.010	[1.07, 1.32]
Math cost						0.05	1.05	0.07	.472	[0.94, 1.18]	0.05	1.05	0.07	.472	[0.94, 1.18]	0.07	1.07	0.07	.321	[0.96, 1.19]
Engl. value						0.03	1.03	0.06	.596	[0.93, 1.14] <sup>a</sup>	0.03	1.03	0.06	.596	[0.93, 1.14] <sup>a</sup>	0.08	1.09	0.06	.165	[0.99, 1.20]
Engl. cost						-0.04	0.96	0.08	.629	[0.84, 1.10]	-0.04	0.96	0.08	.629	[0.84, 1.10]	0.00	1.00	0.07	.973	[0.89, 1.12]
Math SC						0.16	1.17	0.07	.023	[1.05, 1.32]	0.16	1.17	0.07	.023	[1.05, 1.32]	0.16	1.17	0.07	.016	[1.05, 1.31]
Engl. SC						-0.13	0.88	0.08	.084	[0.77, 0.99]	-0.13	0.88	0.08	.084	[0.77, 0.99]	-0.08	0.92	0.07	.280	[0.82, 1.04]

*Note.* The table displays standardized results. Mathematics, natural sciences, and engineering were coded as STEM subjects. Linguistics, cultural studies, law, business, and social sciences were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of studying STEM subjects, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; AC = achievement; Engl. = English; SC = self-concept.

<sup>a</sup>Significantly different parameter estimates between men and women at  $p < .05$ .

Table 11

## Predicting Choice of Mathematically intensive STEM University Major for Men: Results of Multiple-Group Logistic Regression Analyses

Predictor	Model I			Model II			Model III			Model IV										
	$\beta$	OR	SE	$\beta$	OR	SE	$\beta$	OR	SE	$\beta$	OR	SE	$p$	OR	95% CI					
School type (VG = 1)	0.07	1.07	0.41	.661	[0.80, 1.35]	0.12	1.13	0.42	.454	[0.84, 1.42] <sup>a</sup>	0.12	1.13	0.37	.386	[0.88, 1.38]	0.20	1.22	0.37	.179	[0.95, 1.47]
HISEI	-0.26	0.77	0.08	.001	[0.68, 0.88]	-0.20	0.82	0.08	.009	[0.72, 0.93]	-0.20	0.82	0.07	.003	[0.73, 0.91]	-0.15	0.86	0.07	.031	[0.76, 0.96]
Immigration (= 1)	-0.22	0.80	0.46	.154	[0.59, 1.03]	-0.21	0.81	0.49	.207	[0.58, 1.06]	-0.21	0.81	0.45	.161	[0.59, 1.04]	-0.20	0.82	0.45	.189	[0.61, 1.05]
Math AC	0.11	1.11	0.08	.195	[0.97, 1.27]	-0.03	0.97	0.09	.727	[0.84, 1.12]	0.13	1.14	0.08	.091	[1.00, 1.29]	0.01	1.01	0.08	.872	[0.89, 1.15]
Engl. AC	-0.28	0.76	0.07	<.000	[0.67, 0.86]	-0.09	0.92	0.10	.393	[0.78, 1.08]	-0.24	0.79	0.07	.001	[0.70, 0.89]	-0.02	0.98	0.09	.806	[0.84, 1.14]
Realistic						0.47	1.60	0.07	<.000	[1.43, 1.79]	0.42	1.52	0.07	<.000	[1.35, 1.71]	0.42	1.52	0.07	<.000	[1.35, 1.71]
Investigative						-0.30	0.74	0.07	<.000	[0.66, 0.84] <sup>a</sup>	-0.36	0.70	0.08	<.000	[0.61, 0.80]	-0.36	0.70	0.08	<.000	[0.61, 0.80]
Artistic						0.07	1.07	0.08	.351	[0.95, 1.22]	0.13	1.14	0.08	.114	[1.00, 1.30]	0.13	1.14	0.08	.114	[1.00, 1.30]
Social						-0.32	0.73	0.08	<.000	[0.64, 0.82]	-0.39	0.68	0.08	<.000	[0.59, 0.77] <sup>a</sup>	-0.39	0.68	0.08	<.000	[0.59, 0.77] <sup>a</sup>
Enterprising						0.02	1.02	0.09	.798	[0.88, 1.19]	0.05	1.05	0.10	.644	[0.89, 1.23]	0.05	1.05	0.10	.644	[0.89, 1.23]
Conventional						0.01	1.01	0.09	.896	[0.88, 1.17] <sup>a</sup>	0.04	1.04	0.09	.687	[0.89, 1.20]	0.04	1.04	0.09	.687	[0.89, 1.20]
Math value						0.24	1.27	0.18	.176	[0.95, 1.70]	0.17	1.18	0.17	.344	[0.89, 1.57]	0.17	1.18	0.17	.344	[0.89, 1.57]
Math cost						-0.06	0.94	0.16	.697	[0.72, 1.22]	-0.05	0.95	0.14	.731	[0.76, 1.20]	-0.05	0.95	0.14	.731	[0.76, 1.20]
Engl. value						-0.03	0.97	0.15	.813	[0.76, 1.23]	-0.13	0.88	0.15	.387	[0.68, 1.13]	-0.13	0.88	0.15	.387	[0.68, 1.13]
Engl. cost						0.29	1.34	0.16	.066	[1.03, 1.74]	0.39	1.47	0.13	.003	[1.19, 1.82]	0.39	1.47	0.13	.003	[1.19, 1.82]
Math SC						-0.05	0.95	0.15	.719	[0.74, 1.21]	-0.01	1.00	0.14	.971	[0.79, 1.26]	-0.01	1.00	0.14	.971	[0.79, 1.26]
Engl. SC						0.05	1.05	0.15	.754	[0.82, 1.33]	0.13	1.14	0.14	.370	[0.82, 1.04]	0.13	1.14	0.14	.370	[0.82, 1.04]

*Note.* The table displays standardized results. Mathematics, physics, computer science, and engineering were coded as mathematically intensive STEM subjects. Biology and medicine were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of studying a mathematically intensive STEM subject, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; AC = achievement; Engl. = English; SC = self-concept.

<sup>a</sup>Significantly different parameter estimates between men and women at  $p < .05$ .

Table 12

*Predicting Choice of Mathematically intensive STEM University Major for Women: Results of Multiple-Group Logistic Regression Analyses*

Predictor	Model I			Model II			Model III			Model IV										
	$\beta$	OR	SE	$p$	OR	SE	$p$	OR	SE	$p$	OR	SE	$p$	OR	95% CI					
School type (VG = 1)	-0.16	0.86	0.36	.244	[0.65, 1.06]	-0.27	0.77	0.38	.057	[0.57, 0.97] <sup>a</sup>	-0.02	0.98	0.32	.843	[0.78, 1.17]	-0.13	0.88	0.35	.320	[0.68, 1.08]
HISEI	-0.08	0.92	0.07	.243	[0.82, 1.03]	-0.05	0.96	0.07	.547	[0.85, 1.08]	-0.09	0.92	0.07	.192	[0.82, 1.02]	-0.06	0.95	0.07	.417	[0.84, 1.06]
Immigration (= 1)	-0.07	0.94	0.38	.624	[0.73, 1.15]	-0.18	0.83	0.45	.246	[0.61, 1.07]	-0.09	0.91	0.34	.453	[0.73, 1.10]	-0.13	0.87	0.38	.313	[0.68, 1.08]
Math AC	0.06	1.06	0.07	.345	[0.96, 1.19]	-0.19	0.83	0.08	.019	[0.73, 0.95]	0.06	1.07	0.06	.311	[0.96, 1.18]	-0.11	0.89	0.07	.113	[0.79, 1.00]
Engl. AC	-0.18	0.83	0.08	.017	[0.73, 0.94]	-0.10	0.90	0.10	.317	[0.76, 1.07]	-0.09	0.91	0.07	.176	[0.82, 1.02]	-0.05	0.95	0.09	.596	[0.81, 1.11]
Realistic						0.44	1.56	0.06	<.000	[1.41, 1.71]	0.40	1.49	0.07	<.000	[1.34, 1.65]					
Investigative						-0.61	0.54	0.07	<.000	[0.49, 0.61] <sup>a</sup>	-0.64	0.53	0.07	<.000	[0.47, 0.59]					
Artistic						-0.10	0.90	0.06	.082	[0.82, 0.99]	-0.01	0.99	0.07	.843	[0.89, 1.10]					
Social						-0.17	0.84	0.07	.009	[0.75, 0.94]	-0.18	0.83	0.07	.010	[0.74, 0.94] <sup>a</sup>					
Enterprising						0.12	1.13	0.08	.146	[0.98, 1.29]	0.16	1.17	0.09	.083	[1.01, 1.36]					
Conventional						0.22	1.24	0.07	.002	[1.10, 1.39] <sup>a</sup>	0.13	1.14	0.08	.086	[1.01, 1.30]					
Math value						0.41	1.51	0.13	.001	[1.23, 1.86]	0.32	1.38	0.12	.006	[1.14, 1.68]					
Math cost						0.03	1.03	0.13	.829	[0.83, 1.27]	-0.01	0.99	0.14	.918	[0.79, 1.24]					
Engl. value						0.00	1.00	0.11	.995	[0.83, 1.20]	-0.05	0.96	0.12	.688	[0.79, 1.15]					
Engl. cost						-0.04	0.97	0.13	.785	[0.78, 1.19]	0.07	1.07	0.14	.628	[0.85, 1.34]					
Math SC						0.06	1.06	0.13	.644	[0.86, 1.31]	0.07	1.07	0.14	.608	[0.85, 1.35]					
Engl. SC						-0.05	0.95	0.13	.685	[0.76, 1.18]	0.06	1.06	0.15	.711	[0.83, 1.35]					

*Note.* The table displays standardized results. Mathematics, physics, computer science, and engineering were coded as mathematically intensive STEM subjects. Biology and medicine were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of studying a mathematically intensive STEM subject, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; VG = vocational Gymnasium; HISEI = highest international socioeconomic index; AC = achievement; Engl. = English; SC = self-concept.

<sup>a</sup>Significantly different parameter estimates between men and women at  $p < .05$ .

## Discussion

In the present study, we extended existing findings on expectancy-value constructs and vocational interests. Each set of constructs is well established as an important predictor of STEM careers, but less is known about their relative power to predict different STEM careers for men and women. We investigated the differential predictive abilities using a large-scale longitudinal data set, which enabled us to investigate how expectancy-value constructs and vocational interests predict different indicators of STEM careers as well as the moderating role of gender. Overall, our results were in line with our expectations. Expectancy-value constructs and vocational interests were highly predictive of math achievement and the choice of various STEM majors, but their relative predictive power differed depending on their domain-specificity. Whereas expectancies and values were more predictive of math achievement, vocational interests were more predictive of the choice of a STEM university major. In addition, we found gender differences in the mean levels as expected and gender differences in the relative predictive power of the constructs, which we had investigated in an exploratory manner. Gender differences in the mean levels partly explained gender differences in indicators of STEM careers, and whereas the power of the expectancy-value constructs to predict STEM constructs was comparable between men and women, we found differences with respect to vocational interests and math achievement and the choice of STEM majors, respectively.

### *Differential Predictive Power of Expectancy-Value Constructs and Vocational Interests in Predicting STEM Careers*

Our results concerning the predictive power of expectancy-value constructs and vocational interests in predicting STEM careers when investigated separately are in line with previous research (e.g., Ackerman & Heggestad, 1997; Denissen et al., 2007; Päßler & Hell, 2012; Wang et al., 2015). With respect to differential associations of the expectancy-value constructs, such work has shown that expectancies and task values in one domain are related to achievement and choices in the same domain, usually with stronger associations between expectancies and achievement, whereas task values usually show stronger associations with choices (e.g., Denissen et al., 2007; Nagy et al., 2008). In the expectancy-value theory, differences between domain-specific expectancy beliefs and task values are explained as follows: On the one hand, it is assumed that individuals with high competence beliefs in a domain aspire to achieve more ambitious goals, engage more deeply in activities, and show more effort in activities in the same domain, all of which thus results in higher achievement in the end (Wigfield et al., 2015). Such relations have also been assumed in other theoretical

frameworks such as in Bandura's (1986) social cognitive theory. According to social cognitive theory, individuals with high competence beliefs strive to find more challenging task and thus show higher abilities in the end (Bussey & Bandura, 1999). In the expectancy-value theory, subjective task values, on the other hand, are assumed to be more closely related to choices because value beliefs are more strongly associated with the question of whether an individual wants to engage in a task (Wigfield et al., 2015; Wigfield & Cambria, 2010).

Also, differential effects of expectancy-value constructs versus vocational interests are in line with previous evidence. We found that compared with expectancy-value constructs, vocational interests were less predictive of math achievement but more predictive of the choice of a STEM major. This finding may be explained by differences in how competence beliefs are captured in the two models. For vocational interests, Holland (1997) suggested that vocational interests correspond with abilities and ability beliefs because vocational interests predict individuals' choices of different environments, which in turn provide specific opportunities and requirements, leading to the skills as well as the corresponding ability beliefs required in the respective environments. Unlike in expectancy-value theory, ability beliefs are thus not explicitly included as separate constructs in Holland's model. In line with the results of the studies described above, which showed stronger associations between ability beliefs and achievement than between interests and achievement, it is therefore not surprising that expectancy value constructs that include ability beliefs explicitly showed stronger relations with math achievement than vocational interests did in our study.

A second possible explanation for the finding that expectancy-value constructs are stronger predictors of achievement, whereas vocational interests are stronger predictors of STEM choice is related to differences in the level of domain-specificity between expectancy-value constructs, vocational interests, math achievement, and STEM university majors. Expectancy and value beliefs are highly domain-specific and are typically assessed at the specific level of a school subject such as expectancy and value beliefs in mathematics (e.g., Wang et al., 2015). Consequently, they are measured on the same level of specificity as math achievement but on a different level than STEM university majors, which include a broad range of different subjects. By contrast, vocational interests refer to a more general level, capturing interests that are related to different professions (Rounds & Su, 2014). Thus, vocational interests are measured on a level that is comparable to university majors, particularly when a broad range of majors is investigated (e.g., STEM university majors vs. majors outside of STEM), whereas they are conceptualized on a more general level than math achievement. In addition, math achievement is not explicitly associated with a single vocational interest; rather, some of the

vocational interests capture mathematically related aspects (e.g., realistic and investigative interests; Bergmann & Eder, 2005). Relating to the differences in the domain-specificity of the constructs, it is thus reasonable that the expectancy-value constructs were stronger predictors of math achievement than vocational interests were, and vocational interests were stronger predictors of STEM university majors than the expectancy-value constructs were.

### *The Role of Gender*

With respect to gender differences, we found mean-level differences in expectancy-value constructs and vocational interests, which are in line with previous research (Marsh, Craven, & Debus, 1998; Su et al., 2009; Wigfield et al., 1997). There is a large body of research findings suggesting lower scores for women than men in STEM-related constructs, such as math expectancy and value beliefs, realistic interests, and investigative interests. Previous studies have furthermore found that gender differences in (mathematically intensive) STEM fields are related to gender differences in mean levels of expectancy-value constructs and vocational interests (e.g., Schoon & Eccles, 2014; Su et al., 2009). Extending these studies, we provided evidence that both sets of predictors contributed to explaining gender differences in STEM careers, even though vocational interests seemed to be more powerful in explaining gender differences in math achievement and the choice of a STEM major. Nevertheless, our results also support the assumption that gender differences in STEM careers are rooted in complex processes, including a variety of different contextual and individual factors that influence males' and females' choices in different ways (e.g., Cheryan, Ziegler, Montoya, & Jiang, 2016; Eccles, 2009; Hübner et al., 2017).

Our study furthermore indicates that the relative predictive power of these constructs in predicting STEM careers differs in part by gender. Whereas the extent to which the expectancy-value constructs could predict the STEM careers was comparable for men and women, we found different results for vocational interests: especially investigative interests and also social interests predicted the criterion variables differently for men and women. Previous research on differential associations between expectancy-value constructs or vocational interests with STEM careers is limited, and among the existing findings, some studies have found differences in the relative power of expectancies, task values, and vocational interests in predicting STEM careers (Päßler & Hell, 2012; Watt, Richardson, & Devos, 2013) but others have not, at least for expectancy-value constructs (e.g., Guo et al., 2015; Tynkkynen et al., 2012). Such mixed findings might be related to the broad variety of indicators of STEM careers that have been investigated in previous research—including course selection in high school (Watt et al., 2013),

career aspirations (Watt et al., 2013), and different STEM majors at university (Guo et al., 2015; Päßler & Hell, 2012). Our results provide further insights about differences between the genders in the extent to which the constructs predict STEM careers; such differences are crucial for understanding women's underrepresentation in STEM areas. However, more systematic research is needed to explore the moderating role of gender in relations between the constructs explored in this study and various indicators of STEM careers.

### *Implications and Further Research*

Our study provides implications for theory and practice. We provide support for associations between expectancy-value constructs and vocational interests, although their interrelations as well as their relative predictive value for STEM careers suggest important differences. Referring to the different levels of domain-specificity of expectancy-value constructs and vocational interests, the present study suggests that both sets of predictors are highly relevant for different STEM outcomes. Future research, however, should carefully consider congruence in the domain-specificity of predictors and criterion—at least when comparing different predictors—because different levels of specificity might add to a better understanding of the investigated relations. Depending on the investigated criterion, predictors should be chosen carefully as expectancy-value constructs seem to be better predictors of mathematics achievement, and vocational interests seem to be better predictors of the choice of university majors.

The differences in the relative predictive power of expectancy-value constructs and vocational interests also add to discussions about interventions targeting students' interests and expectancy beliefs in STEM areas as well as gender differences in these fields (e.g., Lazowski & Hulleman, 2016; Rosenzweig & Wigfield, 2016). A broad range of motivation interventions have been shown to be very successful at fostering students' motivation (see Lazowski & Hulleman, 2016; Rosenzweig & Wigfield, 2016). Even brief interventions targeting utility values, for instance, have been shown to have positive effects on students' value beliefs in mathematics (Gaspard, Dicke, Flunger, Brisson, et al., 2015). Our findings suggest that the success of interventions might be related to the outcome under investigation. In particular, to increase the number of students who choose STEM majors, targeting their realistic and investigative vocational interests could be more effective than targeting students' value beliefs because these interests were better predictors of STEM choice in our study. The suggestion to target expectancy-value constructs or vocational interests, however, comes along with questions about the factors of influence and the stability of the constructs over time.

A variety of influences on students' expectancy and value beliefs have been identified such as the role of teachers or classroom composition (e.g., Chmielewski, Dumont, & Trautwein, 2013; Schurtz, Pfost, Nagengast, & Artelt, 2014; Wang & Eccles, 2013). Such factors have also been shown to influence gender differences in STEM outcomes (e.g., Hübner et al., 2017). Conversely, much less is known about contextual influences on vocational interests, although a recent study suggested that contextual factors might also affect students' vocational interests as well as gender differences in realistic and investigative interests (Hübner et al., 2017). In addition, little is known about the development of vocational interests during childhood and adolescence, whereas there is a large body of research indicating how students' expectancy and value beliefs develop over time (see Wigfield et al., 2015). There is research indicating that students' vocational interests develop during the early school years (Tracey, 2002). A meta-analysis by Low, Yoon, Roberts, and Rounds (2005), however, suggested that vocational interests remain relatively stable during adolescence when career choices develop and manifest. Thus, even though our study suggests it might be more successful to intervene in students' vocational interests than in their expectancy and value beliefs, little is known about how such interventions need to be designed. In order to understand developmental and contextual influences, more research is necessary in this regard. Further research should thereby also consider the moderating role of gender, as our results suggest that relations between vocational interests and STEM careers might differ between males and females.

### ***Limitations and Further Research***

One limitation of the present study is related to the cross-sectional data with respect to math achievement. We used a powerful data set to explore whether the constructs of interest would differentially predict STEM careers as the data set consisted of two sets of predictors as well as important covariates and varying STEM outcomes simultaneously. Furthermore, the sample size was large enough to investigate different subgroups as well as the moderating effect of gender during the transition from high school to university. Nonetheless, our analyses with respect to math achievement were cross-sectional because math achievement and its predictors were both assessed at the first measurement point, and further research is therefore necessary to investigate whether expectancy-value constructs and vocational interests can longitudinally predict math achievement in this phase of young adults' career pathways.

Our research was also limited to a specific phase in the lives of young adults, that is, the end of high school and the transition to university. This phase is central in young men's and women's career paths because the choice of a university major is one important step toward



entering different professions and has long-lasting consequences on the courses of young adults' lives. However, because previous research has also suggested that preferences for certain careers manifest early in childhood and adolescence (see Schoon & Eccles, 2014), further research should investigate how the interrelations of vocational interests and expectancy-value constructs as well as their power to predict STEM careers might develop over time to provide a more comprehensive understanding of the constraints of STEM careers. This might be important for understanding the different career pathways of men and women in particular because we found indications of differences in their relative power to predict STEM variables for men and women, at least with respect to vocational interests.

Concerning expectancy-value constructs, one might also speculate how results might differ when investigating the constructs in different or various subjects. In our study, we focused on self-concepts and task values in the domains of mathematics and English—two domains that are very important for STEM outcomes (Nagy et al., 2008). Research on expectancy-value constructs, however, has shown that the intraindividual patterns of expectancy and value beliefs across different domains (rather than absolute values in single domains) are related to different career choices (Chow, Eccles, & Salmela-Aro, 2012). Referring to such findings, expectancy-value constructs might thus be even more strongly related to indicators of STEM careers when expectancy-value constructs in other domains (e.g., science) are also considered. The present study thus provides initial evidence for the differential predictive power of expectancy-value constructs in predicting STEM careers over and above vocational interests, but further research is needed to investigate such predictions in more depth by considering expectancy-value constructs in different domains in order to extend the findings of the present study.

## References

- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*(2), 219–245.  
<https://doi.org/10.1037/0033-2909.121.2.219>
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bergmann, C., & Eder, F. (2005). *AIST-R: Allgemeiner Interessen-Struktur-Test mit Umwelt-Struktur-Test (UST-R) - Revision. [GIST-R: General interest structure test and environmental structure test (EST-R) - Revision]*. Göttingen: Beltz.
- Brown, S. D., & Gore Paul A., J. (1994). An evaluation of interest congruence indices: Distribution characteristics and measurement properties. *Journal of Vocational Behavior, 45*(3), 310–327. <https://doi.org/http://dx.doi.org/10.1006/jvbe.1994.1038>
- Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106*(4), 676–713. <https://doi.org/10.1037/0033-295x.106.4.676>
- Cheryan, S., Ziegler, S. A., Montoya, A., & Jiang, L. (2016). Why are some STEM field more gender balanced than others? *Psychological Bulletin, 143*(1), 1-35.  
<https://doi.org/10.1037/bul0000052>
- Chmielewski, a. K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal, 50*(5), 925–957.  
<https://doi.org/10.3102/0002831213489843>
- Chow, A., Eccles, J. S., & Salmela-Aro, K. (2012). Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Developmental Psychology, 48*(6), 1612–1628. <https://doi.org/10.1037/a0030194>
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression. Correlation analyses for the behavioral sciences* (3rd ed.). Mahwah, NJ, London: Lawrence Erlbaum Associates.
- Cole, J. S., Bergin, D. A., & Whittaker, T. A. (2008). Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology, 33*(4), 609–624. <https://doi.org/10.1016/j.cedpsych.2007.10.002>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, Boston, Dordrecht, London, Moscow: Kluwer Academic/Plenum Publishers.

- Denissen, J. J. A., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development, 78*(2), 430–447. <https://doi.org/10.1111/j.1467-8624.2007.01007.x>
- Eccles, J. S. (2005). Subjective task values and the Eccles et al. model of achievement related choices. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105–121). New York, NY: Guilford.
- Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78–89. <https://doi.org/10.1080/00461520902832368>
- Eccles, J. S., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–121). San Francisco, CA: W. H. Freeman & Co.
- Eccles, J. S., & Wigfield, A. (1995). In the mind of an actor: The structure of adolescents' academic achievement related-beliefs and self-perceptions. *Personality and Social Psychology Bulletin, 21*, 215–225. <https://doi.org/https://doi.org/10.1177/0146167295213003>
- 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>
- Enders, C. K. (2010). *Applied missing data analysis. Methodology in the social sciences*. New York, NY: Guilford Press.
- Federal Statistical Office. (2008). *Bildung und Kultur - Studierende an Hochschulen Wintersemester 2007/2008. Fachserie 11, Reihe 4.1 [Education and culture - Students at universities in wintersemester 2007/2008]*. (Statistisches Bundesamt, Ed.). Wiesbaden, Germany: Statistisches Bundesamt.
- 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>
- Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research, 25*(3), 201–239. <https://doi.org/10.1006/ssre.1996.0010>
- Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B., Häfner, I., Nagengast, B., & Trautwein,

- U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology, 51*(9), 1226–1240. <https://doi.org/10.1037/dev0000028>
- Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*(3), 663–677. <https://doi.org/10.1037/edu0000003>
- 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*(8), 1163–1176. <https://doi.org/10.1037/a0039440>
- Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology, 6*(1), 35–45. <https://doi.org/10.1037/h0040767>
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, Fla.: Psychological Assessment Resources.
- Hübner, N., Wille, E., Cambria, J., Oschatz, K., Nagengast, B., & Trautwein, U. (2017). Maximizing gender equality in STEM by minimizing course choice options? Differential effects of obligatory math coursework on girls' and boys' math achievement, math self-concept, and vocational interests. *Journal of Educational Psychology*, (Advance online publication). <https://doi.org/http://dx.doi.org/10.1037/edu0000183>
- Humphreys, L. G., & Yao, G. (2002). Prediction of graduate major from cognitive and self-report test scores obtained during the high school years. *Psychological Reports, 90*(1), 3–30. <https://doi.org/10.2466/PR0.90.1.3-30>
- 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>
- Köller, O., Watermann, R., Trautwein, U., & Lüdtke, O. (2004). *Wege zur Hochschulreife in Baden-Württemberg: TOSCA - eine Untersuchung an allgemein bildenden und beruflichen Gymnasien [Ways to qualification for university entrance. TOSCA – A study at traditional and vocational Gymnasium schools]*. Opladen: Leske + Budrich.
- 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>
- Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). *STEM: Good jobs now and for the future. ESA issue brief #03-11*. US Department of Commerce. 1401 Constitution Avenue NW, Washington, DC 20230. Tel: 202-482-2000; Web site: <http://www.commerce.gov>.
- Larson, L. M., Wu, T.-F., Bailey, D. C., Borgen, F. H., & Gasser, C. E. (2010). Male and female college students' college majors: The contribution of basic vocational confidence and interests. *Journal of Career Assessment, 18*(1), 16–33.  
<https://doi.org/10.1177/1069072709340520>
- 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>
- 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>
- Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, London, New Delhi: Sage.
- Low, K. S. D., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: A quantitative review of longitudinal studies. *Psychological Bulletin, 131*(5), 713–737.  
<https://doi.org/10.1037/0033-2909.131.5.713>
- Marsh, H. W. (1992). *Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self-concept: A test manual and a research monograph*. Sydney, New South Wales, Australia: Macarthur.
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective. *Perspectives on Psychological Science, 1*(2), 133–163. <https://doi.org/10.1111/j.1745-6916.2006.00010.x>
- Marsh, H. W., Craven, R. G., & Debus, R. (1998). Structure, stability, and development of youth children's self-concepts: A multicohort-multioccasion study. *Child Development, 69*(4), 1030–1053. <https://doi.org/10.2307/1132361>
- Marsh, H. W., & O'Neill, R. (1984). Self Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21*(2), 153–174. <https://doi.org/10.1111/j.1745-3984.1984.tb00227.x>

- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>
- 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>
- 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>
- Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*(1), 60–70. <https://doi.org/10.1037//0022-0663.82.1.60>
- Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *The American Statistician, 54*(1), 17–24. <https://doi.org/10.1080/00031305.2000.10474502>
- Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review, 26*(1), 67–82. <https://doi.org/10.1093/esr/jcp006>
- Mullis, I. V. S., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 international mathematics report. Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:TIMSS+2007+International+Mathematics+Report#2>
- Muthén, B. O., & Muthén, L. K. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T., & Trautwein, U. (2011). Who took the “×” out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*(8), 1058–1066. <https://doi.org/10.1177/0956797611415540>
- Nagy, G. (2006). *Berufliche Interessen, kognitive und fachgebundene Kompetenzen. Ihre Bedeutung für die Studienfachwahl und die Bewährung im Studium [Vocational interests, cognitive and domain-specific competencies. Their role for major choices and success at university]*. Berlin: FU. Retrieved from <http://www.diss.fu->

berlin.de/diss/receive/FUDISS\_thesis\_000000002714

- Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., & Eccles, J. S. (2008). Gendered high school course selection as a precursor of gendered careers: The mediating role of self-concept and intrinsic value. In J. S. Eccles & H. M. G. Watt (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 115–143). Washington, DC, US: American Psychological Association. <https://doi.org/10.1037/11706-004>
- National Science Board. (2016). *Science & engineering indicators 2016*. <https://doi.org/10.1002/ejoc.201200111>
- Noonan, R. (2017). *STEM jobs: 2017 update (ESA Issue Brief #02-17)*. Retrieved from <http://www.esa.gov/reports/stem-jobs-2017-update>
- Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2017). Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior, 98*(May 2014), 138–151. <https://doi.org/10.1016/j.jvb.2016.11.002>
- Parker, P. D., 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* (pp. 224–246). Cambridge, U.K.: Cambridge University Press.
- Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multicontext study. *Developmental Psychology, 48*(6), 1629–42. <https://doi.org/10.1037/a0029167>
- Päßler, K., Beinicke, A., & Hell, B. (2014). Gender-related differential validity and differential prediction in interest inventories. *Journal of Career Assessment, 22*(1), 138–152. <https://doi.org/10.1177/1069072713492934>
- 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>
- Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist, 46*(3), 168–184. <https://doi.org/10.1080/00461520.2011.587723>
- Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2011). Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity.

- Science Education*, 95(3), 458–476. <https://doi.org/10.1002/sce.20431>
- Rolfhus, E. L., & Ackerman, P. L. (1996). Self-report knowledge: At the crossroads of ability, interest, and personality. *Journal of Educational Psychology*, 88(1), 174–188. <https://doi.org/10.1037/0022-0663.88.1.174>
- Roloff Henoch, J., Klusmann, U., Lüdtke, O., & Trautwein, U. (2015). Who becomes a teacher? Challenging the “negative selection” hypothesis. *Learning and Instruction*, 36, 46–56. <https://doi.org/10.1016/j.learninstruc.2014.11.005>
- 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>
- Rounds, J., & Su, R. (2014). The nature and power of interests. *Current Directions in Psychological Science*, 23(2), 98–103. <https://doi.org/10.1177/0963721414522812>
- Sáinz, M., & Eccles, J. S. (2012). Self-concept of computer and math ability: Gender implications across time and within ICT studies. *Journal of Vocational Behavior*, 80(2), 486–499. <https://doi.org/http://dx.doi.org/10.1016/j.jvb.2011.08.005>
- 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>
- Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading*, 3(3), 257–279. [https://doi.org/10.1207/s1532799xssr0303\\_4](https://doi.org/10.1207/s1532799xssr0303_4)
- Schoon, I., & Eccles, J. S. (Eds.). (2014). *Gender differences in aspiration and attainment: A life course perspective*. Cambridge, U.K.: Cambridge University Press.
- Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student’s mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction*, 34, 32–41. <https://doi.org/10.1016/j.learninstruc.2014.08.001>
- Schwanzer, A. D., Trautwein, U., Lüdtke, O., & Sydow, H. (2005). Entwicklung eines Instruments zur Erfassung des Selbstkonzepts junger Erwachsener [Development of a new instrument to assess self-concept of young adults]. *Diagnostica*, 51(4), 183–194. <https://doi.org/10.1026/0012-1924.51.4.183>
- Sells, L. W. (1980). Mathematics: The invisible filter. *Engineering Education*, 70, 340–341.
- Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2012). Charting the Eccles’ expectancy-value model from mothers’ beliefs in childhood to youths’ activities in adolescence. *Developmental Psychology*, 48(4), 1019–1032. <https://doi.org/10.1037/a0027468>
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and*



- advanced multilevel modeling*. Los Angeles, CA: Sage.
- Steinmayr, R., & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW) [Construction and initial validation of a scale for the assessment of subjective school values (SESSW)]. *Diagnostica, 56*, 195–211. <https://doi.org/10.1026/0012-1924/a000023>
- Stoll, G., & Trautwein, U. (2017). Vocational interests as personality traits. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 401–417). London, San Diego, Cambridge, Oxford: Elsevier.
- 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*, 1–20. <https://doi.org/10.3389/fpsyg.2015.00189>
- Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-analysis of sex differences in interests. *Psychological Bulletin, 135*(6), 859–884. <https://doi.org/10.1037/a0017364>
- Tracey, T. J. G. (2002). Development of interests and competency beliefs: A 1-year longitudinal study of fifth- to eighth-grade students using the ICA-R and structural equation modeling. *Journal of Counseling Psychology, 49*(2), 148–163. <https://doi.org/10.1037/0022-0167.49.2.148>
- Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology, 98*(2), 438–456. <https://doi.org/10.1037/0022-0663.98.2.438>
- Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763–777. <https://doi.org/10.1037/a0027470>
- Trautwein, U., Neumann, M., Nagy, G., Lüdtke, O., & Maaz, K. (Eds.). (2010). *Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand. [School achievement of upper secondary school students. The rearranging upper secondary school on the trial]*. Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-92037-5>
- Tynkkynen, L., Tolvanen, A., & Salmela-Aro, K. (2012). Trajectories of educational expectations from adolescence to young adulthood in Finland. *Developmental Psychology, 48*(6), 1674–85. <https://doi.org/10.1037/a0027245>
- Van Iddekinge, C. H., Roth, P. L., Putka, D. J., & Lanivich, S. E. (2011). Are you interested?

- A meta-analysis of relations between vocational interests and employee performance and turnover. *Journal of Applied Psychology*, 96(6), 1167–1194.  
<https://doi.org/10.1037/a0024343>
- Volodina, A., & Nagy, G. (2016). Vocational choices in adolescence: The role of gender, school achievement, self-concepts, and vocational interests. *Journal of Vocational Behavior*, 95–96, 58–73. <https://doi.org/10.1016/j.jvb.2016.07.005>
- 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, 1–9.  
<https://doi.org/10.3389/fpsyg.2015.00036>
- Wang, M.-T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, 28, 12–23.  
<https://doi.org/10.1016/j.learninstruc.2013.04.002>
- Wang, M.-T., & Kenny, S. (2014). Gender differences in personal aptitudes and motivational beliefs for achievement in and commitment to math and science fields. In I. Schoon & J. S. Eccles (Eds.), *Gender Differences in Aspirations and Attainment* (pp. 266–284). Cambridge, U.K.: Cambridge University Press.
- Warwas, J., Nagy, G., Watermann, R., & Hasselhorn, M. (2009). The relations of vocational interests and mathematical literacy: On the predictive power of interest profiles. *Journal of Career Assessment*, 17(4), 417–438. <https://doi.org/10.1177/1069072709339284>
- 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>
- Watt, H. M. G., Hyde, J. S., Petersen, J., Morris, Z. A., Rozek, C. S., & Harackiewicz, J. M. (2016). Mathematics a critical filter for STEM-related career choices? A longitudinal examination among Australian and U.S. adolescents. *Sex Roles*, 77, 254–274.  
<https://doi.org/10.1007/s11199-016-0711-1>
- Watt, H. M. G., Richardson, R. W., & Devos, C. (2013). (How) Does Gender Matter in the Choice of a STEM Teaching Career and Later Teaching Behaviours? *International Journal of Gender, Science and Technology*, 5(3), 187–206. Retrieved from <http://genderandset.open.ac.uk/index.php/genderandset/article/viewArticle/331>
- 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>
- Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review, 30*(1), 1–35. <https://doi.org/10.1016/j.dr.2009.12.001>
- 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., Fredricks, J. A., Simpkins, S. D., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. In R. M. Lerner, C. Garcia Coll, & M. Lamb (Eds.), *Handbook of child psychology and developmental science* (pp. 1–44). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118963418.childpsy316>
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. a., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*(3), 451–469. <https://doi.org/10.1037/0022-0663.89.3.451>

## Appendix

### Wording of the Items used to Assess Students' Expectancies for Success and Task Values

---

#### Attainment Value

- I'm really keen to learn a lot in mathematics/English.
- It is important for me to remember what I learn in mathematics/English.
- Mathematics/English is important to me personally.
- It is important to me personally to be a good mathematician/good at English.

#### Intrinsic Value

- I enjoy puzzling over mathematics/English problems.
- I would like to have more mathematics/English lessons.
- When I'm working on a mathematics/English problem, I sometimes don't notice time passing.
- I always look forward to mathematics/English lessons.
- If I can learn something new in mathematics/English, I'm prepared to use my free time to do so.

#### Utility Value

- I'll need good mathematics/English skills for my later life (training, studies, work).
- Good grades in mathematics/English can be of great value to me later.
- We'll often benefit from what we learn in mathematics/English.

#### Cost

- I'd have to sacrifice a lot of free time to be good at mathematics/English.
- I'd have to invest a lot of time to get good grades in mathematics/English.

#### Self-Concept Mathematics

- I have always been good at mathematics.
- I'm good at mathematics.
- I have difficulty understanding everything to do with mathematics.
- I have never been good at tasks that require mathematical thinking.

#### Self-Concept English

- English isn't really my thing.
  - I'm good at English.
  - I have difficulty understanding everything to do with English.
  - I'm just not good at English.
-

## Supplemental Materials

### *1 Additional analyses using the aspiration to choose a STEM major and mathematically intensive STEM major, respectively, as outcomes instead of the actual choice* *(Tables S1-S2)*

In the manuscript, we report results for the choice of different STEM university majors stated 2 years after graduation because the choice of university major is an important step in a STEM career pathway. Nevertheless, we also used the aspiration to choose a certain university major stated at the end of high school as an outcome variable because aspirations are also often used as outcome measures in research on expectancy-value constructs (e.g., Lauermann et al., 2017; Watt et al., 2012). We therefore present results for the aspiration to choose a certain major in the supplemental materials.

### *2 Robustness checks using only the subsample in which task values were assessed* *(Tables S3-S5)*

Task values were assessed in only a subsample of students whose data were used in the present study. We used the full information maximum likelihood approach as implemented in Mplus 7.3 to handle missing data in the analyses reported in the manuscript. As robustness checks, we ran the same analyses for only the subsample for which task values were assessed and present the results in the supplemental materials. The results followed the same pattern as the results presented in the manuscript.

Table S1

*Predicting Aspiration of STEM University Major: Results of Logistic Regression Analyses*

Predictor	Model I				Model II				Model III				Model IV							
	$\beta$	OR	SE	p	OR 95% CI	$\beta$	OR	SE	p	OR 95% CI	$\beta$	OR	SE	p	OR 95% CI					
Gender (f = 1)	-0.25	0.78	0.19	<.001	[0.74, 0.83]	-0.19	0.83	0.20	<.001	[0.78, 0.88]	0.14	1.15	0.20	<.001	[1.09, 1.21]	0.11	1.12	0.20	.001	[1.06, 1.17]
School type (VG = 1)	0.05	1.05	0.18	.354	[0.96, 1.15]	0.02	1.02	0.18	.758	[0.92, 1.11]	-0.04	0.96	0.12	.297	[0.89, 1.02]	-0.07	0.93	0.12	.083	[0.86, 1.00]
HISEI	0.01	1.01	0.02	.716	[0.97, 1.04]	0.01	1.01	0.02	.575	[0.98, 1.24]	0.02	1.02	0.02	.358	[0.99, 1.05]	0.02	1.02	0.02	.417	[0.98, 1.05]
Immigration (= 1)	0.05	1.05	0.22	.354	[0.96, 1.14]	0.03	1.03	0.22	.634	[0.94, 1.11]	0.09	1.10	0.20	.062	[1.01, 1.18]	0.07	1.07	0.20	.186	[0.98, 1.15]
Math AC	0.35	1.42	0.03	<.001	[1.36, 1.49]	0.12	1.13	0.04	.001	[1.07, 1.20]	0.16	1.18	0.03	<.001	[1.13, 1.23]	0.05	1.05	0.03	.065	[1.01, 1.11]
Engl. AC	-0.18	0.83	0.02	<.001	[0.80, 0.87]	-0.02	0.98	0.03	.501	[0.92, 1.03]	-0.11	0.90	0.02	<.001	[0.87, 0.93]	-0.08	0.93	0.03	.011	[0.88, 0.97]
Realistic						0.29	1.34	0.03	<.001	[1.28, 1.40]	0.29	1.34	0.03	<.001	[1.28, 1.40]	0.26	1.30	0.03	<.001	[1.24, 1.36]
Investigative						0.39	1.48	0.03	<.001	[1.42, 1.54]	0.39	1.48	0.03	<.001	[1.42, 1.54]	0.35	1.42	0.03	<.001	[1.36, 1.48]
Artistic						-0.28	0.76	0.02	<.001	[0.73, 0.79]	-0.28	0.76	0.02	<.001	[0.73, 0.79]	-0.22	0.80	0.02	<.001	[0.77, 0.84]
Social						0.09	1.09	0.03	.001	[1.04, 1.14]	0.09	1.09	0.03	.001	[1.04, 1.14]	0.09	1.09	0.03	.001	[1.04, 1.14]
Enterprising						-0.17	0.85	0.03	<.001	[0.81, 0.89]	-0.17	0.85	0.03	<.001	[0.81, 0.89]	-0.16	0.85	0.03	<.001	[0.81, 0.90]
Conventional						-0.26	0.77	0.03	<.001	[0.74, 0.81]	-0.26	0.77	0.03	<.001	[0.74, 0.81]	-0.28	0.75	0.03	<.001	[0.72, 0.79]
Math value						0.29	1.33	0.05	<.001	[1.23, 1.44]						0.18	1.19	0.04	<.001	[1.11, 1.28]
Math cost						-0.01	0.99	0.05	.857	[0.91, 1.08]						0.03	1.03	0.05	.500	[0.96, 1.11]
Engl. value						-0.14	0.87	0.05	.005	[0.80, 0.94]						-0.08	0.93	0.04	.080	[0.86, 1.00]
Engl. cost						0.04	1.04	0.06	.483	[0.95, 1.14]						0.03	1.03	0.05	.485	[0.96, 1.12]
Math SC						0.06	1.07	0.05	.205	[0.98, 1.16]						0.09	1.09	0.04	.035	[1.02, 1.17]
Engl. SC						-0.01	0.99	0.05	.868	[0.92, 1.07]						0.07	1.07	0.04	.093	[1.00, 1.15]
Pseudo-R <sup>2</sup>									.207										.374	
									.117										.357	

*Note.* The table displays standardized results. Mathematics, natural sciences, and engineering were coded as STEM subjects. Linguistics, cultural studies, law, business, and social sciences were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of aspiring a STEM subject, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; HISEI = highest international socioeconomic index; VG = vocational Gymnasium; AC = achievement; Engl. = English; SC = self-concept.

Table S2

## Predicting Aspiration of Mathematically intensive STEM University Major: Results of Logistic Regression Analyses

Predictor	Model I			Model II			Model III			Model IV										
	$\beta$	OR	SE	$\beta$	OR	SE	$\beta$	OR	SE	$\beta$	OR	SE	$p$	OR 95% CI						
Gender (f = 1)	-0.52	0.60	0.22	<.001	0.54, 0.66]	-0.45	0.64	0.23	<.001	0.58, 0.70]	-0.12	0.89	0.27	.038	[0.80, 0.98]	-0.14	0.87	0.29	.023	[0.78, 0.96]
School type (VG = 1)	0.19	1.21	0.24	.020	[1.06, 1.34]	0.17	1.18	0.25	.051	[1.03, 1.33]	0.24	1.27	0.21	.001	[1.14, 1.39]	0.23	1.26	0.22	.001	[1.13, 1.38]
HISEI	-0.12	0.89	0.04	.001	[0.84, 0.94]	-0.09	0.91	0.03	.006	[0.86, 0.96]	-0.11	0.90	0.03	<.001	[0.85, 0.94]	-0.10	0.90	0.03	.001	[0.86, 0.95]
Immigration (= 1)	0.08	1.09	0.28	.283	[0.95, 1.22]	0.01	1.01	0.29	.880	[0.88, 1.15]	0.14	1.15	0.29	.086	[1.01, 1.29]	0.04	1.05	0.26	.558	[0.92, 1.17]
Math AC	0.36	1.43	0.04	<.001	[1.34, 1.53]	0.09	1.09	0.05	.082	[1.01, 1.19]	0.34	1.40	0.04	<.001	[1.32, 1.49]	0.11	1.11	0.05	.017	[1.03, 1.20]
Engl. AC	-0.19	0.83	0.04	<.001	[0.77, 0.88]	-0.03	0.97	0.05	.561	[0.89, 1.06]	-0.11	0.90	0.04	.008	[0.84, 0.96]	0.01	1.01	0.05	.842	[0.93, 1.10]
Realistic						0.54	1.72	0.04	<.001	[1.60, 1.84]	0.48	1.62	0.05	<.001	[1.50, 1.75]					
Investigative						-0.38	0.68	0.04	<.001	[0.64, 0.73]	-0.42	0.66	0.04	<.001	[0.62, 0.70]					
Artistic						-0.09	0.92	0.04	.015	[0.86, 0.97]	0.01	1.01	0.04	.869	[0.94, 1.07]					
Social						-0.19	0.83	0.04	<.001	[0.78, 0.89]	-0.21	0.81	0.05	<.001	[0.75, 0.88]					
Enterprising						0.10	1.10	0.05	.034	[1.02, 1.19]	0.13	1.14	0.05	.003	[1.06, 1.23]					
Conventional						0.05	1.05	0.04	.194	[0.99, 1.12]	-0.02	0.98	0.05	.644	[0.91, 1.06]					
Math value						0.46	1.58	0.05	<.001	[1.45, 1.73]						0.42	1.52	0.06	<.001	[1.38, 1.67]
Math cost						-0.02	0.99	0.07	.838	[0.88, 1.11]						-0.05	0.95	0.07	.422	[0.85, 1.06]
Engl. value						-0.07	0.94	0.06	.296	[0.85, 1.04]						-0.08	0.92	0.06	.210	[0.83, 1.02]
Engl. cost						0.03	1.03	0.06	.663	[0.93, 1.13]						-0.01	0.99	0.06	.915	[0.90, 1.09]
Math SC						0.02	1.02	0.06	.772	[0.92, 1.12]						0.02	1.02	0.06	.767	[0.92, 1.12]
Engl. SC						-0.03	0.98	0.05	.640	[0.89, 1.07]						-0.03	0.98	0.05	.616	[0.90, 1.06]
Pseudo-R <sup>2</sup>									.295					.338					.407	

Note. The table displays standardized results. Mathematics, physics, computer science, and engineering were coded as mathematically intensive STEM subjects. Biology and medicine were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of aspiring a mathematically intensive STEM subject, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; HISEI = highest international socioeconomic index; VG = vocational Gymnasium; AC = achievement; Engl. = English; SC = self-concept.

Table S3

*Predicting Mathematics Achievement for the Subsample in which Task Values were Assessed: Results of Multiple Regression Analyses*

Predictor	Model I			Model II			Model III			Model IV			
	$\beta$	SE	p	$\beta$	SE	p	$\beta$	SE	p	$\beta$	SE	p	95% CI
Gender (f = 1)	-0.20	0.18	<.001	-0.08	0.15	.001	0.03	0.21	.452	-0.04	0.17	.223	[-0.09, 0.01]
School type (VG = 1)	-0.30	0.13	<.001	-0.29	0.11	<.001	-0.33	0.13	<.001	-0.30	0.11	<.001	[-0.39, -0.23]
HISEI	0.04	0.02	.038	0.03	0.02	.096	0.06	0.02	.005	0.03	0.02	.058	[0.01, 0.06]
Immigration (= 1)	-0.15	0.20	.002	-0.10	0.16	.017	-0.11	0.19	.023	-0.09	0.16	.023	[-0.17, -0.02]
Engl. AC	0.29	0.02	<.001	0.31	0.02	<.001	0.30	0.02	<.001	0.29	0.03	<.001	[0.25, 0.33]
Realistic							0.09	0.03	.001	0.04	0.03	.153	[-0.01, 0.08]
Investigative							0.19	0.03	<.001	0.04	0.03	.106	[0.00, 0.08]
Artistic							-0.18	0.02	<.001	-0.03	0.02	.204	[-0.06, 0.01]
Social							0.00	0.03	.971	0.01	0.02	.770	[-0.03, 0.05]
Enterprising							-0.07	0.03	.011	-0.05	0.02	.042	[-0.09, -0.01]
Conventional							0.03	0.03	.250	0.00	0.02	.864	[-0.04, 0.03]
Math value				0.10	0.02	<.001				0.08	0.02	.001	[0.04, 0.12]
Math cost				-0.08	0.02	<.001				-0.07	0.02	.001	[-0.10, -0.04]
Engl. value				0.00	0.03	.912				0.01	0.03	.843	[-0.04, 0.05]
Engl. cost				-0.04	0.02	.086				-0.04	0.02	.079	[-0.08, 0.00]
Math SC				0.36	0.03	<.001				0.36	0.03	<.001	[0.32, 0.40]
Engl. SC				-0.09	0.03	.002				-0.08	0.03	.006	[-0.13, -0.03]
R <sup>2</sup>			.230			.480			.308			.485	

*Note.* The table displays standardized results. CI = confidence interval; HISEI = highest international socioeconomic index; VG = vocational Gymnasium; AC = achievement; Engl. = English; SC = self-concept.



Table S4

*Predicting Choice of STEM University Major for the Subsample in which Task Values were Assessed: Results of Logistic Regression Analyses*

Predictor	Model I				Model II				Model III				Model IV							
	$\beta$	OR	SE	p	OR 95%CI	$\beta$	OR	SE	p	OR 95%CI	$\beta$	OR	SE	p	OR 95%CI					
Gender (f = 1)	-0.26	0.77	0.29	<.001	[0.70, 0.84]	-0.21	0.81	0.28	<.001	[0.74, 0.89]	0.12	1.13	0.31	.034	[1.03, 1.23]	0.10	1.10	0.31	.077	[1.01, 1.21]
School type (VG = 1)	0.25	1.28	0.24	.004	[1.12, 1.42]	0.17	1.19	0.25	.051	[1.03, 1.34]	0.09	1.09	0.20	.214	[0.97, 1.21]	0.03	1.03	0.21	.678	[0.90, 1.15]
HISEI	0.02	1.02	0.04	.579	[0.96, 1.09]	0.04	1.04	0.04	.324	[0.97, 1.46]	0.02	1.02	0.03	.573	[0.96, 1.08]	0.03	1.03	0.03	.323	[0.98, 1.10]
Immigration (= 1)	-0.04	0.96	0.28	.641	[0.82, 1.10]	-0.04	0.96	0.25	.581	[0.83, 1.08]	-0.04	0.96	0.23	.552	[0.84, 1.07]	-0.05	0.95	0.22	.428	[0.83, 1.06]
Math AC	0.39	1.48	0.04	<.001	[1.37, 1.59]	0.16	1.17	0.06	.004	[1.07, 1.28]	0.20	1.23	0.04	<.001	[1.14, 1.32]	0.09	1.10	0.05	.056	[1.01, 1.19]
Engl. AC	-0.07	0.94	0.04	.097	[0.88, 1.00]	0.01	1.01	0.05	.798	[0.94, 1.09]	-0.04	0.97	0.04	.402	[0.90, 1.03]	-0.06	0.94	0.05	.215	[0.86, 1.02]
Realistic						0.23	1.26	0.04	<.001	[1.17, 1.36]	0.23	1.25	0.05	<.001	[1.16, 1.35]					
Investigative						0.43	1.53	0.05	<.001	[1.42, 1.65]	0.37	1.45	0.05	<.001	[1.35, 1.57]					
Artistic						-0.19	0.82	0.05	<.001	[0.76, 0.89]	-0.15	0.86	0.05	.003	[0.80, 0.94]					
Social						0.04	1.04	0.05	.348	[0.97, 1.13]	0.03	1.03	0.05	.526	[0.95, 1.12]					
Enterprising						-0.14	0.87	0.05	.008	[0.80, 0.95]	-0.14	0.87	0.05	.006	[0.80, 0.94]					
Conventional						-0.19	0.83	0.04	<.001	[0.77, 0.89]	-0.23	0.80	0.04	<.001	[0.74, 0.85]					
Math value						0.31	1.36	0.05	<.001	[1.25, 1.48]						0.17	1.19	0.05	.001	[1.09, 1.30]
Math cost						0.02	1.02	0.05	.667	[0.94, 1.12]						0.04	1.04	0.05	.474	[0.95, 1.13]
Engl. value						0.01	1.01	0.05	.868	[0.93, 1.09]						0.08	1.08	0.05	.112	[1.00, 1.17]
Engl. cost						-0.04	0.96	0.05	.408	[0.88, 1.04]						-0.05	0.95	0.05	.320	[0.88, 1.03]
Math SC						0.09	1.09	0.07	.184	[0.98, 1.22]						0.11	1.12	0.07	.108	[1.00, 1.26]
Engl. SC						-0.05	0.95	0.07	.447	[0.86, 1.06]						-0.01	0.99	0.06	.891	[0.89, 1.10]
Pseudo-R <sup>2</sup>									.211					.337					.359	

*Note.* The table displays standardized results. Mathematics, natural sciences, and engineering were coded as STEM subjects. Linguistics, cultural studies, law, business, and social sciences were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of studying a STEM subject, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; HISEI = highest international

Table S5

*Predicting Choice of Mathematically intensive STEM University Major for the Subsample in which Task Values were Assessed: Results of Logistic Regression Analyses*

Predictor	Model I				Model II				Model III				Model IV							
	$\beta$	OR	SE	p	$\beta$	OR	SE	p	$\beta$	OR	SE	p	$\beta$	OR	SE	p				
Gender (f = 1)	-0.64	0.53	0.33	<.001	[0.44, 0.62]	-0.58	0.56	0.36	<.001	[0.46, 0.67]	-0.23	0.80	0.40	.029	[0.66, 0.95]	-0.20	0.82	0.42	.071	[0.67, 0.98]
School type (VG = 1)	-0.09	0.91	0.39	.532	[0.69, 1.14]	-0.11	0.90	0.40	.482	[0.67, 1.14]	-0.20	0.82	0.31	.080	[0.64, 0.99]	-0.19	0.83	0.30	.089	[0.65, 0.99]
HISEI	-0.16	0.85	0.08	.036	[0.75, 0.97]	-0.13	0.88	0.08	.106	[0.78, 1.00]	-0.18	0.83	0.07	.005	[0.75, 0.93]	-0.15	0.86	0.06	.018	[0.77, 0.96]
Immigration (= 1)	-0.14	0.87	0.37	.317	[0.67, 1.08]	-0.19	0.83	0.37	.153	[0.63, 1.03]	-0.34	0.71	0.34	.005	[0.54, 0.88]	-0.37	0.69	0.35	.004	[0.52, 0.87]
Math AC	0.12	1.13	0.08	.112	[1.00, 1.27]	-0.07	0.93	0.08	.405	[0.81, 1.07]	0.13	1.13	0.06	.040	[1.03, 1.25]	-0.04	0.96	0.07	.566	[0.85, 1.08]
Engl. AC	-0.20	0.82	0.07	.005	[0.72, 0.92]	-0.07	0.94	0.10	.493	[0.80, 1.10]	-0.20	0.82	0.07	.006	[0.73, 0.92]	-0.06	0.94	0.08	.469	[0.83, 1.08]
Realistic						0.60	1.82	0.07	<.001	[1.61, 2.05]						0.56	1.75	0.08	<.001	[1.53, 1.99]
Investigative						-0.48	0.62	0.08	<.001	[0.55, 0.71]						-0.55	0.58	0.08	<.001	[0.51, 0.66]
Artistic						-0.15	0.87	0.07	.037	[0.77, 0.97]						-0.04	0.96	0.08	.630	[0.85, 1.09]
Social						-0.35	0.71	0.08	<.001	[0.62, 0.80]						-0.43	0.65	0.08	<.001	[0.57, 0.75]
Enterprising						0.28	1.33	0.09	.001	[1.15, 1.53]						0.32	1.37	0.09	<.001	[1.19, 1.58]
Conventional						-0.04	0.97	0.07	.609	[0.86, 1.08]						-0.07	0.93	0.07	.303	[0.82, 1.04]
Math value						0.30	1.35	0.09	.001	[1.16, 1.57]						0.26	1.29	0.09	.004	[1.12, 1.50]
Math cost						-0.04	0.97	0.09	.699	[0.83, 1.12]						-0.05	0.95	0.07	.419	[0.85, 1.06]
Engl. value						0.03	1.03	0.08	.729	[0.90, 1.18]						-0.06	0.94	0.08	.448	[0.84, 1.07]
Engl. cost						0.09	1.09	0.09	.360	[0.93, 1.27]						0.15	1.17	0.08	.055	[1.02, 1.33]
Math SC						-0.04	0.97	0.09	.696	[0.83, 1.12]						0.03	1.03	0.09	.733	[0.89, 1.19]
Engl. SC						-0.06	0.94	0.11	.560	[0.79, 1.12]						0.03	1.03	0.09	.750	[0.88, 1.20]
Pseudo-R <sup>2</sup>									.217										.345	.398

*Note.* The table displays standardized results. Mathematics, physics, computer science, and engineering were coded as mathematically intensive STEM subjects. Biology and medicine were included in the reference group. Odds Ratios significantly larger than 1 indicate a higher likelihood of studying a mathematically intensive STEM subject, adjusted for all other predictors in the model. OR = Odds ratio; CI = confidence interval; HISEI = highest international socioeconomic index; VG = vocational Gymnasium; AC = achievement; Engl. = English; SC = self-concept.

GENERAL  
DISCUSSION



---

## 5 General Discussion

Women are less likely than men to choose mathematically intensive STEM careers, which is critical in terms of gender equity in rights and opportunities (European Commission, 2013; Eurostat, 2017; National Science Foundation, 2017; Noonan, 2017). Such gender differences have been linked to differences in females' and males' motivation, which are themselves rooted in gendered socialization processes (Eccles, 2009). In order to deepen our understanding of gender differences in mathematically intensive STEM fields, this dissertation investigated the role of selected individual and environmental factors in females' and males' socialization on gender differences in mathematically intensive STEM career pathways. To this end, three empirical studies were conducted. The first two studies were designed to investigate the influence of selected environmental features: Children's experiences of stereotypes in the media that associate mathematics more with males than with females, and coursework requirements that oblige all students to take advanced math courses in high school. The third study was designed to investigate the role of individual motivational factors in gender differences in mathematically intensive STEM careers from a more multidisciplinary perspective: Constructs from expectancy-value theory and Holland's theory of vocational interests were compared in terms of their relative predictive validity for gendered STEM careers. The findings of the three empirical studies are summarized and discussed within their broader research context in the following. General strengths and limitations of this dissertation are identified next, before the implications of this dissertation for future research, policy and practice are discussed in the final section.

## 5.1 Discussion of General Findings

### 5.1.1 Summary of central findings

The three empirical studies that made up this dissertation examined the role of selected individual and environmental factors on gender differences in mathematically intensive STEM career pathways during childhood, adolescence, and young adulthood. In Study 1, a randomized study was conducted to investigate the role of gender stereotypes embedded in a children's television program on 10-year-old girls' and boys' stereotype endorsement, math motivation, and performance. Overall, the results did not indicate that children were strongly affected by the stereotypes embedded in the television program in the short term. However, there were some effects on the different outcomes investigated: Both girls and boys reported higher stereotype endorsement after watching the video that included the stereotypes compared to the control group. Furthermore, whereas there were no effects on girls' motivation, there were some effects on boys' motivation (boys showed a higher sense of belonging but lower social utility after watching the video that included the stereotypes compared to boys in the control condition). Neither girls' nor boys' math performance was affected.

Study 2 investigated the effects of a school reform in a large German state that required all students to take advanced math courses on gender differences in math achievement, math self-concept, realistic and investigative interests at the end of high school, and the choice of STEM majors at university two years after graduation. To this end, data from two cohorts were compared (i.e., students who were in high school two years before and two years after the reform was implemented). Although gender differences favoring young men were identified in all outcomes under investigation in the cohorts before and after the reform, the reform was differentially associated with gender differences in STEM in two respects: First, the reform was associated with differential effects for young women and men. Second, the pattern of these differential effects differed with respect to the studied outcomes. Whereas gender differences in math achievement were smaller in the cohort after the reform, differences in math self-concept and realistic and vocational interests were larger. Yet even here, different patterns were found for self-concept and vocational interests. The larger gender differences in math self-concept were driven by lower self-concepts among young women after the reform, whereas young men's self-concepts in the cohorts before and after the reform did not differ. The larger gender differences in realistic and vocational interests were driven by higher interest among young men, whereas young women's interests did not differ or

differed only slightly before and after the reform. Gender differences with respect to mathematically intensive STEM university majors, however, were similar in the cohorts before and after the reform.

Aiming to gain further insights into the different motivational aspects involved in gendered career pathways, the relative predictive validity of expectancy-value constructs and vocational interests for gender differences in mathematically intensive STEM careers were examined. The results revealed that both sets of variables (i.e., expectancy-value constructs, vocational interests) predicted math achievement as well as the choice of different STEM majors. However, the predictive power differed according to the specification of the constructs: whereas expectancy-value constructs were more predictive of math achievement, vocational interests were more predictive of the choice of STEM majors in general and mathematically intensive STEM majors in particular. Furthermore, differences were found with respect to the role of gender: whereas the predictive power of expectancy-value constructs for STEM careers was invariant over gender, there were some differences in the relative power of vocational interests to predict STEM careers between young women and men.

Taken together, the results of all three empirical studies strengthen the assumption that gendered career pathways within mathematically intensive STEM fields are complex. They contribute to our understanding of how environmental factors influence these gender differences as well as how various motivation and interest constructs are differentially involved in gendered STEM career pathways. Both of these aspects are further discussed with respect to the complexity of gender differences in mathematically intensive STEM careers in the following section.

### **5.1.2 Complexity of gendered career pathways**

It is well-known that gender differences in mathematically intensive STEM careers are rooted in complex processes, and a variety of different factors and constructs have been shown to be related to differences in women's and men's career pathways (Ceci et al., 2009; Cheryan et al., 2016; Eccles, 2009). Eccles et al.'s (1983) expectancy-value model provides a comprehensive framework for investigating gender differences in STEM careers. This theory incorporates a variety of motivational constructs that predict achievement and achievement-related choices such as career choices (i.e., expectancies for success and different task values) as well as a variety of factors that influence gender differences in expectancy and value beliefs (for some recent reviews, see Wang & Degol, 2013; Wigfield et al., 2015). The present

dissertation addressed this complexity by considering a variety of motivational constructs that are related to gendered career pathways as well as two critical environmental factors that may influence gender differences in these constructs. Studies 1 and 2 examined the differential effects of environmental factors (i.e., gender stereotypes embedded in a television program and obligatory advanced math courses) on different constructs related to gendered STEM careers. Study 3 investigated associations between different motivation and interest constructs and gendered STEM careers. The current section further discusses the results of these three studies with respect to the following aspects: (1) differential influences of environmental factors on female and male students, (2) the role of single experiences, and (3) differential associations between motivation and interest constructs and gender differences in STEM careers.

### ***Differential influences of environmental factors on female and male students***

A large body of research has indicated that girls and boys might have different mathematics and science experiences while growing up, which likely influence gender differences in mathematically intensive STEM careers (see Schoon & Eccles, 2014; Wang & Degol, 2013; Watt & Eccles, 2008). These different experiences can stem from being treated differently by those around them, such as their parents and teachers (for reviews see Gunderson, Ramirez, Levine, & Beilock, 2012; Jacobs & Eccles, 2000). For example, parents tend to provide more math and science learning experiences for their sons than their daughters (Jacobs & Bleeker, 2004). In addition, differences in choices between females and males influence gender differences in the amount of time spent on math and science tasks and activities, which can ultimately influence career pathways (see Ceci et al., 2014; Watt & Eccles, 2008). Examples of such different choices include enrollment in advanced math courses in high school and participation in extracurricular science activities, both of which have been reported to be more frequency among boys than girls (Budde, 2009; M. G. Jones, Howe, & Rua, 2000; Kennedy et al., 2014; National Science Board, 2016).

The results of the present dissertation, however, are in line with prior research indicating that girls and boys might experience things differently even when they share the same tasks and activities (Ellis, Fosdick, & Rasmussen, 2016; Ma & Johnson, 2008; Walton & Cohen, 2003). In Study 1, differential effects on girls' and boys' math motivation were found after watching a television program. However, this television program contained also a section on stereotypes that linked mathematics more with males than with females. Previous research has suggested that experiencing such stereotypes might negatively influence girls'



math performance and motivation, but positively influence boys' math performance and motivation, a phenomenon referred to as *stereotype lift* (Johnson, Barnard-Brak, Saxon, & Johnson, 2012; Walton & Cohen, 2003). In a meta-analysis of 43 studies, Walton and Cohen (2003) found that individuals perform better when a test is linked with a negative stereotype about members of an outgroup; for example, men perform better when negative stereotypes about females' math ability are made salient prior to testing. As the video contained stereotypes about gender and mathematics, it is therefore reasonable that children were differently influenced by the stereotypes depending on the group (i.e., gender) they belong to, despite having seen the same video. However, this mechanism might be more complex in Study 1, as different stereotypes were presented (i.e., the stereotype that mathematics is more associated with males than with females and the stereotype that boys who are good at mathematics are geeky), as discussed in detail within the manuscript. Furthermore, children saw the video only once and the section about stereotypes was very short. Thus, although the results provide first insights into how boys and girls might be differently influenced by watching television programs that contain gender stereotypes, more needs to be known about the effects of repeatedly experiencing such stereotypes in television programs.

Whereas the experimental material in Study 1 directly contained a gender-specific feature (i.e., the short clip about stereotypes), this was not the case in Study 2, which investigated the effects of a school reform that required all students to take advanced courses in mathematics. Although young women and men attended all the same courses after the reform, differential effects of the reform on young women's and men's math motivation and achievement were found. The different effects for female and male students might be due to the fact that prior to the reform young women were less likely to choose advanced courses than young men and that high school coursework is differently associated with achievement and motivation (Byun et al., 2015; Gamoran & Mare, 1989; National Science Board, 2016), as discussed in the manuscript. Nevertheless, the findings are in line with prior research, which also found differential effects of math courses on female and male students (Ellis et al., 2016; Ma & Johnson, 2008; Ro & Knight, 2016). For example, Ma and Johnson (2008) found that taking specific math courses (i.e., calculus) in high school predicted the choice of science and engineering majors over and above achievement, career aspirations, and student characteristics (i.e., age, race/ethnicity, and family characteristics) only for young women, not for young men. Furthermore, Ellis et al. (2016) found that taking Calculus I in college was differentially associated with intentions to take further calculus courses among young women and men who had indicated at the start of the course that they planned to enroll in further

calculus courses. When the same students were asked again at the end of the course, women were less likely than men to intend to take further calculus courses after controlling for prior achievement, career aspirations, instruction quality, and institution, and were thus more likely to drop out of the “STEM pipeline.” Therefore, the results of both studies indicate that similar educational factors seem to not necessarily affect all students in similar ways. In the study by Ellis et al. (2016), the authors found some indications that lower math ability beliefs among females than males (even at the same levels of achievement) might be one potential mechanism associated with such differential effects. In line with these findings, the results of Study 2 indicated differential effects of advanced math courses on young women’s and men’s self-concepts. Nevertheless, more research is needed to explore such mechanisms in detail in order to understand the influence of math courses on gender differences in mathematically intensive STEM careers.

### *The role of single experiences*

Although the environmental factors investigated in the present dissertation had differential effects, the impact of these two factors seems to be rather small. There was little support that children were strongly affected by the stereotypes embedded in the television program in Study 1. Furthermore, although there were differential effects on achievement, self-concept, and realistic and investigative interests in Study 2, gender differences in the choice of STEM majors at university did not differ before and after the reform. Possible explanations for these findings are discussed in the manuscript. The results of both studies, however, are in line with prior research suggesting that career choices are complex processes that involve numerous factors, including stereotypes in the media and advanced math courses in high school (see Schoon & Eccles, 2014). Due to the large number of factors that influence gender differences in mathematically intensive STEM domains, the influence of any one factor might thus be rather small. Nevertheless, each individual factor might be one small piece contributing to the whole process of gendered career pathways. In this respect, the cumulative experience of different factors also has to be considered when interpreting the results of the present dissertation. The gender stereotypes investigated in Study 1 were presented only once for a very short period of time. The reform examined in Study 2 influenced students’ upper secondary high school coursework, and thus occurred over a longer time-span. However, the reform influenced students’ coursework in upper secondary school only, not their learning environments in all the school years before. Furthermore, the reform did not affect the wider context students grew up in, including such factors as their

family structure, their role models, or their stereotypical views of STEM professions. Thus, although the influence of the single situation or single aspect considered in Studies 1 and 2, respectively, might seem small at first, it may be repeated experiences that accumulate over time that ultimately lead to gender differences in career pathways (Ridgeway & Correll, 2004; Schoon & Eccles, 2014). However, more research is needed to investigate how specific factors, such as repeated experiences of gender stereotypes in television programs or math learning environments, might influence females and males in a cumulative manner over the life course.

### *The influence of construct symmetry on outcomes*

A variety of different motivational constructs have been found to be associated with gender differences in mathematically intensive STEM careers (e.g., Bussey & Bandura, 1999; Schoon & Eccles, 2014; Stipek & Gralinski, 1991; Su et al., 2009). In order to gain a deeper understanding of the complexity of gender differences in mathematically intensive STEM careers, it is crucial to consider how these different constructs relate to each other and their relative ability to explain gender differences in mathematically intensive STEM careers (see Eccles, Wigfield, et al., 1993; P. K. Murphy & Alexander, 2000; Wigfield et al., 2015). In recent years, how motivational constructs from different theories relate to each other and differently relate to specific achievement-related outcomes has become a growing concern (see Pintrich, 2003a; Wigfield et al., 2015). However, as discussed in the introductory chapter (see Section 1.5), such work has thus far mainly focused on constructs within educational psychology. Study 3 expanded research in this area by examining associations between expectancy-value constructs and vocational interests and their relative associations with gender differences in mathematically intensive STEM careers. In doing so, the results of Study 3 provide further support for the need for gender differences in mathematically intensive STEM fields to be considered differentially. First, gender differences in math achievement, majors within STEM, and majors within mathematically intensive STEM fields all differed from one another. Second, differential associations were found between various motivation- and interest-based constructs (i.e., self-concept, task values, and vocational interests) and different indicators of mathematically intensive STEM careers (math achievement, different STEM majors).

In order to interpret these results more deeply, it might be useful to examine the symmetry of the variables considered. Bruswik (1955) proposed a symmetric lens model in which he argued for symmetry in the variables used to examine relations in the social

sciences. Within this model, he distinguishes between different levels of generality in variables considered within research, and claimed that symmetric variables have maximal predictive validity (Brunswik, 1955). According to this model, symmetry exists when variables considered on a broad level are used to predict variables that are also considered on a broad level. Symmetry is also given when variables considered at a more specific level are used to predict outcome variables that are also on a specific level. In contrast, predictive relations between variables at different levels are considered asymmetrical, and should have less predictive power (Brunswik, 1955). Within psychological research, Brunswik symmetry has been applied in research on intelligence and personality, for instance (Figueredo, Gladden, Sisco, Patch, & Jones, 2015; Wittmann & Süß, 1999; Zech, Bühner, Kröner, Heene, & Hilbert, 2017). Here, differential associations between different measures of intelligence and different performance measures have been found, supporting the notion that symmetry between variables plays a role in their predictive relations (Wittmann & Hattrup, 2004; Wittmann & Süß, 1999; Zech et al., 2017). Applied to the results of Study 3, this means that it is reasonable that the predictive power of vocational interests on STEM majors was higher than that of expectancy-value constructs because of a difference in symmetry: vocational interests and majors were both considered on a more general level, whereas expectancy-value constructs were considered as domain-specific constructs. In addition, it is also reasonable that the predictive power of expectancy-value constructs for math achievement was higher than that of vocational interests, because expectancy-value constructs and math achievement are both domain-specific constructs and are thus more symmetrical than math achievement and vocational interests. The results of Study 3 can therefore be interpreted as the result of different symmetries between the variables. However, more research exploring the associations between expectancy-value constructs as well as vocational interests and gender differences in mathematically intensive STEM careers—at different age groups, with different samples, and on different time spans, for example—is needed to prove such symmetries.

## 5.2 Strengths and Limitations

A few strengths and limitations should be kept in mind when interpreting the results of the three empirical studies conducted as part of this dissertation. In general, the research questions were addressed using strong research designs, including a randomized block design (Study 1) and a lagged cohort control design (Study 2). The analyses were based on strong data sets, particularly with respect to the large-scale data sets used in Study 2 and Study 3. Nevertheless, Study 1 also had a relatively large sample size, which was calculated using power analysis.

High ecological validity is another major strength of the present dissertation, as the influence of environmental factors on gender differences in mathematically intensive STEM fields were investigated in real-world situations: the effects of a television program were examined using a video that had been broadcast on national television, and the influence of encouraging young women to take advanced math courses was examined by investigating the effects of a statewide reform that required all students to take advanced math courses. On the other hand, this high ecological validity came at the price of the complexity of the influencing factors under study: Several forms of stereotypes were presented in the television program (i.e., stereotypes that associate mathematics more with males than with females, and stereotypes about males who are good at mathematics being geeky). Similar, the investigated reform also consisted of multiple aspects, as the courses before and after the reform differed with respect to the reference group, course level, and teaching time. Therefore, the unique effects of individual components of the stereotypes and the reform could not be disentangled within the present dissertation, and the concrete mechanisms underlying these findings are not yet clear.

Building upon previous findings showing that gender differences in mathematically intensive STEM careers develop over childhood, adolescence, and young adulthood (Ceci et al., 2014; Schoon & Eccles, 2014), samples of different age groups were used in the three studies that make up this dissertation. Thus, research on gender differences in mathematically intensive STEM domains was extended with respect to different developmental periods considered to be particularly relevant for the specific research questions investigated. Nevertheless, only a few selected periods were considered, and it is unclear how the environmental factors studied influence girls and boys at different ages and over larger time spans. The same holds for how the development of expectancy-value beliefs and vocational interests relate to each other and their relative predictive power for gendered STEM careers. The present dissertation examined the relative predictive power of young women's and men's

expectancy-value constructs and vocational interests for math achievement and the choice of different STEM majors at the end of high school and two years later. This time period is highly relevant for gender differences in mathematically intensive STEM careers, as university majors determine pathways into different occupations (Ceci et al., 2014). Nevertheless, gender differences in career aspirations seem to already exist by late adolescence or even earlier (Jerrim & Schoon, 2014; Riegle-Crumb et al., 2011; for a review, see Wang & Degol, 2017). In order to understand the role of expectancy-value constructs and vocational interests in the development of gender differences during adolescence, further research is needed to examine these associations at different ages as well as longitudinally.

The present dissertation focused on the domain of mathematics with respect to the environmental factors investigated and the motivational constructs considered. Mathematics is a key domain within mathematically intensive STEM fields, as already apparent from the term “mathematically intensive STEM fields”. Mathematics is the most often and widely studied domain in research on gender differences in such careers, and math achievement and motivation have found to be highly predictive of entering mathematically intensive STEM careers as well as gender differences in these fields (e.g., Guo, Parker, et al., 2015; Meece et al., 1990; Sells, 1980; see also Section 1.3). Thus, the present dissertation extended prior research on gender differences in mathematically intensive STEM fields with respect to a very central domain. Nonetheless, a broader consideration of other environmental factors and motivational constructs could provide further insights on why young women and men differ in their career pathways within mathematically intensive STEM domains. For example, with respect to the influence of environmental factors, other mathematically intensive STEM domains such as physics seem to be even more gender stereotyped and unfavorable for girls than mathematics (Kessels, Rau, & Hannover, 2006). Moreover, even fewer young women relative to men take advanced physics courses in high school compared to mathematics (Cunningham & Hoyer, 2015; National Science Board, 2016). Examining the effects of stereotypes about gender and physics or other STEM domains in television programs as well as the effects of encouraging young women to choose advanced courses in physics and other STEM domains might therefore provide additional information about the influence of stereotypes and high school coursework requirements on gender differences in mathematically intensive STEM fields.

Similarly, the three studies of the present dissertation primarily focused on students’ motivation in mathematics as a motivational construct. Although self-concept and task values in mathematics are highly predictive of mathematically intensive STEM careers (e.g., Guo,

Parker, et al., 2015; Meece et al., 1990; Watt, 2005; see also Section 1.3), previous research has also shown that patterns of intraindividual beliefs in different domains are related to gender differences in STEM careers (e.g., Chow et al., 2012; Lauermann et al., 2015; Parker et al., 2014). For example, Lauermann et al. (2015) found that gender differences in math and science career aspirations were mediated by students' English self-concept, which was negatively associated with aspirations to a math and science career. Furthermore, Chow et al. (2012) found that the extent to which students value math and science compared to other domains predicts aspirations to mathematically intensive STEM careers. Students who reported relatively higher value beliefs about mathematics and science compared to English were most likely to aspire mathematically intensive STEM careers, whereas students who reported valuing math and science less than English were less likely to aspire to such careers. Building upon such findings, Study 3 considered self-concept and task values not only in mathematics but also in English when investigating the predictive power of expectancy-value constructs and vocational interests for STEM careers. Nevertheless, the results are limited to these two domains. For German samples in particular, including students' German self-concept and value beliefs (i.e., their native language) might provide further information about the relative predictive power of expectancy-value constructs and vocational interests on gender differences in mathematically intensive STEM careers. Similarly, investigating the relative value placed on mathematics compared to other domains as might provide greater insight into such associations.

The fact that not only expectancy and value beliefs in mathematics but also intraindividual patterns of beliefs in different domains predict gendered career pathways is also important to consider in terms of the effects of gender stereotypes and coursework requirements examined in Study 1 and Study 2. Here, the question arises as to how environmental factors can influence individuals' expectancy and value beliefs not only in mathematics, but also in other domains, and how they thus might influence gender differences in mathematically intensive STEM career pathways in a more complex manner. The assumption that environmental factors might not only influence students' motivation in the respective domain but also in other related and non-related domains is based on the different processes involved in the development of expectancy and value beliefs: As introduced in the introductory chapter, students' self-concept in one domain is influenced by their evaluation of their prior achievement and by comparing their own achievement with the achievement of others in the same domain (Marsh, 1987; Marsh et al., 2005; see also Sections 1.2.3 and 1.4.2). Students also compare their achievement in one domain with their achievement in

other domains, referred to as *dimensional comparisons* (see Möller & Marsh, 2013). Whereas there are positive associations between achievement and self-concept in matching domains, negative associations have usually been found in non-matching domains, such as mathematics and verbal domains (Möller & Marsh, 2013; Möller et al., 2009). Such associations have been widely studied in self-concept research, but similar patterns have been also found for interest and value beliefs (e.g., Guo, Marsh, Parker, Morin, & Dicke, 2017; Jansen, Schroeders, Lüdtke, & Marsh, 2015; Schurtz et al., 2014). Thus, it seems that individuals tend to feel that they either a “math person” or a “verbal person”, although their achievement in both domains is typically positively correlated. Environmental factors that influence achievement, self-concept, and value in one domain could therefore differently influence students’ motivational beliefs in other domains. For example, Gaspard, Dicke, et al. (2016) found that an intervention that improved students’ value beliefs in mathematics negatively influenced their value beliefs in German. Transferring these findings to gender differences in mathematics, one could assume that environmental factors that negatively influence girls’ math motivation might positively influence their German motivation, and thus might influence gender differences in mathematically intensive STEM careers in two ways.

Taking such arguments into consideration, Study 1 investigated the effects of the stereotypes presented in the television program on children’s motivation not only in mathematics, but also in German. Although there were no effects on motivational outcomes in German, the same arguments discussed in the manuscript with respect to the findings for mathematics (e.g., the video contained a mixture of different stereotypes, the participants’ age group) might also apply to students’ motivation in German. Nevertheless, it remains a strength of the present dissertation that the effects of this stereotyping program were not only investigated in one domain (i.e., mathematics), but also in another important non-matching domain (i.e., German). With respect to the coursework requirements investigated in Study 2, however, the results were limited to the domain of mathematics. As the investigated reform also consisted of changes in coursework requirements for German and the first foreign language, it was not possible to investigate effects of the changes in math coursework on students’ self-concept in verbal domains. Thus, it is unclear how a reform requiring all students to take advanced courses in mathematics might affect students’ self-concept in verbal domains.

With respect to the findings in Study 2, it also should be noted that in terms of expectancy-value constructs only the effects of the reform on self-concept were investigated, because students’ value beliefs were not assessed in the first cohort. As an indicator of



---

students' interest, two vocational interest orientations were examined, namely realistic and investigative interests. As described in Section 1.5, however, there are important differences between value beliefs as conceptualized in the expectancy-value theory and vocational interests (see Eccles, 2005; Holland, 1997). Thus, it is unclear how the changes in math coursework requirements influenced gender differences in value beliefs in mathematics or in other domains. Due to this limitation, further research is needed to investigate the effects of encouraging young women to take advanced math courses in high school on gender differences in value beliefs.

## 5.3 Implications and Future Directions

The findings of the three studies conducted as part of this dissertation have implications not only for future research but also for education policy and practice. The specific implications of each study have already been discussed in Chapters 2, 3, and 4 and are therefore not repeated at this point. Rather, some general implications for future research as well as policy and practice are discussed in this section along with emerging open research questions.

### 5.3.1 Implications and directions for future research

The results of all three studies illustrate the complexity of the factors involved in gendered career pathways, as discussed in the previous sections. This section derives some implications and directions for future research from the findings of this dissertation in terms of the following aspects: (1) the need to investigate multiple outcomes when studying environmental influences on gender differences in mathematically intensive STEM careers, and (2) the need to carefully consider the constructs used when investigating gender differences in mathematically intensive STEM careers.

#### *Investigation of environmental influences on multiple outcomes*

The findings of Study 1 and Study 2 suggest that it is crucial to consider multiple aspects when investigating the effects of environmental factors on gender differences in mathematically intensive career pathways. Prior research on gender stereotypes has often focused on effects on math achievement, although more recent studies have also considered the effects of gender-math stereotypes on different motivational outcomes for female and male students (e.g., Plante et al., 2013; Song et al., 2017). Similarly, research on educational reforms has often focused solely on main effects of achievement, although there is increasing awareness that educational reforms can differentially affect specific subgroups of students (e.g., Domina, McEachin, Penner, & Penner, 2015; Gross, Booker, & Goldhaber, 2009). Furthermore, recent studies have investigated how gender moderates associations between math courses and STEM outcomes (Ellis et al., 2016; Ma & Johnson, 2008). Adding to such findings, the results of Study 1 and Study 2 suggest that a promising avenue for future research on how gender differences in mathematically intensive STEM fields might be influenced by environmental factors would be to broaden the focus in two ways: in terms of how environmental factors differentially influence female and male students, and how the effects of environmental factors are specific to the particular outcomes considered. Research

designs that simultaneously examine the effects of gender stereotypes and high school coursework in mathematics on different outcomes among both female and male students might therefore provide a deeper understanding of how gender differences in mathematically intensive STEM fields develop during students' school years.

In contrast to the assumption that specific experiences accumulate over time (Ridgeway & Correll, 2004; Schoon & Eccles, 2014), it is also still an open question how such experiences develop and accumulate over time. For example, children are likely to repeatedly experience stereotypes that link mathematics more closely with males than with females or portray mathematicians as geeky people in television programs. One popular example presenting such stereotypes is the American television sitcom *The Big Bang Theory* airing on CBS, which has been one of the highest rated and most viewed television shows in the US and many other countries, including Germany, since 2007 (Kirsch, 2011; Patten, 2013). How different motivational outcomes among boys and girls might be influenced by observing such stereotypes repeatedly thus needs to be explored in future research. Similarly, more research exploring the role of high school coursework on gender differences in mathematically intensive STEM careers that considers multiple aspects is needed. Here, the specifics of different educational systems should be a primary focus in future research in order to gain greater insight into how the results of Study 2 might be similar in samples from other Western countries.

Moreover, the use of different assessment tools to measure the effects of stereotypes in television programs and coursework requirements could provide additional insights. For example, eye tracking methodologies could provide further insights into the processes underlying girls' and boys' experience of watching a television program with gender stereotypes. Eye tracking is increasingly used in research on multimedia learning to identify which specific sections of a given piece of material individuals focus on (see van Gog & Scheiter, 2010). Drawing upon such findings, the use of eye tracking could provide insights into specific aspects of television programs that include stereotypes, such as what specific portions of the video boys and girls were looking at in Study 1. Do boys focus more on the geeky boy, whereas girls focus more on the girls who do not want to do their homework? Or is it the other way around, with boys focusing on how the girls in the video react to the geeky boy in particular, and girls focusing more on what the geeky boy is doing? Knowing more about these aspects could help us understand how girls and boys experience such stereotypes in television programs. Observational measures could also provide a deeper understanding of the effects of coursework requirements beyond what is possible with self-reports (see

Fredricks & McColskey, 2012). For example, observational measures assessing how young women and men participate in math classes or how teachers interact with them could provide information about the processes that might mediate the differential effects of advanced courses in high school for young women and men.

### *Consideration of constructs used*

The results of all three studies conducted in this dissertation add to prior findings on the variety of constructs involved in gender differences in mathematically intensive STEM career pathways. A broad range of different motivation and interest constructs have been found to be related to gender differences in mathematically intensive STEM careers, such as expectancy-value constructs and vocational interests (e.g., Schoon & Eccles, 2014; Su et al., 2009). Furthermore, different indicators of STEM careers have been investigated in prior research, including math achievement or aspirations and choice of different STEM majors (see Wang & Degol, 2013). The results of Study 3 add to such prior findings by suggesting that the predictive power of specific motivational and interest constructs might depend on the domain-specificity of the variables considered in analyses. Based on these findings, it seems critical to consider the domain-specificity of variables in research on the associations between motivation and interest constructs and gender differences in mathematically intensive STEM careers. However, the results of Study 3 only provide first insights into the relative predictive power of expectancy-value constructs and vocational interests on mathematically intensive STEM career pathways for young women and men. The relative predictive power of expectancy-value constructs and vocational interests for different STEM outcomes at other time points or other timespans is still an open question. Based on the results of Study 3, the development of expectancy-value constructs and vocational interests in relation to each other over students' school careers is another open question that needs to be investigated in further research. More information about the relative predictive power of expectancy-value constructs and vocational interests over time as well as their relative development over students' school careers is needed in order to design adequate interventions to increase gender equity in mathematically intensive STEM careers, for instance.

Furthermore, the results of Study 3 indicate that it is critical to consider different indicators for mathematically intensive STEM careers, also within STEM domains. Research on gender differences in mathematically intensive STEM careers has put a great deal of focus on the choice of (mathematically intensive) STEM majors compared to non-STEM careers (for a review, see Wang & Degol, 2017). As stated by Wang and Degol (2017), however,

there are important differences within different STEM disciplines, meaning that more research examining gender differences within various STEM fields is also needed (for a similar conclusion, see Su & Rounds, 2015). Whereas women are underrepresented in mathematically intensive STEM disciplines, they pursue careers in the field of life sciences at similar or even higher rates than men (DG Research and Innovation, 2016). The results of Study 3 add to prior research findings in suggesting that students' self-concepts, task values, and vocational interests are predictive not only of choosing a STEM major versus a major outside of STEM, but also the choice of mathematically intensive STEM majors versus STEM majors in the area of life sciences. Thus, these findings provide further support for the need to carefully distinguish among sub-disciplines when discussing gender differences in STEM careers. In extending these findings, however, more also needs to be known about specific sub-disciplines at a finer level. For example, how do expectancy-value constructs and vocational interests predict gender differences in different (mathematically intensive) STEM disciplines such as mathematics, engineering, physics, or biology?

In terms of methodology, the use of person-centered approaches might provide further information about the role of expectancy-value constructs and vocational interests for different indicators of STEM careers. Person-centered techniques such as latent class analyses allow for a consideration of how students prioritize expectancy and value beliefs in different domains or vocational interests (Vermunt & Magidson, 2002). Such approaches assign students into different groups depending on the variables considered in the analyses (e.g., expectancy and value beliefs; Vermunt & Magidson, 2002). In a second step, group membership can be linked to specific outcomes of interest (e.g., choice of STEM majors; Chow et al., 2012; Chow & Salmela-Aro, 2011; Vermunt & Magidson, 2002). In accordance with the role of intraindividual hierarchies of expectancy and value beliefs in different domains (Eccles, 2009), the use of person-centered approaches could thus provide more information about how individuals rank expectancy-value constructs and vocational interests, and how these rank orders predict gender differences in mathematically intensive STEM careers.

### **5.3.2 Implications for educational policy and practice**

As previously discussed, future research is needed to replicate and extend the findings of the present dissertation. Nevertheless, some general implications for policy and practice can be cautiously derived on the basis of the findings of the three studies. In particular, the present dissertation provides further support for significant gender differences in math

motivation at different ages, such as at the beginning of secondary school and the end of high school. Building on previous research, the present dissertation indicates that gender differences in mathematically intensive STEM career pathways need to be considered not only at the transition from high school to actually entering professions (e.g., when choosing university majors). Rather, earlier phases when students are still in school should also remain in focus when it comes to promoting gender equity in the form of economic equity (Ceci et al., 2014; Schoon & Eccles, 2014).

In addition, different factors seem to be able to influence gender differences in mathematically intensive STEM career pathways (i.e., stereotyped television programs and high school coursework requirements). However, the findings of Study 1 indicate that the stereotypes presented in the television program had a rather small influence on children's motivation, although girls and boys reported higher stereotype endorsement (i.e., tended to assume that boys are better in mathematics than girls) after watching the video containing stereotypes. However, the fact that there weren't any effects on girls' motivation does not necessarily mean that girls' motivation is not affected by such stereotyping programs in general. Further research is needed to examine such experiences, as previously discussed. In the meantime, however, one could possibly argue that program developers should at least carefully consider stereotypes in television programs for children, as stereotypes might influence girls' and boys' stereotype endorsement and math motivation (as indicated by the effects found for boys).

With respect to high school coursework in mathematics, the results of Study 2 indicate that gender differences in mathematically intensive STEM fields might be influenced by coursework requirements, although gender differences in the choice of STEM majors after high school were similar before and after the reform. Nevertheless, the reform was associated with smaller gender differences in math achievement but larger gender differences in math self-concept, which are important predictors of the later pursuit of mathematically intensive STEM careers. Thus, when altering coursework requirements, the possibility of differential effects on achievement and motivation among female and male students should be considered and ideally balanced in education policy.

In addition, the results of Study 3 provide information that could be useful for conducting interventions targeting gender differences in mathematically intensive STEM career pathways in high school. Interventions targeting students' motivation (e.g., task value beliefs) have found to be highly successful (see Lazowski & Hulleman, 2016; Rosenzweig & Wigfield, 2016). Based on the findings of Study 3, it seems to be useful to carefully think

about the constructs targeted in such interventions. In order to address gender differences in mathematically intensive STEM majors, for instance, it might be more effective to foster students' vocational interests rather than their value beliefs. Nevertheless, the malleability of expectancy-value constructs and vocational interests should also be considered, and here only little is known about the development of vocational interests might be targeted in interventions.

All in all, the findings of the three studies of this dissertation indicate that gender differences in mathematically intensive STEM careers are rooted in complex processes. A variety of motivational factors are involved in such processes, which themselves can be influenced by different environmental factors. The findings of this dissertation provide some further insights into these complex processes with regard to the specific research questions addressed. Various points are relevant for achieving gender equality in mathematically intensive STEM domains. However, based on the findings of the present dissertation, it seems that, as already stated by Schoon and Eccles (2014, p. 23), “to eliminate gender inequality it is not sufficient to address or to eliminate any single factor or process, [...] [rather], an integrative effort is needed”.

## References

- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*(2), 219–245.  
<https://doi.org/10.1037/0033-2909.121.2.219>
- Ackerman, P. L., Kanfer, R., & Beier, M. E. (2013). Trait complex, cognitive ability, and domain knowledge predictors of baccalaureate success, STEM persistence, and gender differences. *Journal of Educational Psychology*, *105*(3), 911–927.  
<https://doi.org/10.1037/a0032338>
- Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science*, *12*(5), 385–390. <https://doi.org/10.1111/1467-9280.00371>
- Anderson, M. Z., Tracey, T. J. G., & Rounds, J. (1997). Examining the invariance of Holland's vocational interest model across gender. *Journal of Vocational Behavior*, *50*(3), 349–364. <https://doi.org/10.1006/jvbe.1996.1550>
- Archambault, I., Eccles, J. S., & Vida, M. N. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from grades 1 through 12. *Journal of Educational Psychology*, *102*(4), 804–816. <https://doi.org/10.1037/a0021075>
- Armstrong, P. I., & Vogel, D. L. (2010). Theoretical and methodological issues with testing the SCCT and RIASEC models: Comment on Lent, Sheu, and Brown (2010) and Lubinski (2010). *Journal of Counseling Psychology*, *57*(2), 239–247.  
<https://doi.org/10.1037/a0019177>
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*(6), 359–372. <https://doi.org/10.1037/h0043445>
- Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Barron, K. E., & Hulleman, C. S. (2015). Expectancy-value-cost model of motivation. In *International Encyclopedia of the Social and Behavioral Sciences* (pp. 503–509). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.26099-6>
- Battle, E. S. (1965). Motivational determinants of academic task persistence. *Journal of Personality and Social Psychology*, *2*(2), 209–218.  
<https://doi.org/http://dx.doi.org/10.1037/h0022442>
- Battle, E. S. (1966). Motivational determinants of academic competence. *Journal of Personality and Social Psychology*, *4*(6), 634–642. <https://doi.org/10.1037/h0022442>
- Benbow, C. P., Lubinski, D., Shea, D. L., & Eftekhari-sanjani, H. (2000). Sex differences in



- mathematical reasoning ability at age 13: Their status 20 years later. *Psychological Science*, *11*(6), 474–480. <https://doi.org/10.1111/1467-9280.00291>
- Bleeker, M. M., & Jacobs, J. E. (2004). Achievement in math and science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology*, *96*(1), 97–109. <https://doi.org/10.1037/0022-0663.96.1.97>
- Bong, M. (2001a). Between-and within-domain relations of academic motivation among middle and high school students: self-efficacy, task-value, and achievement goals. *Journal of Educational Psychology*, *93*(1), 23–34. <https://doi.org/10.1037//0022-0663.93.1.23>
- Bong, M. (2001b). Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions. *Contemporary Educational Psychology*, *26*(4), 553–570. <https://doi.org/10.1006/ceps.2000.1048>
- Bong, M., Cho, C., Ahn, H. S., & Kim, H. J. (2012). Comparison of self-beliefs for predicting student motivation and achievement. *The Journal of Educational Research*, *105*(5), 336–352. <https://doi.org/10.1080/00220671.2011.627401>
- 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>
- Bonnot, V., & Croizet, J.-C. (2007). Stereotype internalization and women's math performance: The role of interference in working memory. *Journal of Experimental Social Psychology*, *43*(6), 857–866. <https://doi.org/10.1016/j.jesp.2006.10.006>
- Brunello, G., & Checchi, D. (2007). Does school tracking affect equality of opportunity? New international evidence. *Economic Policy*, *22*(52), 782–861. <https://doi.org/10.1111/j.1468-0327.2007.00189.x>
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, *62*(3), 193–217. <https://doi.org/10.1037/h0047470>
- Budde, J. (2009). *Mathematikunterricht und Geschlecht - Empirische Ergebnisse und pädagogische Ansätze [math class and gender - empirical findings and pedagogical approaches]*. BMBF Bildungsforschung (Vol. 30). Retrieved from <http://www.bmbf.de>
- Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106*(4), 676–713. <https://doi.org/10.1037/0033-295x.106.4.676>
- Byun, S., Irvin, M. J., & Bell, B. A. (2015). Advanced math course taking: Effects on math achievement and college enrollment. *Journal of Experimental Education*, *83*(4), 439–

468. <https://doi.org/10.1080/00220973.2014.919570>
- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectance on performance. *European Journal of Social Psychology, 33*, 267–285. <https://doi.org/10.1002/ejsp.145>
- Cambria, J., Brandt, H., Nagengast, B., & Trautwein, U. (2017). Frame of reference effects on values in mathematics: Evidence from German secondary school students. *ZDM, 49*(3), 435–447. <https://doi.org/10.1007/s11858-017-0841-0>
- Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest, 15*(3), 75–141. <https://doi.org/10.1177/1529100614541236>
- Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions in Psychological Science, 19*(5), 275–279. <https://doi.org/10.1177/0963721410383241>
- Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*(2), 218–261. <https://doi.org/10.1037/a0014412>
- Cheryan, S., Plaut, V. C., Handron, C., & Hudson, L. (2013). The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women. *Sex Roles, 69*(1–2), 58–71. <https://doi.org/10.1007/s11199-013-0296-x>
- Cheryan, S., Ziegler, S. A., Montoya, A., & Jiang, L. (2016). Why are some STEM field more gender balanced than others? *Psychological Bulletin, 143*(1), 1–35. <https://doi.org/10.1037/bul0000052>
- Chmielewski, a. K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal, 50*(5), 925–957. <https://doi.org/10.3102/0002831213489843>
- Chow, A., Eccles, J. S., & Salmela-Aro, K. (2012). Task value profiles across subjects and aspirations to physical and IT-related sciences in the United States and Finland. *Developmental Psychology, 48*(6), 1612–1628. <https://doi.org/10.1037/a0030194>
- Chow, A., & Salmela-Aro, K. (2011). Task-values across subject domains: A gender comparison using a person-centered approach. *International Journal of Behavioral Development, 35*(3), 202–209. <https://doi.org/10.1177/0165025411398184>
- Chow, A., & Salmela-Aro, K. (2014). Does priority matter? Gendered patterns of subjective task values across school subject domains. In I. Schoon & J. S. Eccles (Eds.), *Gender*

- differences in aspirations and attainment* (pp. 247–265). Cambridge, U.K.: Cambridge University Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences. Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc. <https://doi.org/10.1234/12345678>
- Cole, J. S., Bergin, D. A., & Whittaker, T. A. (2008). Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology, 33*(4), 609–624. <https://doi.org/10.1016/j.cedpsych.2007.10.002>
- Collins, R. L. (2011). Content analysis of gender roles in media: Where are we now and where should we go? *Sex Roles, 64*(3), 290–298. <https://doi.org/10.1007/s11199-010-9929-5>
- Conley, A. M. (2012). Patterns of motivation beliefs: Combining achievement goal and expectancy-value perspectives. *Journal of Educational Psychology, 104*(1), 32–47. <https://doi.org/10.1037/a0026042>
- Cundiff, J. L., Vescio, T. K., Loken, E., & Lo, L. (2013). Do gender-science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Social Psychology of Education, 16*, 541–554. <https://doi.org/10.1007/s11218-013-9232-8>
- Cunningham, B. C., & Hoyer, K. M. (2015). *Gender differences in science, technology, engineering, and mathematics (STEM) interest, credits earned, and NAEP performance in the 12th grade. (NCES 2015-075)*. Washington, DC. Retrieved from <https://nces.ed.gov/pubs2015/2015075.pdf>
- Cvencek, D., Kapur, M., & Meltzoff, A. N. (2015). Math achievement, stereotypes, and math self-concepts among elementary-school students in Singapore. *Learning and Instruction, 39*, 1–10. <https://doi.org/10.1016/j.learninstruc.2015.04.002>
- 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>
- Daniels, Z. (2008). *Entwicklung schulischer Interessen im Jugendalter [Development of educational interest in adolescence]*. Münster, New York, München, Berlin: Waxmann.
- Darcy, M. U. A., & Tracey, T. J. G. (2007). Circumplex structure of Holland’s RIASEC interests across gender and time. *Journal of Counseling Psychology, 54*(1), 17–31. <https://doi.org/10.1037/0022-0167.54.1.17>
- Davies, P. G., Spencer, S. J., Quinn, D. M., & Gerhardstein, R. (2002). Consuming images:

- How television commercials that elicit stereotype threat can restrain women academically and professionally. *Personality and Social Psychology Bulletin*, 28(12), 1615–1628. <https://doi.org/10.1177/014616702237644>
- Deaux, K. (1985). Sex and gender. *Annual Review of Psychology*, 36, 49–81. <https://doi.org/10.1146/annurev.biochem.64.1.721>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, Boston, Dordrecht, London, Moscow: Kluwer Academic/Plenum Publishers.
- Deci, E. L., & Ryan, R. M. (Eds.). (2004). *Handbook of self-determination research*. Rochester, NY: The University of Rochester Press.
- Denissen, J. J. A., Zarrett, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development*, 78(2), 430–447. <https://doi.org/10.1111/j.1467-8624.2007.01007.x>
- DG Research and Innovation. (2016). *She figures 2015*. Luxemburg.
- Dickhäuser, O., & Stiensmeier-Pelster, J. (2002). Gender differences in computer work: Evidence for the model of achievement-related choices. *Contemporary Educational Psychology*, 27(3), 486–496. <https://doi.org/10.1006/ceps.2001.1106>
- Diekmann, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, 21(8), 1051–1057. <https://doi.org/10.1177/0956797610377342>
- Diekmann, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., & Steinberg, M. (2011). Malleability in communal goals and beliefs influences attraction to stem careers: Evidence for a goal congruity perspective. *Journal of Personality and Social Psychology*, 101(5), 902–918. <https://doi.org/10.1037/a0025199>
- Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). Aiming high and falling short: California's eighth-grade algebra-for-all effort. *Educational Evaluation and Policy Analysis*, 37(3), 275–295. <https://doi.org/10.3102/0162373714543685>
- Doyle, R. A., & Voyer, D. (2016). Stereotype manipulation effects on math and spatial test performance: A meta-analysis. *Learning and Individual Differences*, 47, 103–116. <https://doi.org/10.1016/j.lindif.2015.12.018>
- Dweck, C. S., & Elliott, E. S. (1983). Achievement motivation. In P. H. Mussen (Ed.), *Handbook of child psychology* (pp. 643–691). New York: Wiley.

- Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin*, *15*(4), 543–558.  
<https://doi.org/0803973233>
- Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly*, *11*(2), 135–172. <https://doi.org/10.1111/j.1471-6402.1987.tb00781.x>
- 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. (2005). Subjective task values and the Eccles et al. model of achievement related choices. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105–121). New York, NY: Guilford.
- Eccles, J. S. (2007a). Families, schools, and developing achievement-related motivations and engagement. In J. E. Grusec & P. D. Hastings (Eds.), *Handbook of socialization. Theory and research* (pp. 665-691). New York, London: The Guilford Press.
- Eccles, J. S. (2007b). Where are all the women? Gender differences in participation in physical science and engineering. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more women in science?* (pp. 199-210) Washington, DC: American Psychological Association.
- Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, *44*(2), 78–89.  
<https://doi.org/10.1080/00461520902832368>
- Eccles, J. S. (2011). Understanding women's achievement choices: Looking back and looking forward. *Psychology of Women Quarterly*, *35*(3), 510–516.  
<https://doi.org/10.1177/0361684311414829>
- Eccles, J. S. (2014). Gendered socialization of STEM interests in the family. *International Journal of Gender, Science and Technology*, *7*(2), 116–132. Retrieved from <http://genderandset.open.ac.uk/index.php/genderandset/article/view/419>
- Eccles, J. S., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–121). San Francisco, CA: W. H. Freeman & Co.
- Eccles, J. S., & Harold, R. D. (1991). Gender differences in sport involvement: Applying the eccles' expectancy-value model. *Journal of Applied Sport Psychology*, *3*(1), 7–35.  
<https://doi.org/10.1080/10413209108406432>

- Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. In R. Ames & C. Ames (Eds.), *Research on motivation in education: Vol. 3. Goals and cognitions* (pp. 139–181). New York: Academic Press. [https://doi.org/10.1163/\\_q3\\_SIM\\_00374](https://doi.org/10.1163/_q3_SIM_00374)
- Eccles, J. S., 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. *American Psychologist*, *48*(2), 90–101. <https://doi.org/10.1037/0003-066X.48.2.90>
- Eccles, J. S., & Roeser, R. W. (2009). Schools, academic motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (3rd ed., pp. 404–434). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470479193.adlpsy001013>
- 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. (1995). In the mind of an actor: The structure of adolescents' academic achievement related-beliefs and self-perceptions. *Personality and Social Psychology Bulletin*, *21*, 215–225. <https://doi.org/https://doi.org/10.1177/0146167295213003>
- 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. S., Wigfield, A., Harold, R. D., & Blumenfeld, P. C. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*(3), 830–847. <https://doi.org/10.1111/j.1467-8624.1993.tb02946.x>
- Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon & N. Eisenberg (Eds.), *Handbook of child development* (5th ed., pp. 1017–1095). New York: Wiley.
- Ellis, J., Fosdick, B. K., & Rasmussen, C. (2016). Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. *PLoS ONE*, *11*(7), 1–14. <https://doi.org/10.1371/journal.pone.0157447>
- 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>
- Else-Quest, N. M., Mineo, C. C., & Higgins, A. (2013). Math and science attitudes and

- achievement at the intersection of gender and ethnicity. *Psychology of Women Quarterly*, 37(3), 293–309. <https://doi.org/10.1177/0361684313480694>
- European Commission. (2013). *European research area. Progress report 2013*. <https://doi.org/10.2777/3450>
- Eurostat. (2017). *Tertiary education statistics*. Retrieved from [http://ec.europa.eu/eurostat/statistics-explained/index.php/Tertiary\\_education\\_statistics#Participation\\_by\\_level](http://ec.europa.eu/eurostat/statistics-explained/index.php/Tertiary_education_statistics#Participation_by_level)
- Feather, N. T. (1982). *Expectations and actions: Expectancy-value models in psychology*. Hillsdale, NJ, England: Erlbaum.
- Figueredo, A. J., Gladden, P. R., Sisco, M. M., Patch, E. A., & Jones, D. N. (2015). The unholy trinity: The Dark Triad, coercion, and Brunswik-Symmetry. *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behavior*, 13(2), 435–54. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/26054294>
- Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology*, 53(1), 25–44. <https://doi.org/10.1016/j.jsp.2014.10.002>
- Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology*, 38(4), 519–533. <https://doi.org/10.1037//0012-1649.38.4.519>
- Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 763–782). Boston, MA: Springer US. [https://doi.org/10.1007/978-1-4614-2018-7\\_37](https://doi.org/10.1007/978-1-4614-2018-7_37)
- 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>
- Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics —A “hopeless” issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, 22(April 2005), 497–514. <https://doi.org/10.1007/BF03173468>

- Gamoran, A., & Mare, R. D. (1989). Secondary school tracking and educational inequality: Compensation, reinforcement, or neutrality? *American Journal of Sociology*, *94*(5), 1146-1183. <https://doi.org/10.1086/229114>
- Gao, Z. (2007). *Understanding students' motivation in physical education: Integration of expectancy-value model and self-efficacy theory [Dissertation]*. Louisiana State University and Agricultural and Medical College.
- Gaspard, H., Dicke, A.-L., Flunger, B., Häfner, I., Brisson, B., Trautwein, U., & Nagengast, B. (2016). Side effects of motivational interventions? Effects of an intervention in math classrooms on motivation in verbal domains. *AERA Open*, *2*(2), 1–14. <https://doi.org/10.1017/CBO9781107415324.004>
- Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology*, *107*(3), 663–677. <https://doi.org/10.1037/edu0000003>
- Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., & Nagengast, B. (2016). Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology. Advance Online Publication*. <https://doi.org/10.1016/j.cedpsych.2016.09.003>
- Good, J. J., Woodzicka, J. a, & Wingfield, L. C. (2010). The effects of gender stereotypic and counter-stereotypic textbook images on science performance. *The Journal of Social Psychology*, *150*(2), 132–47. <https://doi.org/10.1080/00224540903366552>
- Goodenow, C. (1993). Classroom belonging among early adolescent students: relationships to motivation and achievement. *The Journal of Early Adolescence*, *13*(1), 21-43. <https://doi.org/10.1177/0272431693013001002>
- Graham, S., & Williams, C. (2009). An attributional approach to motivation in school. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 11–33). New York, NY, US: Routledge.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*(1), 3–25. <https://doi.org/10.1037/0033-295X.109.1.3>
- Gross, B., Booker, T. K., & Goldhaber, D. (2009). Boosting student achievement: The effect of comprehensive school reform on student achievement. *Educational Evaluation and*



- Policy Analysis*, 31(2), 111–126. <https://doi.org/10.3102/0162373709333886>
- Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, 95(1), 124–136. <https://doi.org/10.1037/0022-0663.95.1.124>
- 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(3–4), 153–166. <https://doi.org/10.1007/s11199-011-9996-2>
- Guo, J., Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations, and attainment: A longitudinal study. *American Educational Research Journal*, 52(2), 371–402. <https://doi.org/10.3102/0002831214565786>
- Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J. S., & Dicke, T. (2017). Extending expectancy-value theory predictions of achievement and aspirations in science: Dimensional comparison processes and expectancy-by-value interactions. *Learning and Instruction*, 49, 81–91. <https://doi.org/10.1016/j.learninstruc.2016.12.007>
- Guo, J., Nagengast, B., Marsh, H. W., Kelava, A., Gaspard, H., Brandt, H., ... Trautwein, U. (2016). Probing the unique contributions of self-concept, task values, and their interactions using multiple value facets and multiple academic outcomes. *AERA Open*, 2(1), 233285841562688. <https://doi.org/10.1177/2332858415626884>
- 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(8), 1163–1176. <https://doi.org/10.1037/a0039440>
- Hackett, G., Betz, N. E., Casas, J. M., & Rocha-Singh, I. A. (1992). Gender, ethnicity, and social cognitive factors predicting the academic achievement of students in engineering. *Journal of Counseling Psychology*, 39(4), 527–538. <https://doi.org/10.1037/0022-0167.39.4.527>
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbache, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, 8(1), 1–51. <https://doi.org/10.1111/j.1529-1006.2007.00032.x>
- Hanushek, E. A., & Wössmann, L. (2006). Does educational tracking affect performance and inequality? Differences- in-differences evidence across countries. *The Economic Journal*, 116(510), C63–C76. <https://doi.org/10.1111/j.1468-0297.2006.01076.x>

- Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94*(3), 562–575. <https://doi.org/10.1037/0022-0663.94.3.562>
- Heyman, K. (2008). Talk nerdy to me. *Science, 320*(5877), 740–741. <https://doi.org/10.1126/science.320.5877.740>
- Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review, 1*(2), 69–82. <https://doi.org/10.1016/j.edurev.2006.09.001>
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172–177. <https://doi.org/10.1111/j.1750-8606.2008.00061.x>
- Holland, J. L. (1966). *The psychology of vocational choice: A theory of personality types and model environments*. Waltham, MA: Blaisdell.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, Fla.: Psychological Assessment Resources.
- 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>
- Hulleman, C. S., Durik, A. M., Schweigert, S. B., & Harackiewicz, J. M. (2008). Task values, achievement goals, and interest: An integrative analysis. *Journal of Educational Psychology, 100*(2), 398–416. <https://doi.org/10.1037/0022-0663.100.2.398>
- Humphreys, L. G., & Yao, G. (2002). Prediction of graduate major from cognitive and self-report test scores obtained during the high school years. *Psychological Reports, 90*(1), 3–30. <https://doi.org/10.2466/PR0.90.1.3-30>
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*(6), 581–592. <https://doi.org/10.1037/0003-066X.60.6.581>
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139–155. <https://doi.org/10.1037/0033-2909.107.2.139>
- Jacobs, J. E., & Bleeker, M. M. (2004). Girls' and boys' developing interests in math and science: Do parents matter? *New Directions for Child and Adolescent Development, 2004*(106), 5–21. <https://doi.org/10.1002/cd.113>
- Jacobs, J. E., Davis-Kean, P. E., 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, U.K.: Cambridge University Press. <https://doi.org/doi:10.1017/CBO9780511614446.013>
- Jacobs, J. E., & Eccles, J. S. (2000). Parents, task values, and real-life achievement-related choices. In C. Samsone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 405–439). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-012619070-0/50036-2>
- 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., & Marsh, H. W. (2015). Contrast and assimilation effects of dimensional comparisons in five subjects: An extension of the I/E model. *Journal of Educational Psychology, 107*(4), 1086–1101. <https://doi.org/10.1037/edu0000021>
- Jerrim, J., & Schoon, I. (2014). Do teenagers want to become scientists? A comparison of gender differences in attitudes toward science, career expectations, and academic skills across 29 countries. In I. Schoon & J. S. Eccles (Eds.), *Gender differences in aspirations and attainment* (pp. 203–223). Cambridge, U.K.: Cambridge University Press.
- Jodl, K. M., Michael, A., Malanchuk, O., Eccles, J. S., & Sameroff, A. (2001). Parents' roles in shaping early adolescents' occupational aspirations. *Child Development, 72*(4), 1247–1266. <https://doi.org/10.1111/1467-8624.00345>
- Johnson, H. J., Barnard-Brak, L., Saxon, T. F. ., & Johnson, M. K. . (2012). An experimental study of the effects of stereotype threat and stereotype lift on men and women's performance in mathematics. *Journal of Experimental Education, 80*(2), 137–149. Retrieved from <http://10.0.4.56/00220973.2011.567312>
- Jones, L. R., Wheeler, G., & Centurino, V. A. (2015). TIMSS 2015 science framework. In I. V. S. Mullis & M. O. Martin (Eds.), *TIMSS 2015 assessment framework* (pp. 29–60). Boston: TIMSS & PIRLS International Study Center, Boston College.
- Jones, M. G., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science Education, 84*(2), 180–192. [https://doi.org/10.1002/\(SICI\)1098-237X\(200003\)84:2<180::AID-SCE3>3.0.CO;2-X](https://doi.org/10.1002/(SICI)1098-237X(200003)84:2<180::AID-SCE3>3.0.CO;2-X)
- Kelly, S. (2004). Are teachers tracked? On what basis and with what consequences. *Social Psychology of Education, 7*(1984), 55–72.

<https://doi.org/10.1023/B:SPOE.0000010673.78910.f1>

- Kennedy, J., Lyons, T., & Quinn, F. (2014). The continuing decline of science and mathematics enrolments in Australian high schools. *Teaching Science*, 60(2), 34–46.
- Kessels, U. (2015). Bridging the gap by enhancing the fit: How stereotypes about STEM clash with stereotypes about girls. *International Journal of Gender, Science and Technology*, 7(July), 280–296. Retrieved from <http://genderandset.open.ac.uk/index.php/genderandset/article/viewFile/392/687>
- Kessels, U., Rau, M., & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, 76(4), 761–780. <https://doi.org/10.1348/000709905X59961>
- Kirsch, J. (2011). “The Big Bang Theory” - nächster Glückstreffer für ProSieben? [“The big bang theory” - the next lucky strike for ProSieben?]. Retrieved February 23, 2017, from <http://www.quotenmeter.de/cms/?p1=n&p2=48491>
- Köller, O., Baumert, J., & Schnabel, K. U. (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>
- Köller, O., Trautwein, U., Lüdtke, O., & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe [On the interplay of academic achievement, self-concept, and interest in upper secondary schools]. *Zeitschrift Für Pädagogische Psychologie*, 20(1), 27–39. <https://doi.org/10.1024/1010-0652.20.12.27>
- Köller, O., Watermann, R., Trautwein, U., & Lüdtke, O. (2004). *Wege zur Hochschulreife in Baden-Württemberg: TOSCA - eine Untersuchung an allgemein bildenden und beruflichen Gymnasien [Ways to qualification for university entrance. TOSCA – A study at traditional and vocational Gymnasium schools]*. Opladen: Leske + Budrich.
- 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>
- Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12(4), 383–409. [https://doi.org/10.1016/S0959-4752\(01\)00011-1](https://doi.org/10.1016/S0959-4752(01)00011-1)
- Langdon, D., McKittrick, G., Beede, D., Khan, B., & Doms, M. (2011). *STEM: Good jobs now and for the future. ESA issue brief #03-11*. US Department of Commerce. 1401

- Constitution Avenue NW, Washington, DC 20230. Tel: 202-482-2000; Web site: <http://www.commerce.gov>.
- Lapan, R., Shaughnessy, P., & Boggs, K. R. (1996). Efficacy expectations and vocational interests as mediators between sex and choice of math/Science college majors: A longitudinal study. *Journal of Vocational Behavior*, 49(3), 277–91. <http://dx.doi.org/10.1006/jvbe.1996.00>
- Larson, L. M., Wu, T.-F., Bailey, D. C., Borgen, F. H., & Gasser, C. E. (2010). Male and female college students' college majors: The contribution of basic vocational confidence and interests. *Journal of Career Assessment*, 18(1), 16–33. <https://doi.org/10.1177/1069072709340520>
- Lauermann, F., Chow, A., & Eccles, J. S. (2015). Differential effects of adolescents' expectancy and value beliefs about math and English on math/science-related and human-services-related career plans. *International Journal of Gender, Science and Technology*, 7(2), 205–228. Retrieved from: <http://genderandset.open.ac.uk/index.php/genderandset/article/view/393>
- 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*. <https://doi.org/10.1037/dev0000367>
- 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 *Handbook of child psychology and developmental science* (pp. 1–48). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118963418.childpsy219>
- Lent, R. W., Lopez, A. M., Lopez, F. G., & Sheu, H. Bin. (2008). Social cognitive career theory and the prediction of interests and choice goals in the computing disciplines. *Journal of Vocational Behavior*, 73, 52–62. <https://doi.org/10.1016/j.jvb.2008.01.002>
- Liem, G. a. D., Marsh, H. W., Martin, A. J., McInerney, D. M., & Yeung, A. S. (2013). The big-fish-little-pond effect and a national policy of within-school ability streaming: alternative frames of reference. *American Educational Research Journal*, 50(2), 326–370. <https://doi.org/10.3102/0002831212464511>
- Lips, H. M. (2008). *Sex and gender: An introduction*. Long Grove, Illinois: Waveland Press, Inc.
- Lucas, S. R. (2001). Effectively maintained inequality: Education transitions, track mobility,

- and social background effects. *American Journal of Sociology*, 106(6), 1642–1690.  
<https://doi.org/10.1086/321300>
- Ma, X., & Johnson, W. (2008). Mathematics as the critical filter: Curricular effects on gendered career choices. In H. M. G. Watt & J. S. Eccles (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 55–83). Washington, DC: American Psychological Association.
- Maehr, M. L., & Zusho, A. (2009). Achievement goal theory. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 77–104). New York, NY, US: Routledge.
- Mariani, M. (1999). Replace with a database: O\*NET replaces the dictionary of occupational titles. *Occupational Outlook Quarterly*, Spring, 3–9. Retrieved from <http://www.bls.gov/opub/ooq/1999/Spring/art01.pdf>
- Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, 38(1), 299–337.  
<https://doi.org/10.1146/annurev.psych.38.1.299>
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23(1), 129–149.  
<https://doi.org/10.3102/00028312023001129>
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79(3), 280–295. <https://doi.org/10.1037/0022-0663.79.3.280>
- Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology*, 81(3), 417–430.  
<http://dx.doi.org/10.1037/0022-0663.81.3.417>
- Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *Zeitschrift Für Pädagogische Psychologie*, 19(3), 119–129. <https://doi.org/10.1024/1010-0652.19.3.119>
- Marsh, H. W. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester, UK: British Psychological Society.
- Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F., Nagengast, B., ... Abu-Hilal, M. M. (2014). The internal/external frame of reference model of self-concept and achievement relations: Age-cohort and cross-cultural differences. *American Educational Research Journal*, 52(1), 168–202.  
<https://doi.org/10.3102/0002831214549453>
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective. *Perspectives on Psychological Science*, 1(2), 133–163. <https://doi.org/10.1111/j.1745-6916.2006.00010.x>

- Marsh, H. W., Craven, R. G., & Debus, R. (1998). Structure, stability, and development of youth children's self-concepts: A multicohort-multioccasion study. *Child Development, 69*(4), 1030–1053. <https://doi.org/10.2307/1132361>
- Marsh, H. W., & Hau, K.-T. (2003). Big-fish-little-pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist, 58*(5), 364–376. <https://doi.org/10.1037/0003-066X.58.5.364>
- Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*(1), 59–77. <https://doi.org/10.1348/000709910X503501>
- Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107–123. [https://doi.org/10.1207/s15326985ep2003\\_1](https://doi.org/10.1207/s15326985ep2003_1)
- Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal, 44*(3), 631–669. <https://doi.org/10.3102/0002831207306728>
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>
- Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: Relations to academic self-concept and achievement. *American Educational Research Journal, 34*(4), 691–720. <https://doi.org/10.3102/00028312034004691>
- 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>
- Master, A., Cheryan, S., & Meltzoff, A. N. (2015). Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of Educational Psychology, 108*(3), 424–437. <https://doi.org/10.1037/edu0000061>
- Meece, J. L., Glienke, B. B., & Askew, K. (2009). Gender and motivation. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 411–431). New York, NY, US: Routledge.

- Meece, J. L., & Scantlebury, K. S. (2006). Gender and schooling: Progress and persistent barriers. In J. Worrell & C. Goodheart (Eds.), *Handbook of girls' and women's psychological health* (pp. 283–291). New York: Oxford University Press.
- Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, *82*(1), 60–70. <https://doi.org/10.1037//0022-0663.82.1.60>
- Mercier, E. M., Barron, B., & Connor, K. M. O. (2006). Images of self and others as computer users: The role of gender and experience. *Journal of Computer Assisted Learning*, *22*, 335–348. <https://doi.org/10.1111/j.1365-2729.2006.00182.x>
- Miller, D. I., Eagly, A. H., & Linn, M. C. (2014). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology, Advance on*. <https://doi.org/http://dx.doi.org/10.1037/edu0000005>
- Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, *120*(3), 544–60. <https://doi.org/10.1037/a0032459>
- Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, *79*(3), 1129–1167. <https://doi.org/10.3102/0034654309337522>
- Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat. *Psychological Science*, *18*(10), 879–885. <https://doi.org/10.1111/j.1467-9280.2007.01995.x>
- Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology*, *25*(1), 3–53. <https://doi.org/http://dx.doi.org/10.1006/ceps.1999.1019>
- 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/http://dx.doi.org/10.1080/13803611.2015.1057161>
- Muzzatti, B., & Agnoli, F. (2007). Gender and mathematics: Attitudes and stereotype threat susceptibility in Italian children. *Developmental Psychology*, *43*(3), 747–759. <https://doi.org/10.1037/0012-1649.43.3.747>
- Nagengast, B., Marsh, H. W., Scalas, L. F., Xu, M. K., Hau, K.-T., & Trautwein, U. (2011). Who took the “×” out of expectancy-value theory? A psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological*



- Science*, 22(8), 1058–1066. <https://doi.org/10.1177/0956797611415540>
- Nagy, G. (2006). *Berufliche Interessen, kognitive und fachgebundene Kompetenzen. Ihre Bedeutung für die Studienfachwahl und die Bewährung im Studium [Vocational interests, cognitive and domain-specific competencies. Their role for major choices and success at university]*. Berlin: FU. Retrieved from [http://www.diss.fu-berlin.de/diss/receive/FUDISS\\_thesis\\_000000002714](http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000002714)
- Nagy, G., Garrett, J., Trautwein, U., Cortina, K. S., Baumert, J., & Eccles, J. S. (2008). Gendered high school course selection as a precursor of gendered careers: The mediating role of self-concept and intrinsic value. In J. S. Eccles & H. M. G. Watt (Eds.), *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences* (pp. 115–143). Washington, DC, US: American Psychological Association. <https://doi.org/10.1037/11706-004>
- Nagy, G., Trautwein, U., Baumert, J., Köller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, 12(4), 323–345. <https://doi.org/10.1080/13803610600765687>
- 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 Center for Education Statistics, (Ies). (2017). *The condition of education 2017 (NCES 2017-144)*. Washington, DC.
- National Reserach Council. (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: The National Academic Press. <https://doi.org/https://doi.org/10.17226/12190>
- National Science Board. (2016). *Science & engineering indicators 2016*. <https://doi.org/10.1002/ejoc.201200111>
- National Science Foundation. (2017). *Women, minorities, and persons with disabilities in science and engineering: 2017. Special Report NSF 17-310*. Arlington, VA.
- Neuville, E., & Croizet, J.-C. (2007). Can salience of gender identity impair math performance among 7-8 years old girls? The moderating role of task difficulty. *European Journal of Psychology of Education*, 22(3), 307–316. <https://doi.org/10.1007/BF03173428>
- Nguyen, H.-H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of

- minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, 93(6), 1314–1334. <https://doi.org/10.1037/a0012702>
- Niepel, C., Brunner, M., & Preckel, F. (2014). The longitudinal interplay of students' academic self-concepts and achievements within and across domains: Replicating and extending the reciprocal internal/external frame of reference model. *Journal of Educational Psychology*, 106(4), 1170–1191. <https://doi.org/10.1037/a0036307>
- Noonan, R. (2017). *STEM jobs: 2017 update (ESA Issue Brief #02-17)*. Retrieved from <http://www.esa.gov/reports/stem-jobs-2017-update>
- Nosek, B. a, Greenwald, A. G., & Banaji, M. R. (2007). The implicit association test at age 7: A methodological and conceptual review. *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes*, 265–292. <https://doi.org/10.1016/j.mrfmmm.2009.01.007>
- Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science*, 7(4), 384–403. <https://doi.org/10.1177/1745691612449021>
- Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2017). Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior*, 98, 138–151. <https://doi.org/10.1016/j.jvb.2016.11.002>
- OECD. (2012). *PISA 2012 results: Ready to learn. Students' engagement, drive and self-beliefs*. (Vol. III). Paris: OECD Publishing. <https://doi.org/10.1787/9789264201170-en>
- OECD. (2013). *PISA 2012 results in focus. What 15-years-olds know and what they can do with what they know*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264208070-en>
- OECD. (2015). *The ABC of gender equality in education. Aptitude, behaviour, confidence*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264229945-en>
- OECD. (2016). *PISA 2015 results (Volume I). Excellence and equity in education*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264266490-en>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4), 543–578. <https://doi.org/10.3102/00346543066004543>
- Parker, P. D., 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* (pp. 224–246). Cambridge, U.K.: Cambridge University Press.

- Parker, P. D., Schoon, I., Tsai, Y.-M., Nagy, G., Trautwein, U., & Eccles, J. S. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: a multicontext study. *Developmental Psychology, 48*(6), 1629–42. <https://doi.org/10.1037/a0029167>
- Päßler, K., Beinicke, A., & Hell, B. (2014). Gender-related differential validity and differential prediction in interest inventories. *Journal of Career Assessment, 22*(1), 138–152. <https://doi.org/10.1177/1069072713492934>
- 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>
- Passolunghi, M. C., Rueda Ferreira, T. I., & Tomasetto, C. (2014). Math–gender stereotypes and math-related beliefs in childhood and early adolescence. *Learning and Individual Differences, 34*, 70–76. <https://doi.org/10.1016/j.lindif.2014.05.005>
- Patall, E. a., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: A systematic review of research (1985–2009). *Review of Educational Research, 80*(3), 401–436. <https://doi.org/10.3102/0034654310377086>
- Patrick, L., Care, E., & Ainley, M. (2011). The relationship between vocational interests, self-efficacy, and achievement in the prediction of educational pathways. *Journal of Career Assessment, 19*, 61–74. <https://doi.org/10.1177/1069072710382615>
- Patten, D. (2013). Full 2012–2013 TV season series rankings. Retrieved February 23, 2017, from <http://deadline.com/2013/05/tv-season-series-rankings-2013-full-list-506970/>
- 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 W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology* (pp. 103–122). New York, NY: John Wiley & Sons, Inc. <https://doi.org/10.1002/0471264385.wei0706>
- Pintrich, P. R., Marx, R. W., & Boyle, R. a. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167–199. <https://doi.org/10.3102/00346543063002167>
- 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*(3), 225–235.

- <https://doi.org/10.1016/j.cedpsych.2013.03.004>
- Régner, I., Steele, J. R., Ambady, N., Thinus-Blanc, C., & Huguet, P. (2014). Our future scientists: A review of stereotype threat in girls from early elementary school to middle school. *Revue Internationale de Psychologie Sociale*, 27(3–4), 13–51.
- Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of national assessment of educational progress assessments. *Journal of Educational Psychology*, 107(3), 645–662.  
<https://doi.org/10.1037/edu0000012>
- Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist*, 46(3), 168–184.  
<https://doi.org/10.1080/00461520.2011.587723>
- Renninger, K. A., Nieswandt, M., & Hidi, S. (Eds.). (2015). *Interest in mathematics and science learning*. Washington, DC: American Educational Research Association.  
<https://doi.org/10.3102/978-0-935302-42-4>
- 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>
- Rideout, V. J. (2015). *The common sense census: Media use by tweens and teens*. San Francisco, CA: Common Sense Media. <https://doi.org/10.1017/CBO9781107415324.004>
- Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). *Media in the lives of 8-18 year olds*. Menlo Parc, CA. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender & Society*, 18(4), 510–531.  
<https://doi.org/10.1177/0891243204265269>
- Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Social cognitive constructs are just as stable as the big five between grades 5 and 8. *AERA Open*, 3(3), 1-9. <https://doi.org/10.1177/2332858417717691>
- 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>
- Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2011). Who wants to have a career in science or math? Exploring adolescents’ future aspirations by gender and race/ethnicity. *Science Education*, 95(3), 458–476. <https://doi.org/10.1002/sci.20431>

- Ro, H. K., & Knight, D. B. (2016). Gender differences in learning outcomes from the college experiences of engineering students. *Journal of Engineering Education, 105*(3), 478–507. <https://doi.org/10.1002/jee.20125>
- Rolfhus, E. L., & Ackerman, P. L. (1996). Self-report knowledge: At the crossroads of ability, interest, and personality. *Journal of Educational Psychology, 88*(1), 174–188. <https://doi.org/10.1037/0022-0663.88.1.174>
- Roloff Henoch, J., Klusmann, U., Lüdtke, O., & Trautwein, U. (2015). Who becomes a teacher? Challenging the “negative selection” hypothesis. *Learning and Instruction, 36*, 46–56. <https://doi.org/10.1016/j.learninstruc.2014.11.005>
- 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>
- Rounds, J., & Su, R. (2014). The nature and power of interests. *Current Directions in Psychological Science, 23*(2), 98–103. <https://doi.org/10.1177/0963721414522812>
- Sáinz, M., & Eccles, J. S. (2012). Self-concept of computer and math ability: Gender implications across time and within ICT studies. *Journal of Vocational Behavior, 80*(2), 486–499. <https://doi.org/http://dx.doi.org/10.1016/j.jvb.2011.08.005>
- Scheerens, J., & Hendriks, M. (2014). State of the art of time effectiveness. In J. Scheerens (Ed.), *SpringerBriefs in Education. Effectiveness of Time Investments in Education* (pp. 7–29). Cham, Germany: Springer International Publishing.
- Schoon, I. (2015). Explaining persisting gender inequalities in aspirations and attainment: An integrative developmental approach. *International Journal of Gender, Science and Technology, 7*(2), 151–165. Retrieved from <http://genderandset.open.ac.uk/index.php/genderandset/article/view/401>
- Schoon, I., & Eccles, J. S. (Eds.). (2014). *Gender differences in aspiration and attainment: A life course perspective*. Cambridge, U.K.: Cambridge University Press.
- Schoon, I., Ross, A., & Martin, P. (2007). Science related careers: Aspirations and outcomes in two British cohort studies. *Equal Opportunities International, 26*(2), 129–143. <https://doi.org/10.1108/02610150710732203>
- Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student’s mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction, 34*, 32–41. <https://doi.org/10.1016/j.learninstruc.2014.08.001>
- Sells, L. W. (1980). Mathematics: The invisible filter. *Engineering Education, 70*, 340–341.

- Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review, 13*(2), 147-183. <https://doi.org/10.1006/drev.1993.1007>
- 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), 70-83. <https://doi.org/10.1037/0012-1649.42.1.70>
- Skaalvik, E. M., & Skaalvik, S. (2008). Self-concept and self-efficacy in mathematics: Relation with mathematics motivation and achievement. *New Developments in the Psychology of Motivation., 3*(3), 105-128.
- Smith, J. L., Brown, E. R., Thoman, D. B., & Deemer, E. D. (2015). Losing its expected communal value: How stereotype threat undermines women's identity as research scientists. *Social Psychology of Education, 18*(3), 443-466. <https://doi.org/10.1007/s11218-015-9296-8>
- Smith, J. L., Sansone, C., & White, P. H. (2007). The stereotyped task engagement process: The role of interest and achievement motivation. *Journal of Educational Psychology, 99*(1), 99-114. <https://doi.org/10.1037/0022-0663.99.1.99>
- Snyder, T. D., & Hoffmann, C. M. (2001). *Digest of Education Statistics [NCES Publication No. 2001-034]*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Song, J., Zuo, B., Wen, F., & Yan, L. (2017). Math-gender stereotypes and career intentions: an application of expectancy-value theory. *British Journal of Guidance & Counselling, 45*(3), 328-340. <https://doi.org/10.1080/03069885.2017.1283680>
- Song, J., Zuo, B., & Yan, L. (2016). Effects of gender stereotypes on performance in mathematics: A serial multivariable mediation model. *Social Behavior and Personality: An International Journal, 44*(6), 943-952. <https://doi.org/10.2224/sbp.2016.44.6.943>
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28. <https://doi.org/10.1006/jesp.1998.1373>
- Spinath, B., Spinath, F. M., Harlaar, N., & Plomin, R. (2006). Predicting school achievement from general cognitive ability, self-perceived ability, and intrinsic value. *Intelligence, 34*(4), 363-374. <https://doi.org/10.1016/j.intell.2005.11.004>
- Spinath, B., & Steinmayr, R. (2008). Longitudinal analysis of intrinsic motivation and competence beliefs: Is there a relation over time? *Child Development, 79*(5), 1555-1569. <https://doi.org/10.1111/j.1467-8624.2008.01205.x>

- Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *The American Psychologist*, *52*(6), 613–629. <https://doi.org/10.1037/0003-066X.52.6.613>
- Steele, C. M., & Aronson, J. M. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. <https://doi.org/http://dx.doi.org/10.1037/0022-3514.69.5.797>
- Steffens, M. C., Jelenec, P., & Noack, P. (2010). On the leaky math pipeline: Comparing implicit math-gender stereotypes and math withdrawal in female and male children and adolescents. *The Journal of Educational Psychology*, *102*(4), 947–963. <https://doi.org/10.1037/a0019920>
- Steinke, J. (1997). A portrait of a woman as a scientist: Breaking down barriers created by gender-role stereotypes. *Public Understanding of Science*, *6*(4), 409–428. <https://doi.org/10.1088/0963-6625/6/4/006>
- Steinke, J., Lapinski, M. K., Crocker, N., Zietsman-Thomas, A., Williams, Y., Evergreen, S. H., & Kuchibhotla, S. (2007). Assessing media influences on middle school aged children's perceptions of women in science using the draw-a-scientist test (DAST). *Science Communication*, *29*(1), 35–64. <https://doi.org/10.1177/1075547007306508>
- Steinmayr, R., & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality*, *22*(3), 185–209. <https://doi.org/10.1002/per.676>
- Stipek, D. J., & Gralinski, J. H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology*, *83*(3), 361–371. <https://doi.org/10.1037/0022-0663.83.3.361>
- Stoll, G., & Trautwein, U. (2017). Vocational interests as personality traits. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 401–417). London, San Diego, Cambridge, Oxford: Elsevier.
- Su, R., Golubovich, J., & Robbins, S. B. (2015). Bridging science and practice: toward a standard, evidence-based framework of 21st century skills. *Industrial and Organizational Psychology*, *8*(2), 289–294. <https://doi.org/http://dx.doi.org/10.1017/iop.2015.36>
- 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*(189), 1–20. <https://doi.org/10.3389/fpsyg.2015.00189>
- Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: A meta-

- analysis of sex differences in interests. *Psychological Bulletin*, 135(6), 859–884.  
<https://doi.org/10.1037/a0017364>
- The College Board. (2017). SAT suite of assessments. Retrieved July 22, 2017, from  
<https://collegereadiness.collegeboard.org/sat>
- The Further Mathematics Support Programme. (2016). *Girls' participation in A level mathematics and further mathematics*. Retrieved from  
<http://furthermaths.org.uk/files/FMSP-Girls-in-Maths.pdf>
- 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>
- Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology*, 47(4), 943–949. <https://doi.org/10.1037/a0024047>
- Tracey, T. J. G., & Rounds, J. (1993). Evaluating Holland's and Gati's vocational-interest models: A structural meta-analysis. *Psychological Bulletin*, 113(2), 229–246.  
<https://doi.org/10.1037/0033-2909.113.2.229>
- Trautwein, U., Köller, O., Lüdtke, O., & Baumert, J. (2005). Student tracking and the powerful effects of opt-in courses on self-concept: Reflected-glory effects do exist after all. In H. W. Marsh, R. Craven, & D. M. McInerney (Eds.), *New frontiers for self research* (pp. 307–327). Greenwich, CT: Information Age Press.
- Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, 98(4), 788–806. <https://doi.org/10.1037/0022-0663.98.4.788>
- Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology*, 98(2), 438–456. <https://doi.org/10.1037/0022-0663.98.2.438>
- Trautwein, U., Marsh, H. W., Nagengast, B., Lüdtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Journal of Educational Psychology*, 104(3), 763–777.  
<https://doi.org/10.1037/a0027470>
- Trautwein, U., Neumann, M., Nagy, G., Lüdtke, O., & Maaz, K. (Eds.). (2010). *Schulleistungen von Abiturienten: Die neu geordnete gymnasiale Oberstufe auf dem Prüfstand. [School achievement of upper secondary school students. The rearranged*



- upper secondary school on the trial*]. Wiesbaden: VS Verlag für Sozialwissenschaften.  
<https://doi.org/http://dx.doi.org/10.1007/978-3-531-92037-5>
- Tynkkynen, L., Tolvanen, A., & Salmela-Aro, K. (2012). Trajectories of educational expectations from adolescence to young adulthood in Finland. *Developmental Psychology, 48*(6), 1674–85. <https://doi.org/10.1037/a0027245>
- Tyson, W., Lee, R., Borman, K. M., & Hanson, M. A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework and postsecondary degree attainment. *Journal of Education for Students Placed at Risk, 12*(3), 243–270. <https://doi.org/10.1080/10824660701601266>
- van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction, 20*(2), 95–99.  
<https://doi.org/http://dx.doi.org/10.1016/j.learninstruc.2009.02.009>
- Van Iddekinge, C. H., Roth, P. L., Putka, D. J., & Lanivich, S. E. (2011). Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. *Journal of Applied Psychology, 96*(6), 1167–1194.  
<https://doi.org/10.1037/a0024343>
- Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. Hagenaars & A. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89–106). Cambridge, U.K.: Cambridge University Press.
- Volodina, A., & Nagy, G. (2016). Vocational choices in adolescence: The role of gender, school achievement, self-concepts, and vocational interests. *Journal of Vocational Behavior, 95–96*, 58–73. <https://doi.org/10.1016/j.jvb.2016.07.005>
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174–1204. <https://doi.org/10.1037/a0036620>
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39*(5), 456–467. [https://doi.org/10.1016/S0022-1031\(03\)00019-2](https://doi.org/10.1016/S0022-1031(03)00019-2)
- 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., Chow, A., Degol, J. L., & Eccles, J. S. (2017). Does everyone's motivational beliefs about physical science decline in secondary school? Heterogeneity of adolescents' achievement motivation trajectories in physics and chemistry. *Journal of Youth and Adolescence, 46*(8), 1821–1838. <https://doi.org/10.1007/s10964-016-0620-1>
- 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). <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, 1–9. <https://doi.org/10.3389/fpsyg.2015.00036>
- 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), 770–5. <https://doi.org/10.1177/0956797612458937>
- Wang, M.-T., & Kenny, S. (2014). Gender differences in personal aptitudes and motivational beliefs for achievement in and commitment to math and science fields. In I. Schoon & J. S. Eccles (Eds.), *Gender Differences in Aspirations and Attainment* (pp. 266–284). Cambridge, U.K.: Cambridge University Press.
- 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), 1556–74. <https://doi.org/10.1111/j.1467-8624.2004.00757.x>
- Watt, H. M. G. (2005). Explaining gendered math enrollments for NSW Australian secondary school students. *New Directions for Child and Adolescent Development*, 2005(110), 15–29. <https://doi.org/10.1002/cd.147>
- Watt, H. M. G., & Eccles, J. S. (Eds.). (2008). *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences*. Washington, DC: American Psychological Association. <https://doi.org/10.1037/11706-000>
- Watt, H. M. G., Hyde, J. S., Petersen, J., Morris, Z. A., Rozek, C. S., & Harackiewicz, J. M. (2016). Mathematics a critical filter for STEM-related career choices? A longitudinal examination among Australian and U.S. adolescents. *Sex Roles*, 1–18. <https://doi.org/10.1007/s11199-016-0711-1>
- Watt, H. M. G., Richardson, P. W., Klusmann, U., Kunter, M., Beyer, B., Trautwein, U., & Baumert, J. (2012). Motivations for choosing teaching as a career: An international

- comparison using the FIT-Choice scale. *Teaching and Teacher Education*, 28(6), 791–805. <https://doi.org/10.1016/j.tate.2012.03.003>
- 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>
- 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-de Lisi & R. De Lisi (Eds.), *Biology, sociology, and behavior: The development of sex differences in cognition* (pp. 93–124). Greenwich, CT: Ablex.
- Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, 30(1), 1–35. <https://doi.org/10.1016/j.dr.2009.12.001>
- 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. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 91–120). San Diego, San Francisco, New York, Boston, London, Sydney, Tokyo: Academic Press.
- Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S. D., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. In R. M. Lerner, C. Garcia Coll, & M. Lamb (Eds.), *Handbook of child psychology and developmental science* (pp. 1–44). Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118963418.childpsy316>
- Wigfield, A., Eccles, J. S., Iver, D. Mac, Reuman, D. A., & Midgley, C. (1991). Transitions during early adolescence: Changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology*, 27(4), 552–565. <https://doi.org/http://dx.doi.org/10.1037/0012-1649.27.4.552>
- Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. E. (2006). Development of achievement motivation. In N. Eisenberg (Ed.), *Handbook of child*

- psychology, Vol 3: Social, emotional and personality development* (Vol. 3, pp. 933–1012). Hoboken, NJ: Wiley. <https://doi.org/10.1016/B978-012750053-9/50008-5>
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbretton, A. J. a., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*(3), 451–469. <https://doi.org/10.1037/0022-0663.89.3.451>
- Wigfield, A., Tonks, S., & Klauda, S. T. (2009). Expectancy-value theory. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 55–57). New York, NY: Routledge.
- Wittmann, W. W., & Hatstrup, K. (2004). The relationship between performance in dynamic systems and intelligence. *Systems Research and Behavioral Science, 21*(4), 393–409. <https://doi.org/10.1002/sres.653>
- Wittmann, W. W., & Süß, H.-M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problemsolving performances via Brunswik symmetry. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 77–104). Washington, DC: American Psychological Association. <https://doi.org/10.1037/10315-004>
- Yazilintas, D., Svensson, J., & Vries, G. De. (2013). Gendered study choice: a literature review. A review of theory and research into the unequal representation of male and female students in mathematics, science, and technology. *Educational Research and Evaluation, 19*(6), 525–545. <https://doi.org/http://dx.doi.org/10.1080/13803611.2013.803931>
- Zech, A., Bühner, M., Kröner, S., Heene, M., & Hilbert, S. (2017). The impact of symmetry: Explaining contradictory results concerning working memory, reasoning, and complex problem solving. *Journal of Intelligence, 5*(2), 22. <https://doi.org/10.3390/jintelligence5020022>
- Zhang, T., Solmon, M. A., & Gu, X. (2012). The role of teachers' support in predicting students' motivation and achievement outcomes in physical education. *Journal of Teaching in Physical Education, 31*(4), 329–343. <https://doi.org/10.1123/jtpe.31.4.329>