education sciences

Special Issue Reprint

Sticking with STEM

Who Comes, Who Stays, Who Goes, and Why?

Edited by
Erin Mackenzie, Bernhard Ertl, Christine R. Starr,
Silke Luttenberger and Manuela Paechter

mdpi.com/journal/education

MDPI

Sticking with STEM: Who Comes, Who Stays, Who Goes, and Why?

Sticking with STEM: Who Comes, Who Stays, Who Goes, and Why?

Editors

Erin Mackenzie
Bernhard Ertl
Christine R. Starr
Silke Luttenberger
Manuela Paechter

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Editors
Erin Mackenzie
Western Sydney University
Penrith, NSW
Australia

Bernhard Ertl
Universität der Bundeswehr
München
Neubiberg
Germany

Christine R. Starr
University of
Wisconsin—Madison
Madison, WI
USA

Silke Luttenberger
University College of Teacher
Education Styria
Graz
Austria

Manuela Paechter
University of Graz
Graz
Austria

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Education Sciences* (ISSN 2227-7102) (available at: https://www.mdpi.com/journal/education/special_issues/STEM_4W).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-2447-2 (Hbk)
ISBN 978-3-7258-2448-9 (PDF)
doi.org/10.3390/books978-3-7258-2448-9

Contents

About the Editors

Erin Mackenzie

Dr Erin Mackenzie is a Senior Lecturer in Educational Psychology and STEM in the School of Education at Western Sydney University. Her research focuses on gender imbalances in STEM education, with emphasis on the role of psychological and social factors in predicting adolescent girls' participation in STEM. Her research also addresses adolescent wellbeing in school contexts, including adolescent online behaviors, coping, mental health, and peer interactions. Erin currently lectures in educational psychology, educational research methods, and STEM education, working with both pre- and in-service teachers. She is a co-author of *Educational Psychology for Learning and Teaching* (7th Edition).

Bernhard Ertl

Bernhard Ertl is Professor for Learning and Teaching with Media at Universität der Bundeswehr München. He earned his Ph.D. at Ludwig-Maximilian University in 2003. He works in several research projects to design and evaluate methods of instructional design with a particular focus on the individual's aptitudes and learning prerequisites. In this context, he has developed a specific research focus on cognitive tools that support learners' collaborative knowledge construction in internet-enriched learning environments, as well as on tools that provide instructional means for facilitating gender equality in STEM fields. This interventional focus is complemented by a meta-analytic perspective, as exemplified, for example, by the DFG-funded project "The impact of interest congruence on females' and males' study progress in STEM subjects". He is member of the European Association for Research on Learning and Instruction (EARLI), the American Educational Research Association (AERA), the International Society of the Learning Sciences (ISLS), and several national professional societies.

Christine R. Starr

Christy Starr is an Assistant Professor in the Human Development Area at the University of Wisconsin, Madison, having earned her Ph.D. at the University of California, Santa Cruz, in 2019. Her research delves into how adolescents and young adults make career decisions, choose courses, and declare majors. She investigates the disparities in gender, race/ethnicity, and other demographic categories across different majors and careers, focusing on understanding why some fields exhibit high equity while others remain unbalanced.

Her work is particularly concentrated on STEM fields (science, technology, engineering, and math), where she examines the impact of barriers such as stereotypes and the role of support systems like parents. Christy's research aims to uncover how these factors influence young people's confidence, interest, and persistence in STEM. Through her research, she seeks to develop strategies to improve equity, benefiting practitioners, parents, teachers, and society as a whole.

Silke Luttenberger

Silke Luttenberger is Professor of Educational Psychology at the University College of Teacher Education Styria, Austria. Her research primarily centers on teaching and learning processes, focusing on performance enhancement and the motivational-affective aspects of learning. She is particularly interested in the interplay between personal characteristics (e.g., attitudes, self-efficacy, interest) and environmental factors (both school-based and extracurricular learning experiences). The empowerment and resilience of learners are at the focus of her work, aiming to prepare them for a constantly changing world and its associated challenges. Her research features cross-disciplinary, competence-based, and motivationally supportive teaching and learning contexts, with a strong emphasis on STEM. She also addresses gender stereotypes and gender-sensitive educational and career decisions. She is member of the European Association for Research on Learning and Instruction (EARLI) and several national professional societies.

Manuela Paechter

Manuela Paechter is a Professor of Educational Psychology at the University of Graz, Austria. In her research, she investigates learning processes from both individual and instructional perspectives. She adopts a holistic approach that considers personality factors, attitudes, motivation, and self-assessments in interaction with the environment and educators. Her research also concerns the development and evaluation of education settings and learning environments. She focuses on creating learning environments involving digital media, communication tools, and AI that enhance learning, incite motivation, and facilitate effective communication among participants. Gender is a central theme in her research. She aims to mitigate gender disparities and provide insights into how gender affects educational and professional development. Manuela currently lectures various topics in educational psychology for psychology students as well as pre-service teachers.

Preface

This reprint has its origin in the biannual "Gender and STEM" conference held at the Universität der Bundeswehr München in Neubiberg (Germany), particularly that held in July 2022. This conference provided a pivotal opportunity for researchers and practitioners to share insights and develop the ideas that laid the foundations for the Special Issue "Sticking with STEM: Who Comes, Who Stays, Who Goes, and Why?" The discussions and collaborations at this conference underscored the ongoing need to address gender disparities in STEM through interdisciplinary approaches.

With ten contributions, this reprint encompasses a broad array of topics, from early interest in STEM subjects to career persistence, and from the impact of cultural wealth and identity to the role of self-concept and stereotypes. By integrating research and practical strategies, it illuminates the pathways that support diverse and inclusive participation in STEM fields.

Our motivation for compiling this reprint stems from a recognition of the persistent barriers faced by under-represented groups in STEM. By providing a comprehensive understanding of these challenges and showcasing effective interventions, we hope to inspire educators, policymakers, and researchers to foster environments that support all students.

We extend our gratitude to the contributing authors for their exceptional work and to our readers for their dedication to this important issue.

Erin Mackenzie, Bernhard Ertl, Christine R. Starr, Silke Luttenberger, and Manuela Paechter
Editors

education sciences

Editorial

From Early Interest to Career Persistence: Understanding and Supporting STEM Pathways

Manuela Paechter [1,*], Silke Luttenberger [2], Christine R. Starr [3], Erin Mackenzie [4] and Bernhard Ertl [5]

1 Institute of Psychology, University of Graz, 8010 Graz, Austria
2 Practical Education and Action Research, University College of Teacher Education Styria, 8010 Graz, Austria; silke.luttenberger@phst.at
3 Department of Educational Psychology, University of Wisconsin—Madison, Madison, WI 53706, USA; starr3@wisc.edu
4 Centre for Educational Research, Western Sydney University, Penrith, NSW 2751, Australia; e.mackenzie@westernsydney.edu.au
5 Learning and Teaching with Media, Institute of Education, Department of Human Sciences, Universität der Bundeswehr München, 85579 Neubiberg, Germany; bernhard.ertl@unibw.de
* Correspondence: manuela.paechter@uni-graz.at

Citation: Paechter, M.; Luttenberger, S.; Starr, C.R.; Mackenzie, E.; Ertl, B. From Early Interest to Career Persistence: Understanding and Supporting STEM Pathways. *Educ. Sci.* **2024**, *14*, 767. https://doi.org/10.3390/educsci14070767

Received: 12 June 2024
Accepted: 11 July 2024
Published: 14 July 2024

Knowledge, competencies, and reflective attitudes regarding STEM (science, technology, engineering, and mathematics) are essential for navigating global and societal changes in the economy and workforce [1]. As we advance further into the 21st century, knowledge and skills in STEM become increasingly important. New professional fields are emerging that require a blend of mathematics, computer science, natural sciences, engineering, and technology. The value of STEM knowledge and skills goes beyond academic settings and personal needs, significantly impacting society overall [1].

Despite this importance, a long line of research highlights numerous obstacles hindering the development of students' STEM competencies [2,3]. From elementary through high school, STEM subjects often rank low in popularity and are not commonly favored for academic pursuits or career paths, especially for girls in the Global North [4]. A deficit in STEM proficiency undermines innovation, economic growth, and competitiveness, hampering the development of critical thinking, problem-solving abilities, and digital literacy, which are essential for addressing complex global challenges and driving progress in a rapidly evolving world. However, it is not only knowledge and skills that contribute to learning and applying STEM. Personal attitudes like interest, values, and positive self-assessments are also crucial for engagement with STEM.

Against this background, the overarching question of this Special Issue emerged: "Sticking with STEM: Who Comes, Who Stays, Who Goes, and Why?" Answering this question should help to better understand the "leaky pipeline" in STEM [5], which involves the loss of interested and skilled students from STEM starting from early childhood and continuing through adolescence and into adulthood. This leaky pipeline manifests in academic choices against STEM in school, career choices outside of STEM in adolescence and adulthood, or a lack of persistence in staying on a STEM career path. The nine articles in this Special Issue address this question, aiming to provide solutions to attract young people to STEM and ensure they stay on the STEM pathway.

When young people choose a STEM subject at school or embark on an education and/or career pathway in STEM, they have already encountered various experiences, support, challenges, and barriers, as illustrated in Figure 1.

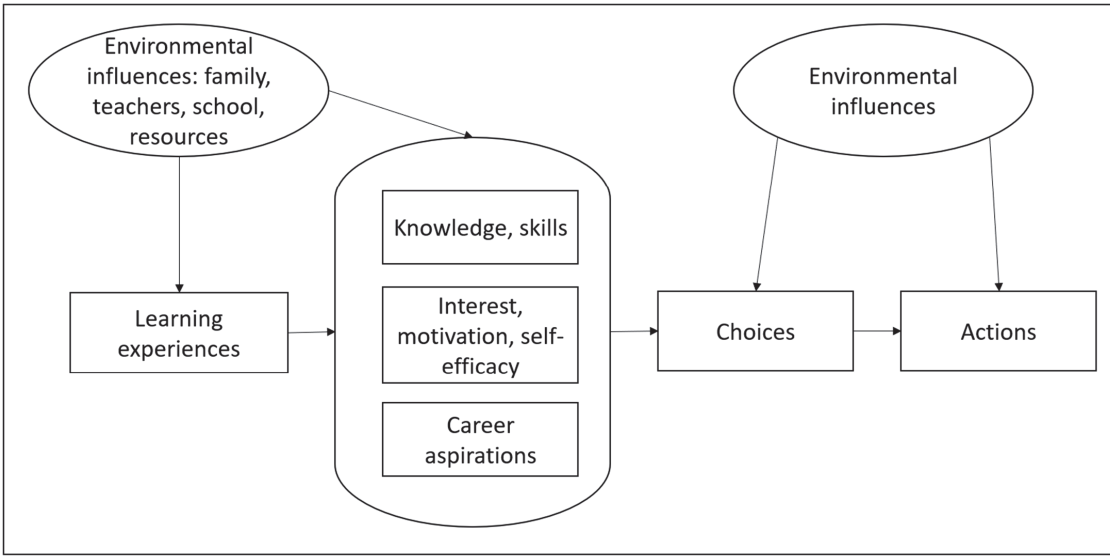


Figure 1. Experiences and influences on the path to STEM.

Positive learning experiences are key to developing favorable attitudes and good skills in STEM. Children and young people engage with STEM learning experiences in different contexts, including school and family. Supporting positive learning experiences in school and other educational settings is one way of making STEM attractive. It is well documented that girls in particular face obstacles in STEM due to external attributions, stereotypes, and lesson designs that appeal more to boys [6–8]. One way of appealing to the target group of girls and motivating them to take up STEM is to make STEM lessons more interdisciplinary, combining popular subjects such, as biology, with more challenging ones, such as physics (see contribution by Bahr and Zinn).

Learning experiences accumulate over time and ideally form positive STEM capital, associated with positive mindsets and STEM identities (see contribution by Davis and Wilson-Kennedy). Such capital significantly impacts student retention and persistence in STEM disciplines [9]. However, not all students have the same chance to develop STEM capital. There are specific at-risk groups, such as low-income students, who struggle to build positive STEM capital due to various obstacles and face particular challenges in their STEM studies due to external barriers (see contributions by Davis and Wilson-Kennedy, Endendijk, and Preuß et al.).

Educators play a crucial role in the STEM development of children and young people. However, they often harbor critical attitudes towards STEM, negative self-assessments, and gender stereotypes. This issue is particularly pronounced among educators of younger children (e.g., pre-school or elementary education), who are predominantly female [10,11]. This raises the need for professional training for all educators focused on gender-sensitive STEM didactics, aiming to improve not only educators' skills but also attitudes and/or self-concepts in STEM (see contribution by Feierabend and colleagues).

The nine articles of the Special Issue concur that developing a positive mindset, self-concept, and self-efficacy expectations are essential for a long-term engagement with STEM. It is not just knowledge and skills in STEM that make these fields attractive for school or career choices. Personal attitudes and attributes, such as interest, aspiration, self-concept of ability, and a sense of belonging to the STEM community, play significant roles in STEM decisions and persistence (see contributions by Reichardt et al. and Hofer et al.).

Another set of variables concerns gender and science cognitions, including implicit gender stereotypes, explicit gender identity, and explicit occupational self-concept. These attitudes towards STEM affect different age groups, from school students (see contributions by Hofer et al. and Reichardt et al.) to university students (see contributions by Endendijk and Mouton et al.), as well as professionals such as educators (see contribution by Feierabend et al.).

Not all children, young people, or adults have equal access to STEM. Several articles in this Special Issue examine at-risk groups. Drawing on numerous studies and a long tradition of research, this naturally includes girls and women, who face gender stereotypes [2,5,7]. Other affected groups include individuals from lower socio-economic backgrounds with less access to STEM experiences or university students who face difficult external barriers in their studies [12]. Educators and teachers in different educational institutions can benefit from diagnostic tools that identify individual risks, such as profiles of at-risk students (see contribution by Mouton et al.). Furthermore, educators and students can benefit from support measures tailored to at-risk individuals (see contributions by Reichardt et al. and Endendijk).

Finally, the pathway to STEM and the decision to remain on or leave this path should be considered within a temporal framework [7,12]. Interest in or rejection of STEM begins in early childhood, with gender stereotypes taking effect at a young age. Critical phases require understanding and support to foster sustained engagement and success in STEM fields. For example, early adolescents tend to be more gender egalitarian or favor their own gender, but by late adolescence, stereotypes typically shift towards the traditional view that boys are better at STEM (see contribution by Starr et al.).

Ultimately, it is essential to provide students with favorable conditions on the pathway to STEM to ensure individuals embark on and remain on this path. This involves creating an environment rich in supportive resources, both social and material, that nurture interests, motivation, and positive self-assessments in STEM from an early age. By addressing and mitigating the effects of stereotypes and anxiety, particularly for at-risk groups, and supporting STEM educators, we can foster more inclusive and sustained engagement in STEM fields.

Current research has thoroughly examined the reasons and profiles of students staying in or leaving STEM. Coming back to Blickenstaff's leaky pipeline [5], future research may examine the pipeline aspect longitudinally and aim to reveal specific critical events or triggers that lead to decisions for or against STEM.

Author Contributions: Conceptualization, M.P.; writing, M.P. and B.E.; writing—review and editing, M.P., B.E., S.L., C.R.S. and E.M. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Organisation for Economic Cooperation and Development. PISA 2024 Strategic Vision and Direction for Science. 2020. Available online: https://www.oecd.org/pisa/publications/PISA-2024-Science-Strategic-Vision-Proposal.pdf (accessed on 10 February 2021).
2. Lane, C.; Kaya-Capocci, S.; Kelly, R.; O'Connell, T.; Goos, M. Fascinating or dull? Female students' attitudes towards STEM subjects and careers. *Front. Psychol.* **2022**, *13*, 959972. [CrossRef] [PubMed]
3. Mazana, M.Y.; Montero, C.S.; Casmir, R.O. Investigating students' attitude towards learning mathematics. *Int. Electron. J. Math. Educ.* **2019**, *14*, 207–231. [CrossRef] [PubMed]
4. Hasenhütl, S.; Luttenberger, S.; Macher, D.; Eichen, L.; Eglmaier, M.W.T.; Paechter, M. Empowering educators: A training for pre-service and in-service teachers on gender-sensitive STEM instruction. *EURASIA J. Math. Sci. Technol. Educ.* **2024**, *20*, em2452. [CrossRef] [PubMed]
5. Blickenstaff, J.C. Women and science careers: Leaky pipeline or gender filter? *Gender and Education.* **2005**, *17*, 369–386. [CrossRef]
6. Bieg, M.; Goetz, T.; Wolter, I.; Hall, N.C. Gender stereotype endorsement differentially predicts girls' and boys' trait-state discrepancy in math anxiety. *Front. Psychol.* **2015**, *6*, 1404. [CrossRef] [PubMed]
7. Ertl, B.; Luttenberger, S.; Paechter, M. The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* **2017**, *8*, 703. [CrossRef] [PubMed]

8. Martynenko, O.O.; Pashanova, O.V.; Korzhuev, A.V.; Prokopyev, A.I.; Sokolova, N.L.; Sokolova, E.G. Exploring attitudes towards STEM education: A global analysis of university, middle school, and elementary school perspectives. *EURASIA J. Math. Sci. Technol. Educ.* **2023**, *19*, em2234. [CrossRef] [PubMed]
9. Godec, S.; Archer, L.; Moote, J.; Watson, E.; DeWitt, J.; Henderson, M.; Francis, B. A Missing piece of the puzzle? Exploring whether science capital and STEM identity are associated with STEM study at university. *Int. J. Sci. Math. Educ.* **2024**. [CrossRef]
10. Bayanova, A.R.; Orekhovskaya, N.A.; Sokolova, N.L.; Shaleeva, E.F.; Knyazeva, S.A.; Budkevich, R.L. Exploring the role of motivation in STEM education: A systematic review. *EURASIA J. Math. Sci. Technol. Educ.* **2023**, *19*, em2250. [CrossRef] [PubMed]
11. Ortiz-Revilla, J.; Ruiz-Martín, Á.; Greca, I.M. Conceptions and attitudes of pre-school and primary school teachers towards STEAM education in Spain. *Educ. Sci.* **2023**, *13*, 377. [CrossRef]
12. Potvin, P.; Hasni, A. Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Stud. Sci. Educ.* **2014**, *50*, 85–129. [CrossRef]

education sciences

Article

Gender Differences in the New Interdisciplinary Subject Informatik, Mathematik, Physik (IMP)—Sticking with STEM?

Tobias Bahr * and Bernd Zinn

Institute of Educational Science, Department of Vocational Education focused on Teaching Technology, University of Stuttgart, 70174 Stuttgart, Germany; zinn@ife.uni-stuttgart.de
* Correspondence: bahr@ife.uni-stuttgart.de

Abstract: The current state of research in computer science education outlines gender differences in motivation, interest, and elective subject decisions in favor of male students. This study takes an exploratory approach to examine the gender differences in the interdisciplinary STEM profile subject Informatik, Mathematik, Physik (in short: subject IMP), which combines the three subjects of computer science, mathematics, and physics. A survey was conducted involving n = 336 (m = 236, f = 88, o = 12) subject IMP students in the 10th grade attending a Gymnasium in Baden-Württemberg, Germany. The deciding factors for choosing the subject, subject interest, motivation, and more were measured using a questionnaire. Overall, the subject IMP is most chosen by male students. For those students choosing the subject IMP, no statistically significant gender differences in subject interest in IMP, mathematics, and the STEM area or in motivation and vocational orientation in natural science and engineering were found in contrast to the state of research. The interdisciplinary character of the subject IMP could be more appealing to girls than computer science by itself. We conclude that, with a higher participation rate of female students, the subject IMP could be a first step in getting more women into STEM fields.

Keywords: computer science education; gender differences; motivation; interest; IMP; subject specific analyses; STEM education; STEM careers; gender motivation self-concept interests

Citation: Bahr, T.; Zinn, B. Gender Differences in the New Interdisciplinary Subject Informatik, Mathematik, Physik (IMP)—Sticking with STEM?. *Educ. Sci.* **2023**, *13*, 478. https://doi.org/10.3390/educsci13050478

Academic Editors: Erin Mackenzie, Bernhard Ertl, Christine R. Starr, Silke Luttenberger and Manuela Paechter

Received: 20 December 2022
Revised: 11 January 2023
Accepted: 13 January 2023
Published: 9 May 2023

1. Introduction

The current gender gap exists in many STEM (science, technology, engineering, and mathematics) fields at different educational levels in Germany. Beginning from subjects in secondary schools—such as computer science (CS) [1], physics, engineering/technology, chemistry, biology, and mathematics—until the university [2], the most prominent fields with a gender gap are CS and engineering [2]. A similar distribution can be observed on the job market [2–5]. The percentages can be as low as 13.7% for women working in computer science professions in Germany [2]. At the same time, there exists a shortage of skilled professionals in many STEM fields (such as computer science, physics, and mathematics) in Germany [5–7]. Combining the two facts mentioned above, education has to adapt. Monitoring those changes to better understand why students choose the subjects they do and to give research-based recommendations to teachers and educators has never been more important. Furthermore, topics such as AI, cryptography, simulation, data literacy, and computational thinking have become more important for today's society and that of the future [8–10]. When dealing with those changes and when seeking to shape the society of the future, the gender gap plays an important role, since the aforementioned topics bring many ethical, legal, and social implications (ELSIs) with them. As one way to address the shortage of skilled professionals and to teach students the basic knowledge needed to understand the fundamentals of CS and acquire future skills, new interdisciplinary subjects, programs, and courses have been created by politicians and educational experts to foster students' interest in STEM at different schools, levels, and countries [11,12]. Nevertheless, in many federal states in Germany, CS and other interdisciplinary subjects related to CS are not compulsory [13]. Overall, the structure of CS as a subject in secondary schools in Germany is diverse [13]. CS is an elective subject in some grades and compulsory in others. In some federal states, it is only an elective subject throughout middle school, while in others it is compulsory, such as in Baden-Württemberg, where CS is taught for one hour per week in the 7th grade. Since 2018, students in the 8th grade at secondary schools called Gymnasien in Baden-Württemberg have been able to choose the new interdisciplinary profile subject IMP. This new subject combines the three subjects of CS, mathematics, and physics as a continuation of the CS profile and, more specifically, a STEM profile [14,15]. As one subject of the four-hour profile, students can choose between a third foreign language, another interdisciplinary natural science and technology subject called Naturwissenschaft und Technik (NwT)—with a focus on technology that complements the natural sciences—and IMP, which is the focus of this article. Students studying IMP can learn the skills of CS, mathematics, and physics, along with the interdisciplinary intersections between those subjects, such as computational physics and numerical methods in mechanics, the physical fundamentals of information systems, and cryptography [15]. In the summer of 2021, the first IMP students reached the end of the three-year-long profile subject. Afterwards, they had the option to continue with CS as a chosen, basic, or performance course (Figure 1) for 2, 3, or 5 h per week, respectively, in the higher grades.

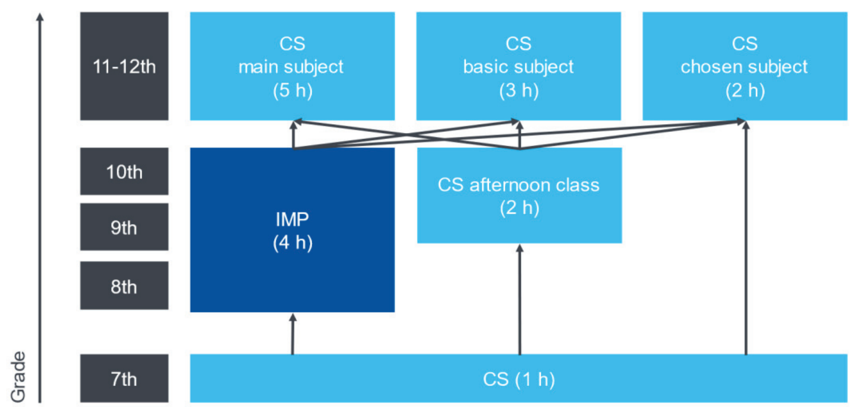


Figure 1. Structure of CS classes in Baden-Württemberg, edited; original by Koch, A. and Mittag, A., licensed under CC BY-SA 4.0 [13].

The course selection in Baden-Württemberg allows for different profiles, with some rules. The students must choose courses totaling at least 32 h per week, including three performance subjects, two courses from German, mathematics, one foreign language, and one natural science (biology, chemistry, or physics), as well as some basic or chosen subjects [16]. CS as a performance subject or basic subject is currently in a school pilot phase, and the subject is only available in some schools. Due to the lack of CS teachers in Baden-Württemberg [17], it is likely that not all of the schools can provide the two new subjects.

1.1. Theory

This study focused on the learning characteristics of IMP students. Learning characteristics are individual prerequisites such as interest, motivation, and self-concept. Starting with the situated expectancy value theory (SEVT) developed by Wigfield and Eccles [18], many empirical studies [19–21] have shown the interdependencies described in models. As in the model described above, learning characteristics influence the course selection and vocational orientation of students.

The theoretical basics of interest, along with the subject-specific interest in IMP for this study, are described by the construct of interest developed by Krapp et al. [22], with reference to the self-determination theory (SDT) of Deci and Ryan [23]. With a situational incentive quality, the attention of the students should be focused on the subject of learning. Objects of learning can be activities, topics, and areas of expertise [24]. Situational interest can be developed through the learning situation or the learning object. Fostering situational interest in the classroom can lead to individual interest as a personality-specific characteristic (e.g., subject interest) in students.

Studies have shown the influence of the academic self-concept on students' career choices and educational decisions, as well as on their performance [25–27]. In addition to other learning characteristics, studies have shown that the subject-specific self-concept is the most significant moderating variable for the variance of the subject interest [28]. As one facet of the academic self-concept, the subject-specific self-concept differentiates the academic self-concept for the specific subjects [25].

To measure the students' motivation (amotivated, extrinsic, introjected, identified, intrinsic, or interested), we fell back on the known scales described by Prenzel [29], based on the SDT of Deci and Ryan [23], according to which motivation can vary with respect to the self-determination aspects of learning. Students who identify with the subject of study and take pleasure in learning more about it are intrinsically motivated. "Interested" is the state where the students themselves want to learn more about the subject of study [24].

The vocational orientation and the influencing factors of the vocational choices were assessed in this study. To better understand the reasons behind the vocational choices, the quantitative results of the influencing factors of educational choices from the questionnaire (adapted from [30,31]) were supplemented with the analysis of the qualitative results from the open questions asking about the influencing factors of the choice of IMP and CS courses in the higher grades. In the best case, the qualitative analysis can explain and indicate trends or add information to the results of the questionnaire items [32].

1.2. State of Research

According to the current state of research, there is no systematic method to collect data on the learning characteristics of IMP students. Therefore, the only option is to view the state of research for each individual STEM subject or comparable interdisciplinary subject. In this subsection, the gender differences in learner characteristics in STEM subjects in Germany are presented. Gender differences occur in some countries but not others, due to differences in educational systems, interventions, programs, influences from the social background, etc. For the state of research, studies regarding the gender differences in the learning characteristics of students from Germany were viewed.

The subject interest in CS [33], physics [28,34,35], and mathematics [36] in different grades of the school is higher for male students than for female students, while the subject interest in biology is higher for female students than for male students [37]. The state of research on chemistry education in Germany does not show an unambiguous preference towards one gender, and Krapp and Prenzel reported the importance of the differentiation of thematic fields within all STEM subjects with respect to the subject interest [38]. Hence, the state of research in STEM education regarding subject's interest can also be seen through a critical lens, as the interest might not be the same for all topics within a subject and different interest profiles can be measured [28,38,39].

The state of research regarding the academic and subject-specific self-concept is mostly homogeneous for the STEM subjects, with the exception of biology. In CS, gender differences appear as early as primary school [40] and in the programming-related self-concept [41]. They can be found in secondary schools [33], universities [42], and even among computing professionals [43]. In physics and chemistry, the same gender differences are prevalent in middle school [25]. Especially in mathematics, large-scale studies such as PISA and TIMSS have reported gender differences in the math-related self-concept as well as the STEM-related self-concept [26,44,45]. The gender differences mentioned thus far are stereotypically all in favor of male students, whereas the biology self-concept favors female students [44].

Another aspect to consider is students' negative associations with STEM stereotypes as well as role models in STEM fields [1]. The latter could influence women's implicit STEM cognitions [46], foster a sense of belonging and raise their expectations of success [47]. Maybe regardless of their gender [48,49], teachers could apply strategies to foster girls' interest in CS [1,49], introduce students to positive role models and programs, or be a role model themselves [1,46,50].

To extend the scientific discourse on CS and STEM education and gender, the focus of this study is on the following two research questions:

(RQ1) By which self-concept, motivation, interests, vocational orientation, and factors influencing educational choices can IMP students in the 10th grade be characterized?

(RQ2) What gender-specific differences in the self-concept, motivation, interests, vocational orientation, and factors influencing the educational choices of IMP students exist between male and female IMP students?

2. Materials and Methods

2.1. Participants and Procedure

The study was conducted from 27 April 2022 to 4 July 2022 via an online survey in Baden-Württemberg, Germany. Overall, n = 336 (*male* = 236, *female* = 88, *other* = 12) IMP students at the end of the 10th grade with an average age of 16 years (*Min* = 15.0, *Max* = 19.0, *M* = 15.98, *SD* = 0.70) attending a Gymnasium participated in the study. In total, 20% of the IMP students in the 10th grade during that period of time participated in the study (information from September 2022 from the Ministry of Education, Cultural Affairs, Youth and Sports of Baden-Württemberg). From all 99 Gymnasien (information from September 2020 from the Ministry of Education, Cultural Affairs, Youth and Sports of Baden-Württemberg) that introduced IMP in the school year 2019/2020 or earlier, 31 from the different administrative districts (Freiburg n = 5, Karlsruhe n = 9, Stuttgart n = 12, Tübingen n = 5) of Baden-Württemberg participated. The Gymnasien received an e-mail with a link to the online survey and a declaration of consent for the parents. The participation in the online survey was anonymous and voluntary, as explained in the covering letter to the schools, parents, and students.

2.2. Instruments

The original version of the questionnaire, in German, can be found in the Supplementary section of this article. We assessed demographic data, such as a code for re-identification if a participant wished to revoke their participation, along with age, gender (male, female, other), and their school. To measure the students' motivation (amotivated, extrinsic, introjected, identified, intrinsic, or interested), we adapted the known scales described by Prenzel et al. [24]. The chosen subjects and their interest in them, along with the subject-specific self-concept (adapted from [27,51,52]), academic self-concept (adapted from [53]), and performance-related expectations associated with CS in higher grades (adapted from [27]), area-specific interest (adapted from [22]), subject interest in IMP (adapted from [54]), factual interest, vocational orientation (adapted from [30,31]), and factors influencing their educational choices [31,55] were also assessed. At the end of the questionnaire, we asked the IMP students about grades and some open questions. As indicated in Table 1 below, the Cronbach's α values to measure the internal consistency of all applied instruments were reliable ($\alpha > 0.75$ [56]), except for extrinsic motivation ($\alpha = 0.684$ [56]). This result was to be expected, as we used tested and validated instruments. Not a single item was excluded. We analyzed the results using SPSS 28.0.0.0 and R 4.3.1.

Table 1. Test quality criteria of the applied scales.

Scale	Subscale	Cronbach's α	Number of Items	M	SD	*n*
Subject-specific self-concept		0.950	10	3.05	0.78	321
Academic self-concept		0.964	11	3.43	0.93	331
Performance-related attitudes towards CS in higher grades		0.945	7	2.71	0.88	225
Motivation	Amotivated	0.820	3	2.38	1.17	333
	Extrinsic	0.684	3	2.28	1.04	333
	Introjected	0.751	3	3.22	1.05	332
	Identified	0.757	3	3.37	1.08	331
	Intrinsic	0.816	3	2.70	1.15	332
	Interested	0.886	3	2.84	1.20	333
Area-specific interest	Linguistic–literary–artistic	0.893	4	2.27	0.85	331
	Social science	0.884	4	2.74	0.79	329
	STEM	0.890	4	3.18	0.77	330
Subject interest in IMP		0.916	18	2.37	0.67	299

The answers to the open questions such as "name the three most important reasons for choosing the profile subject IMP?" and "name the three most important reasons why you decided to take or not to take the CS course in higher grades?" were analyzed using the systematic, rule-guided, qualitative content analysis described by Mayring [57]. The deductive categories (see the Supplementary) were the course selection and deselection motives described by Eitemüller and Walpuski [55]. The only two inductive categories were added after 25% of the analyzed answers were coded, since many students named "Programming" and "Future relevance" among their reasons for choosing IMP or CS in higher grades. After adding these categories, the answers were analyzed again. Programming, as one aspect of the CS part of IMP, was therefore excluded from the categories "Subject and factual interest" and "interest in working methods". If the students did not explicitly mention the importance for their future career or study and only stated that the content of the subject and/or the subject itself would be relevant in the future, the answer was assigned to the category "Future relevance". No answer was coded in more than one category. In addition to the factors influencing the educational choices within the questionnaire, the open questions provided the IMP students with the option to state their own reasons without being restricted by the questionnaire. The answers to the open questions have the benefit of being more specific with respect to the selection of IMP and CS courses in higher grades. The interrater agreement between the first rater (i.e., the first named author of this article) and the second rater (i.e., a trained student assistant) was 88% for the motives for selecting IMP, 81% for the motives for selecting CS in higher grades, and 80% for the motives for dropping CS in higher grades. The Cohen's kappa values for the three categories were substantial [58,59]. We used MAXQDA 20.4.2 for the data analysis.

3. Results

To answer RQ2, the following subsections cover the gender differences in the motivational and affective determinants of the IMP students in this cohort.

3.1. Gender Differences

The significance of gender differences in the subject-specific self-concept in IMP ($U = 7803$, $p = 0.030$, $Z = -2.165$) and academic self-concept ($U = 7263$, $p < 0.001$, $Z = -3.768$) was assessed using the Mann–Whitney U test. There was homogeneity of the variances as assessed by Levene's test for the subject-specific self-concept ($p = 0.887$) and the academic self-concept ($p = 0.179$) (Table 2). None of the populations were normally distributed according to the Shapiro–Wilk test ($p < 0.001$). The analysis of the effect size r indicated that female IMP students had a slightly lower subject-specific self-concept in IMP ($r = -0.124$) and academic self-concept ($r = -0.211$) than male IMP students in this cohort. Instead of computing Cohen's d, it is recommended to report the effect size r for non-parametric tests such as the Mann–Whitney U test [60,61]. The analysis showed that for this cohort of IMP students, gender-related differences in the self-concepts were prevalent, as also shown in other STEM fields in secondary schools in Germany [25,26]. As with other STEM subjects [28,34,35], male IMP students in this cohort were more interested in physics ($U = 8018$, $p = 0.002$, $Z = -3.131$, $r = -0.175$) and CS ($U = 7303$, $p < 0.001$, $Z = -3.531$, $r = -0.200$), while female IMP students were more interested in biology ($U = 7413$, $p < 0.001$, $Z = -3.930$, $r = -0.219$) and in the linguistic–literary–artistic areas ($U = 5570$, $p < 0.001$, $Z = -6.193$, $r = -0.348$), as assessed by the Mann–Whitney U test (Table 2). None of the populations were normally distributed according to the Shapiro–Wilk test ($p < 0.001$). Considering the unequal sample sizes, the group differences were assessed by the Mann-Whitney U test [62–64].

Table 2. Differences between male and female IMP students.

	Shapiro–Wilk Test	**Levene's Test**	**Mann–Whitney U Test/ *t*-Test/Chi-Squared Test**	**Gender *n* (Male, Female)**
Subject-specific self-concept	$p < 0.001$	$p = 0.887$	$U = 7803$, $p = 0.030$, $Z = -2.165$, $r = -0.124$	226, 83
Academic self-concept	$p < 0.001$	$p = 0.179$	$U = 7263$, $p < 0.001$, $Z = -3.768$, $r = -0.211$	232, 87
Performance-related expectations towards CS in higher grades	$p < 0.001$	$p = 0.120$	$U = 2758$, $p = 0.006$, $Z = -2.733$, $r = -0.186$	172, 44
Amotivated	$p < 0.001$	$p = 0.241$	$U = 9055$, $p = 0.155$, $Z = -1.422$	234, 87
Extrinsic motivation	$p < 0.001$	$p = 0.158$	$U = 9189$, $p = 0.253$, $Z = -1.143$	235, 86
Introjected motivation	$p < 0.001$	$p = 0.227$	$U = 8554$, $p = 0.050$, $Z = -1.961$	234, 86
Identified motivation	$p < 0.001$	$p = 0.442$	$U = 9801$, $p = 0.855$, $Z = -0.182$	233, 86
Intrinsic motivation	$p < 0.001$	$p = 0.675$	$U = 9272$, $p = 0.286$, $Z = -1.066$	233, 87
Interested	$p < 0.001$	$p = 0.775$	$U = 8855$, $p = 0.091$, $Z = -1.693$	234, 87
Interest in linguistic–literary–artistic areas	$p < 0.001$	$p = 0.392$	$U = 5570$, $p < 0.001$, $Z = -6.193$, $r = -0.348$	231, 88
Interest in social science areas	$p < 0.001$	$p = 0.045$	$U = 7600$, $p = 0.001$, $Z = -3.275$, $r = -0.184$	231, 87
Interest in STEM areas	$p < 0.001$	$p = 0.406$	$U = 8888$, $p = 0.135$, $Z = -1.495$	231, 87
Subject interest in IMP	$p = 0.122$	$p = 0.733$	$t\ (287) = 1.794$, $p = 0.074$	210, 79
Interest in CS	$p < 0.001$	$p = 0.186$	$U = 7303$, $p < 0.001$, $Z = -3.531$, $r = -0.200$	266, 87
Interest in mathematics	$p < 0.001$	$p = 0.996$	$U = 10176$, $p = 0.864$, $Z = -0.172$	236, 88
Interest in physics	$p < 0.001$	$p = 0.050$	$U = 8018$, $p = 0.002$, $Z = -3.131$, $r = -0.175$	235, 88
Interest in biology	$p < 0.001$	$p = 0.512$	$U = 7413$, $p < 0.001$, $Z = -3.930$, $r = -0.219$	235, 88
Interest in implementing code	$p < 0.001$	$p = 0.917$	$U = 7802$, $p = 0.002$, $Z = -3.065$, $r = -0.172$	230, 86
Interest in algorithms	$p < 0.001$	$p = 0.187$	$U = 6948$, $p < 0.001$, $Z = -4.074$, $r = -0.229$	231, 87
Interest in requirement analysis	$p < 0.001$	$p = 0.359$	$U = 7814$, $p = 0.002$, $Z = -3.029$, $r = -0.170$	231, 87
Interest in software projects	$p < 0.001$	$p = 0.917$	$U = 7151$, $p < 0.001$, $Z = -3.845$, $r = -0.217$	231, 86

Table 2. *Cont.*

	Shapiro–Wilk Test	Levene's Test	Mann–Whitney U Test/ *t*-Test/Chi-Squared Test	Gender *n* (Male, Female)
Vocational orientation: pursuing a career or study in mathematics or CS	$p < 0.001$	$p = 0.130$	$U = 8264$, $p = 0.018$, $Z = -2.360$, $r = -0.133$	230, 87
Vocational orientation: pursuing a career or study in science or engineering/technology	$p < 0.001$	$p = 0.189$	$U = 8700$, $p = 0.088$, $Z = -1.706$	229, 87
Factors influencing the educational choice: career prospects	$p < 0.001$	$p = 0.148$	$U = 9146$, $p = 0.038$, $Z = -2.079$, $r = -0.118$	225, 86
Choice of CS in higher grades	$p < 0.001$	$p = 0.008$	$X^2(1, n = 313) = 6.278$, $p = 0.012$, $\phi = -0.142$, $p = 0.012$	231, 86

3.2. No Gender Differences

In contrast to the state of research in other STEM subjects [28,33,36], there were no significant gender differences in any of the six facets of motivation [65–67], the interest in the STEM area, or the subject interest in IMP and mathematics, as assessed by the Mann–Whitney U test and *t*-test (Table 2). The subject with the highest interest was mathematics, with no gender difference. As the IMP students chose a STEM profile subject with four hours per week in the 8th grade, it was not a big surprise that there were no gender differences in interest in this area and subject. Mathematics is one of the key subjects comprising the theoretical background of the interdisciplinary topics mentioned in Section 1, as well as being a central part of the subject IMP. This could explain why, in contrast to the state of research in mathematics education [36], there were no gender differences in the interest in mathematics. The interdisciplinary context of IMP, which is built around CS, could be more appealing to female students, and this could explain why there were no gender differences in any of the six facets of motivation—even intrinsic motivation and interest (Table 2). Looking at the vocational orientation, there were no significant gender differences for the item "pursuing a career or study in science or engineering/technology", as assessed by the Mann–Whitney U test ($U = 8700$, $p = 0.088$, $Z = -1.706$). The sample size of only 336 (20% of all IMP students in the 10th grade at that time) could be a limiting factor, but the fact that there were no gender differences stands in contrast to the current gender gap in universities (24.2% female students in the first semester studying CS, 21.5% female students with bachelor's degrees in CS, and 23.5% female students with master's degrees in CS at German universities in 2021 [68]) and on the job market [1,2,5]; however, the overall lack of a gender difference could imply that the IMP students tend to choose an education or career in science or engineering/technology (Figure 2).

Overall, most of the IMP students tend to choose an education or career in the fields of science, engineering/technology, mathematics, or CS (Figure 2). Since IMP is the only way for students in Baden-Württemberg to continue with a subject in the field of CS, we also examined the gender differences in the choice of CS in higher grades.

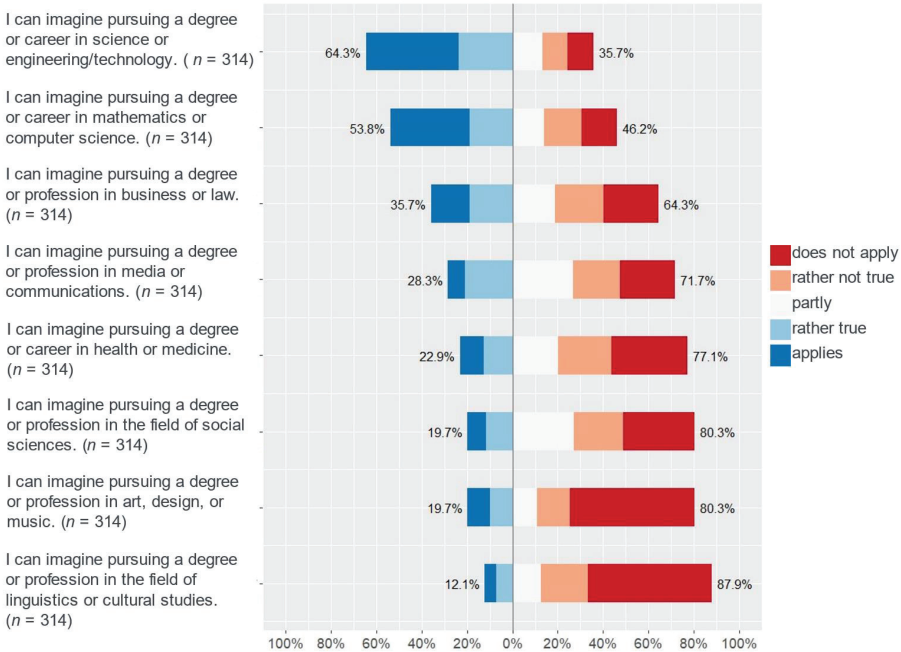


Figure 2. Vocational orientation of IMP students.

3.3. Gender Differences in the Choice of CS in Higher Grades

Overall, 40.7% of IMP students who chose CS in higher grades in this cohort had higher intrinsic motivation, a higher subject-specific and academic self-concept, better IMP grades, higher performance-related expectations towards CS in higher grades, were more interested, and had a higher subject interest in IMP than IMP students who did not choose CS in higher grades, as assessed by the Mann–Whitney U test (Table 3). There was a small but significant gender difference ($X^2(1, n = 313) = 6.278$, $p = 0.012$ [69]) in the choice of CS in higher grades. Consistent with the state of research in CS education, female IMP students were less likely to select CS as a subject in higher grades ($\phi = -0.142$, $p = 0.012$). However, while female students usually drop out of CS with increasing age, 30.7% of the female IMP students in this cohort did not drop out of the CS field after IMP. Since the IMP students choose a STEM profile for 4 h per week very early on, a positive selection bias occurs. Even so, the fact that female students who choose IMP tend not to drop out at the same rate as in other subjects is an interesting trend.

On the other hand, the analysis of the performance-related expectations towards CS in higher grades showed a significant gender difference according to the Mann–Whitney U test ($U = 2758$, $p = 0.006$, $Z = -2.733$, $r = -0.186$) in favor of male IMP students. The results show that male IMP students had slightly higher performance-related expectations towards CS in higher grades. We also analyzed the influencing factors of the choice of CS in higher grades named by the IMP students in the open items, so as to provide an in-depth description of the factors influencing the educational choices according to the questionnaire. A total of 115 ($m = 86, f = 27, o = 2$) IMP students named the three most important reasons for choosing CS in higher grades. In accordance with the SEVT [18], subject and factual interest, vocational orientation, and a positive self-concept were the most commonly mentioned influencing factors for choosing CS in higher grades (Figure 3). Included in the subject interest, “programming” was also named as an influencing factor by the IMP students. The third most frequently mentioned reason was “contribution obligations”. Students chose CS because they “Did not want to select other subjects in higher grades”. Due to the rules regarding the profile choices for students in Baden-Württemberg, as mentioned in the Introduction, the students have to choose at least 32 h of courses per week. The contribution obligations (i.e., the obligation to choose one of the profile subjects) could also be influenced by the expectation of success [18], which could be higher in CS than in other subjects, with CS being seen as the “lesser evil” (54 Pos. 2). Another reason for choosing CS in higher grades was the relevance of the topics or the subject CS itself for the future. The IMP students also named several individual reasons for choosing CS, such as CS being a “young science” (303 Pos. 2) and “[I] had the chance through IMP, so I took it” (266 Pos. 2).

Table 3. Comparison of IMP students who chose CS in higher grades (a) and those who did not (b).

Constructs	Shapiro–Wilk Test	Levene's Test	Mann–Whitney U Test	*n* (a, b)
Subject-specific self-concept	$p < 0.001$	$p < 0.001$	$U = 5507$, $p < 0.001$, $Z = -8.038$, $r = -0.455$	129, 183
Academic self-concept	$p < 0.001$	$p = 0.494$	$U = 5934$, $p < 0.001$, $Z = -8.071$, $r = -0.450$	133, 189
Performance-related expectations towards CS in higher grades	$p < 0.001$	$p < 0.581$	$U = 1787$, $p < 0.001$, $Z = -8.653$, $r = -0.586$	129, 89
IMP grade	$p < 0.001$	$p < 0.001$	$U = 8481$, $p < 0.001$, $Z = -5.438$, $r = -0.300$	133, 194
Intrinsic motivation	$p < 0.001$	$p = 0.466$	$U = 6038$, $p < 0.001$, $Z = -8.010$, $r = -0.445$	131, 193
Interest	$p < 0.001$	$p = 0.030$	$U = 7202$, $p < 0.001$, $Z = -6.631$, $r = -0.368$	132, 192
Subject-specific interest	$p = 0.432$	$p = 0.791$	$U = 4630$, $p < 0.001$, $Z = -8.117$, $r = -0.474$	122, 171

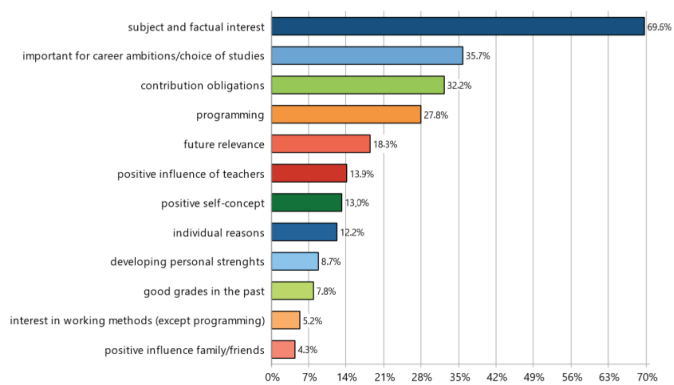


Figure 3. Motives of IMP students for selecting CS in higher grades.

According to the questionnaire, from the factors influencing the educational choices of the IMP students, two gender differences were identified, with “career prospects” being slightly more important for male IMP students in this cohort according to the Mann–Whitney U test ($U = 8164$, $p = 0.038$, $Z = -2.079$, $r = 0.118$), while “teaching” was slightly more important for female IMP students in this cohort according to the Mann–Whitney U test ($U = 8427$, $p = 0.037$, $Z = -2.086$, $r = 0.118$). For “teaching”, there was no normal distribution as assessed by the Shapiro–Wilk test ($p < 0.001$) and no homogeneity of variance according to Levene's test ($p = 0.029$), but there was a homogeneity of variance for “career prospects” according to Levene's test ($p = 0.148$). Therefore, it could be beneficial to examine gender differences in the choice of CS in higher grades (Figure 3). Both male and female IMP students named subject interest and factual interest as their most important reasons for choosing CS (Figure 4). Career ambitions were as important for male IMP students as they were for female IMP students in the choice of CS in higher grades ($X^2(1, n = 321) = 0.002$, $p = 0.966$). The quantitative results show that the positive influence of teachers was mentioned significantly more frequently by female IMP students; however, there were no gender differences in this respect as a reason for choosing CS in higher grades ($X^2(1, n = 321) = 0.921$, $p = 0.337$). Hence, these two gender differences in the influencing factors did not apply for the choice of CS in higher grades in the same way as they did for general educational choices. Female students did not choose CS more due to contribution obligations ($X^2(2, n = 321) = 5.148$, $p = 0.075$). There were also no gender differences in the influence of future relevance with respect to the choice of CS in higher grades ($X^2(1, n = 321) = 0.048$, $p = 0.872$).

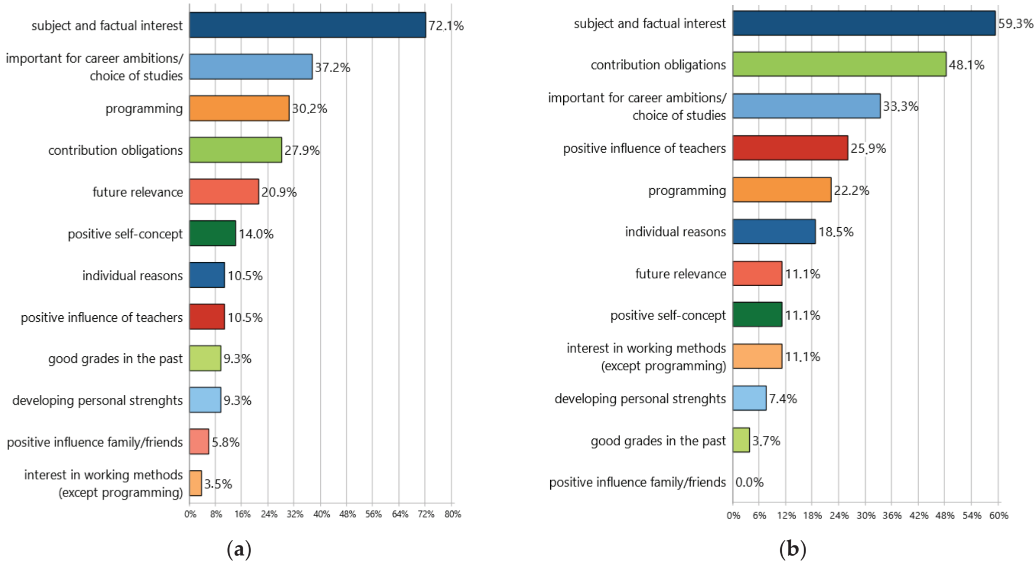


Figure 4. Motives of IMP students for selecting CS in higher grades: (**a**) male IMP students; (**b**) female IMP students.

After looking at the motives of IMP students in choosing CS in higher grades, we analyzed their deselection motives as well (Figure 4) in order to answer the question of why IMP students do not select CS in higher grades after choosing IMP. In total, 187 ($m = 123$, $f = 54$, $o = 10$) IMP students named their three most important reasons for not choosing CS in higher grades. Consistent with the state of research on the course selection motives of students [55], the students named a lack of subject interest or factual interest and a negative self-concept as their main reasons for not choosing CS in higher grades (Figure 5).

When looking at the gender differences in the students' deselection motives, it appears that there were no major differences between female and male IMP students. Only the categories "high workload/complexity of the topics" and "contribution obligations" differed in the order in which they were named by female IMP students, but there was no significant gender difference. Hence, there were no significant associations between gender and the deselection motives "high workload/complexity of the topics" ($X^2(1, n = 321) = 0.194$, $p = 0.660$) and "contribution obligations" ($X^2(2, n = 321) = 5.184$, $p = 0.075$). Individual reasons such as "Basic knowledge from IMP is enough for me" (34 Pos. 2) and "Three years are enough" (45 Pos. 2) indicate that some of the students were already satisfied with their knowledge or did not want to pursue a career in CS. Furthermore, there were gender differences in the item "pursuing a career or study in mathematics or CS" (Table 2), supporting the theory that the choice of CS in higher grades depends on the career choice.

Since female IMP students have a high interest in biology, for example, CS could also be a tool for them to later professionalize in their main field of interest. In summary, female IMP students in this cohort mainly dropped out of CS in higher grades due to their lack of interest, having already picked too many other courses, having a negative self-concept, the high workload, and the complexity of the topics associated with CS. Therefore, it could be interesting to see the paths in other STEM subjects so as to get an idea of which STEM areas the IMP students choose in higher grades and to provide a broader view of the educational choices of the IMP students in this cohort.

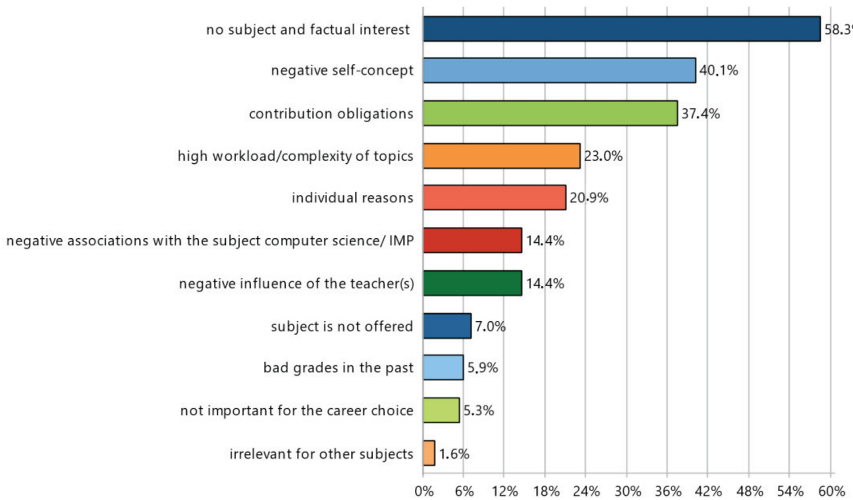


Figure 5. Motives of IMP students for dropping CS in higher grades.

3.4. Educational Choices of IMP Students

To answer RQ1, after taking a closer look at the gender differences in the subsections above, we summarize the data collected in this study in the following two subsections.

Figure 6 gives an overview of the educational choices of the IMP students in this cohort. Gender differences are marked with superscript 1. Male IMP students in this cohort tended to choose CS (basic and performance subjects) (Table 2) and the performance subject physics ($X^2(1, n = 317) = 6.203$, $p = 0.013$) in higher grades significantly more frequently than female students. However, there were no significant gender differences for the performance subjects of biology ($X^2(1, n = 316) = 1.478$, $p = 0.224$) and chemistry ($X^2(1, n = 316) = 0.955$, $p = 0.328$). There were no gender differences for IMP students choosing CS in higher grades in terms of vocational orientation, nor for the fields “science or engineering/technology” ($U = 1170$, $p = 0.328$, $Z = -0.978$), mathematics, or computer science ($U = 1048$, $p = 0.052$, $Z = -1.940$). Of the IMP students choosing a STEM subject other than CS, 62–85% said they could imagine pursuing a career in science or engineering/technology, with no significant gender differences for the IMP students who chose physics ($U = 561$, $p = 0.337$, $Z = -0.960$), biology ($U = 178$, $p = 0.622$, $Z = -0.494$), or chemistry ($U = 182$, $p = 0.193$, $Z = -1.302$). For the field “mathematics or computer science”, there were also no significant gender differences for the IMP students who chose physics ($U = 504$, $p = 0.127$, $Z = -1.524$), biology ($U = 195$, $p = 0.978$, $Z = -0.027$), or chemistry ($U = 213$, $p = 0.645$, $Z = -0.461$).

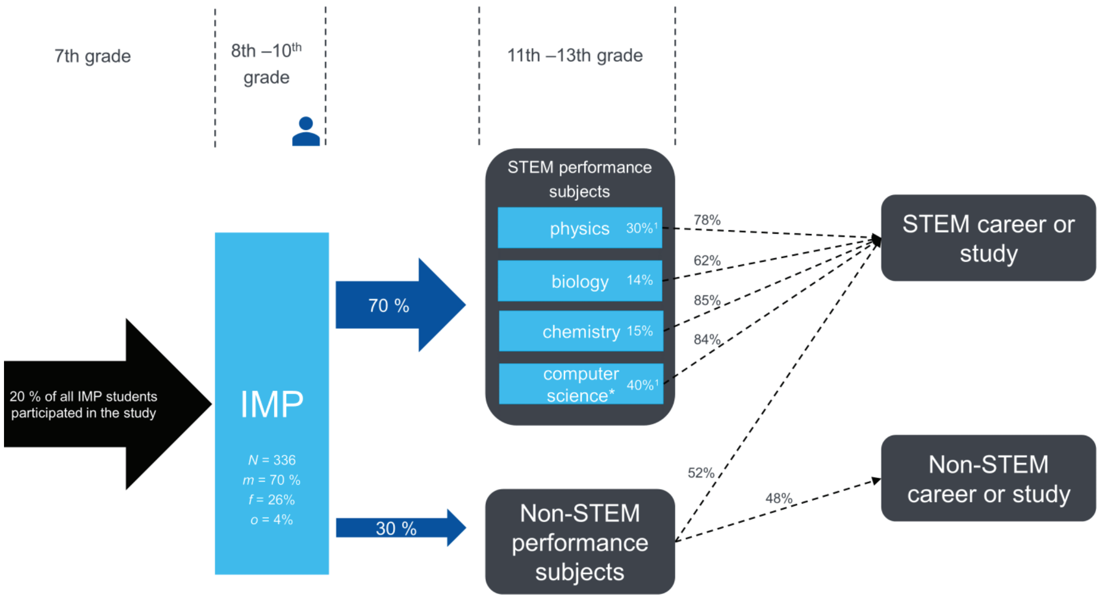


Figure 6. Educational choices of IMP students. [1] significant gender differences * computer science basic and performance subject --► current vocational orientation ➡ course selection of IMP students participating in the study ➡ percentage of IMP students in the 10th grade participating in the study amount of 14% IMP students choosing a subject in higher grades.

At first glance, Figure 6 might appear to suggest that the IMP students in the non-STEM performance subjects group did not choose any STEM subject. Since this is not the case, we wish to refer to the structure of the educational system in Baden-Württemberg, as mentioned in Section 1 (Figure 1). First and foremost, we did not assess whether students picked mathematics as a performance subject in this study since some schools told us after their participation that they could not yet choose the courses for the higher level; therefore, the students only stated which courses they were currently taking, and mathematics is the only STEM performance course that the students can take when they are in middle school. Due to the rules limiting the course selection—as students must choose two subjects out of German, mathematics, a foreign language, and natural science—it is likely that many IMP students chose mathematics as a performance subject because it was the subject in which they had the highest interest. In addition to the performance subjects, IMP students can also choose NwT or one of the other STEM subjects as a basic subject for 3 h per week, in which case they would not be "out" of the STEM field. This could also explain why 52% of the IMP students did not choose one of the aforementioned STEM performance subjects (Figure 6), saying that they were considering a career or education in science or engineering/technology. Only 8.5% of the IMP students chose the performance subject, CS. Due to the subject being in a pilot phase in many schools and the shortage of CS teachers in Baden-Württemberg, as well as the fact that the subject was only introduced in 2021, it is expected that not many students will choose it. Since the basic subject CS is also new and in the school pilot phase, we added CS in Figure 6 without considering it a performance subject. Another factor influencing the choice of CS was the performance-related expectations with respect to CS in higher grades (point-biserial correlation $r = 0.588$, $p < 0.001$), with male IMP students assessing themselves more highly than female IMP students in this cohort (Table 2), which could explain the gender difference in the choice of CS in higher grades. In addition, small gender differences were found in the factual interest within the areas of "implementing code", "algorithms", "requirement analysis", and "software projects". Since those areas are explored in greater depth in CS in higher grades, it is more likely that students who are not interested in these areas will not choose the subject.

The educational choices shown in Figure 6 also visualize the current metaphor of educational choices in the STEM field as a "highway" [70] rather than a "leaky pipeline", since students have the option to go in and out of the STEM field at multiple points in their lives and do not have cumulative disadvantages by dropping out of the STEM field [70].

3.5. Motivation and Interests of IMP Students

The most frequent motivational characteristics of IMP students were introjected and identified (Figure 7). Hence, the motivation of most IMP students in our cohort was mostly without any identification or external goal of their actions [29]. An exception were the IMP students who chose CS in higher grades. Considering that "contribution obligations" were mentioned by 43.9% of the IMP students as their second most important reason for choosing IMP, many of these students chose IMP because they did not like the other profile subjects, as explained by the qualitative data.

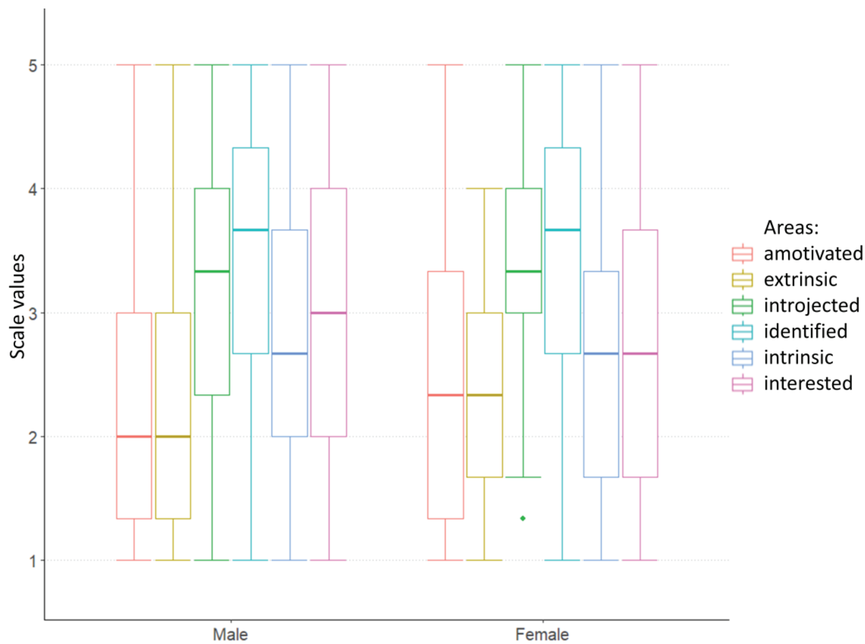


Figure 7. Motivation of IMP students.

The subject-specific interest in IMP (*M* = 2.37, *SD* = 0.67) was relatively low compared to the interest in the STEM area in general (*M* = 3.18, *SD* = 0.77). This result can be traced back to other findings since the IMP students choose various performance subjects in higher grades, with mathematics being the subject with the greatest interest among the IMP students (Table 1).

4. Discussion

Consistent with the state of research on CS education, male IMP students had a slightly higher self-concept (subject-specific, academic, and gender-specific) than female IMP students in this cohort (Table 2). Thus, even when choosing IMP and going through the profile subject for four hours per week, female IMP students—especially in the direct gender comparison to male IMP students—had a lower estimation of their own abilities.

As many literature reviews state, the lack of gender-sensitive CS [1,71–75] could explain the findings in this study.

The gender differences in the choice of CS in higher grades can be explained by the students' reasons for not choosing the subject (Figure 4). Apart from the two previously mentioned influencing factors—interest and self-concept—48% of female IMP students named the curricular requirements and elective options mentioned in Section 1 as limiting factors [16]. Thus, these IMP students did not choose CS on top of their normal schedule. Additionally, 35% named the high workload and complexity of the subject as a reason. This could be because the IMP students associate CS with a high level of effort and an accompanying high workload, as well as because of their rather minimal prior knowledge as they only study CS for one hour per week in the 7th grade. Finally, 13% of the female IMP students named their personal preferences with regard to the teachers as a reason for not choosing CS in higher grades, but without elaborating further.

In this study, we found the relative lack of gender differences interesting, as the new interdisciplinary profile subject IMP might provide an opportunity for more gender equality within the STEM area if more female students choose the subject in the future [75]. After three years, the IMP students in this cohort showed no significant gender differences in motivation (all six facets) or in their subject interests in IMP, mathematics, and the general STEM area. Additionally, 64% of the IMP students said they could imagine pursuing a degree or career in the fields of science or engineering/technology. The second most commonly named field was "mathematics or CS". With 40.7% of the IMP students continuing with CS in higher grades, 30.9% of the IMP students choosing the performance subject physics, 14% choosing the performance subject chemistry, 14% choosing the performance subject biology, and only 30% not choosing one of the aforementioned subjects, IMP could provide a new opportunity for students to choose a STEM profile early on and stick with it (Figure 6). With only 26% of the IMP students being female, there is more work to be done in advertising the subject to female students and creating a gender-sensitive environment in IMP.

4.1. Limitations

Limitations of this study include the lack of a comparison group, the sample size of 336 IMP students (20% of the possible population), and the fact that only 31 schools (31.31% of the total number of schools that have IMP as an option) participated in the survey. Due to the low participation of IMP students who categorized their gender as "other" (3.6%), we only analyzed gender differences for the male and female subgroups. This results in the limitation that not all genders are represented in this study. Since the IMP students chose a STEM profile very early on, a positive selection bias also occurs. The sampling may be a positive selection since 10 schools declined the offer to participate in the survey due to their resources, and 58 did not respond at all.

4.2. Research Desideratum

Different research desiderata exist after taking a closer look at the learning characteristics of IMP students. The use of a comparison group could be an interesting method to add information to the descriptive knowledge of the learning characteristics of IMP students and the students in secondary schools in Baden-Württemberg. Another research desideratum is to assess the subject knowledge in CS as one of the key facets of competence in order to see which factors moderate the CS competence of IMP students and add more information to the learning outcomes regarding the utilization of learning opportunities [76]. Assessing more than the choice of CS in higher grades to see how many students stick with STEM after IMP—and especially, how many female students stick with STEM after IMP—is another topic for future research. The interdisciplinary character of the subject IMP might be more interesting to girls than CS itself; whether this factor has an influence on the course selection in higher grades, and whether it has an impact on motivation and subject interest is another topic for future research. In accordance with the SEVT [18], we assessed the subject-specific interest in IMP, the subject-specific self-concept, the IMP grade, and the educational choices and their motives for choosing IMP and CS in higher grades, as well as the motives for dropping CS in higher grades, and took a look through the lens of gender to describe the current status of gender differences among the IMP students in this cohort. Another research desideratum is to use the aforementioned data for cluster analysis to further characterize the IMP students in this cohort and to better foster the enrolment of different students on the IMP course, along with a research-based recommendation for IMP teachers and school development experts to enhance the IMP classes for everyone.

5. Conclusions

In this section, implications for STEM education and—more specifically—IMP education and didactical implications are presented.

The findings of some studies on the state of research in STEM education can be implemented in the classroom. Teachers building a growth mindset and self-efficacy in STEM subjects in middle school [77] and gender-sensitive CS classes [1] are currently the most prominent recommendations. Male students also benefit from measures to increase the recruitment of female students, as shown in physics by the study of Häußler and Hoffmann [78]. Starting points to increase the attractiveness of IMP could include more woman-oriented teaching contexts or (female) teachers as role models [46–49]. The findings of this study, despite its limitations, indicate no gender differences in the motives for dropping CS in higher grades. Therefore, addressing the lack of subject interest and the negative self-concept through teachers building a growth mindset and self-efficacy in IMP could encourage female and male IMP students alike and make the selection of CS in higher grades more appealing.

Contribution obligations, as the third most significant influencing factor for not selecting CS in higher grades, could be addressed by changing the selection rules in the educational system in Baden-Württemberg. For instance, addressing the shortage of skilled professionals in CS with middle school subjects such as IMP is a first step. Students in higher grades should then be able to continue on the STEM path consistently with performance subjects beyond the natural sciences (i.e., biology, chemistry, and physics). Extending the educational rules so that the interdisciplinary subjects Naturwissenschaft und Technik (NwT) and CS are added to the cluster of what students can choose could bring more students into the STEM area or—more specifically—into CS and engineering in higher grades.

Another area that can be addressed is the subject interest of IMP students and the interdisciplinary content. Since the subject with the third highest interest among female IMP students is biology, followed by chemistry in fourth place, while physics and computer science lag behind, more interdisciplinary STEM topics—such as bioengineering, environmental engineering, and interdisciplinary topics between chemistry and CS—could be integrated into the IMP class to better address the interests of female IMP students. Considering that the fourth most frequently named deselection motive was high workload and the complexity of the topics (Figure 5), the already full schedule should not be further expanded but rather refined. The subject was only selectable in 99 out of the 457 Gymnasien in Baden-Württemberg, and in those 99 Gymnasien, students could choose at least between two and sometimes four profile subjects. Due to the shortage of CS teachers in Germany, IMP is not yet implemented throughout Baden-Württemberg. The aforementioned suggestions can be implemented in the process of expanding the subject throughout Baden-Württemberg.

To conclude, this study provides a first description of the learning characteristics of IMP students and the gender differences among them in the Baden-Württemberg context. The initial trends show no gender differences in the motivation, subject interest in IMP, or interest in the STEM area except in the choice of the performance subjects, physics and CS, and no gender differences in the vocational orientation in the fields of natural science and engineering/technology, in contrast to the state of research in other STEM subjects. The new interdisciplinary profile subject IMP, with a higher participation of female students, could be a first step in getting more people—and probably more female students—into STEM.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci13050478/s1, Document S1: Questionaire of the study; Document S2: Category system.

Author Contributions: Conceptualization, T.B.; methodology, T.B.; validation, T.B.; formal analysis, T.B.; investigation, T.B.; resources, T.B.; data curation, T.B.; writing—original draft preparation, T.B.; writing—review and editing, T.B.; visualization, T.B.; supervision, T.B. and B.Z.; project administration, T.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the University of Stuttgart (Az. 23-014 and date of approval 17 April 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Because the data used in this study are currently being analyzed as part of a doctoral thesis, the dataset will be available from the corresponding author upon request after the completion of the dissertation.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Happe, L.; Buhnova, B.; Koziolek, A.; Wagner, I. Effective measures to foster girls' interest in secondary computer science education. *Educ. Inf. Technol.* **2021**, *26*, 2811–2829. [CrossRef]
2. Anger, C.; Kohlisch, E.; Plünnecke, A. *MINT-Herbstreport 2021. Mehr Frauen für MINT Gewinnen: Herausforderungen von Dekarbonisierung, Digitalisierung und Demografie Meistern*; Institut der deutschen Wirtschaft: Cologne, Germany, 2021.
3. United Nations Educational. *UNESCO Science Report: Towards 2030*; Second Revised Edition; UNESCO: Paris, France, 2016; ISBN 978-92-3-100129-1.
4. Eurostat. More Women Join Science and Engineering Ranks. Available online: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/edn-20220211-2 (accessed on 9 September 2022).
5. Statistik der Bundesagentur für Arbeit. *Berichte: Blickpunkt Arbeitsmarkt—Fachkräfteengpassanalyse 2019*; Berichte: Blickpunkt Arbeitsmarkt, Bundesagentur für Arbeit, Nürnberg; 2020. Available online: https://statistik.arbeitsagentur.de/Statistikdaten/Detail/201912/arbeitsmarktberichte/fachkraefte-engpassanalyse/fachkraefte-engpassanalyse-d-0-201912-zip.zip?__blob=publicationFile&v=5 (accessed on 9 September 2022).
6. Statistik der Bundesagentur für Arbeit. Fachkräfte für Deutschland: Zwischenbilanz und Fortschreibung, Bundesagentur für Arbeit, Nürnberg. 2020. Available online: https://www.arbeitsagentur.de/datei/dok_ba013186.pdf (accessed on 20 December 2022).
7. Statistik der Bundesagentur für Arbeit. *Berichte: Blickpunkt Arbeitsmarkt MINT-Berufe*; Berichte: Blickpunkt Arbeitsmarkt, Bundesagentur für Arbeit, Nürnberg; 2019. Available online: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Statistiken/Themen-im-Fokus/Berufe/Generische-Publikationen/Broschuere-MINT.pdf?__blob=publicationFile#:~:text=Insgesamt%20waren%202018%20durchschnittlich%20259.000,(%2D8%2C7%20Prozent) (accessed on 20 December 2022).
8. Autorengruppe Offensive Digitale Schultransformation. Offensive Digitale Schultransformation: 7 Handlungsempfehlungen. *Gesellschaft für Informatik e. V., Germany.* 2020. Available online: https://offensive-digitale-schultransformation.de/ (accessed on 20 December 2022).
9. De Freitas, A.A.; Weingart, T.B. I'm Going to Learn What? Teaching Artificial Intelligence to Freshmen in an Introductory Computer Science Course. In Proceedings of the 52nd ACM Technical Symposium on Computer Science Education, Virtual, 13–20 March 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 198–204, ISBN 9781450380621.
10. Gadeib, A.; Noller, S. Wichtige Kompetenzen für eine lebenswerte Zukunft. In *INFOS 2021-19. GI-Fachtagung Informatik und Schule*; Humbert, L., Ed.; Gesellschaft für Informatik: Bonn, Germany, 2021; pp. 21–24.
11. OECD. OECD Future of Education and Skills. Available online: https://www.oecd.org/education/2030-project (accessed on 9 September 2022).
12. Bocconi, S.; Chioccariello, A.; Kampylis, P.; Dagienė, V.; Wastiau, P.; Engelhardt, K.; Earp, K.; Horvath, M.; Jasutė, M.; Malagoli, C.; et al. *Reviewing Computational Thinking in Compulsory Education: State of Play and Practices from Computing Education*; Publications Office of the European Union: Luxembourg, 2022; ISBN 978-92-76-47208-7.
13. Schwarz, R.; Hellmig, L.; Friedrich, S. (Eds.) *Eine Synopse zum Informatikunterricht in Deutschland im Jahr 2020*; Gesellschaft für Informatik: Bonn, Germany, 2021.

14. Ministerium für Kultus, Jugend und Sport Baden Württemberg. *Neues Profilfach "Informatik, Mathematik, Physik" (IMP) Startet zum Neuen Schuljahr an 56 Allgemeinbildenden Gymnasien*; Ministerium für Kultus, Jugend und Sport Baden Württemberg: Stuttgart, Germany, 2018.
15. Ministerium für Kultus, Jugend und Sport Baden Württemberg. Bildungsplan zum Profilfach Informatik, Mathematik, Physik (IMP) 2018. *Ministerium für Kultus, Jugend und Sport Baden Württemberg: Stuttgart, Germany*. 2018. Available online: http://bildungsplaene-bw.de/site/bildungsplan/get/documents/lsbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GYM_IMP.pdf (accessed on 20 December 2022).
16. Ministerium für Kultus, Jugend und Sport Baden Württemberg. Leitfaden für die gymnasiale Oberstufe: Abitur 2022. Available online: https://km-bw.de/site/pbs-bw-km-root/get/documents_E-763794728/KULTUS.Dachmandant/KULTUS/KM-Homepage/Publikationen%202019/20191113%20Leitfaden_Abitur_2022.pdf (accessed on 13 December 2022).
17. Staatsministerium Baden-Württemberg Pressestelle der Landesregierung. *Mehr Lehrerstellen im neuen Schuljahr*; Staatsministerium Baden-Württemberg: Stuttgart, Germany, 2020. Available online: https://www.baden-wuerttemberg.de/de/service/presse/pressemitteilung/pid/mehr-lehrerstellen-im-neuen-schuljahr-1 (accessed on 20 December 2022).
18. Eccles, J.S.; Wigfield, A. From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemp. Educ. Psychol.* **2020**, *61*, 101859. [CrossRef]
19. Ertl, B.; Luttenberger, S.; Paechter, M. Stereotype als Einflussfaktoren auf die Motivation und die Einschätzung der eigenen Fähigkeiten bei Studentinnen in MINT-Fächern. *Gr. Organ.* **2014**, *45*, 419–440. [CrossRef]
20. Götsch, M. "Das fängt natürlich an mit irgendwelchen Spielekonsolen"—Oder: Was dazu motiviert, Informatik (nicht) zu studieren. *Informatik-Spektrum* **2013**, *36*, 267–273. [CrossRef]
21. Ihsen, S.; Höhle, E.; Baldin, D. *Spurensuche!: Entscheidungskriterien für Natur- bzw. Ingenieurwissenschaften und mögliche Ursachen für frühe Studienabbrüche von Frauen und Männern an TU9-Universitäten*; LIT Verlag Münster: Münster, Germany, 2013.
22. Krapp, A. *Interesse, Lernen, Leistung: Neuere Ansätze der Pädagogisch-Psychologischen Interessenforschung*; Aschendorff: Münster, Germany, 1992; ISBN 3402045966.
23. Deci, E.L.; Ryan, R.M. Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Z. Pädagogik* **1993**, *39*, 223–238. [CrossRef]
24. Prenzel, M.; Drechsel, B. Ein Jahr kaufmännische Erstausbildung: Veränderungen in Lernmotivation und Interesse. *Unterrichtswissenschaft* **1996**, *24*, 217–234. [CrossRef]
25. Jansen, M.; Schroeders, U.; Lüdtke, O. Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learn. Individ. Differ.* **2014**, *30*, 11–21. [CrossRef]
26. Nagy, G.; Watt, H.M.G.; Eccles, J.S.; Trautwein, U.; Lüdtke, O.; Baumert, J. The Development of Students' Mathematics Self-Concept in Relation to Gender: Different Countries, Different Trajectories? *J. Res. Adolesc.* **2010**, *20*, 482–506. [CrossRef]
27. Hülsmann, C. *Kurswahlmotive im Fach Chemie: Eine Studie zum Wahlverhalten und Erfolg von Schülerinnen und Schülern in der gymnasialen Oberstufe*; Hochschulschrift: Berlin, Germany, 2015.
28. Hoffmann, L.; Häussler, P.; Lehrke, M. *Die IPN-Interessenstudie Physik*; IPN: Kiel, Germany, 1998; p. 158. ISBN 3-89088-114-9.
29. Krapp, A. Interesse, Lernen und Leistung. Neue Forschungsansätze in der Pädagogischen Psychologie. *Zeitschrift für Pädagogik* **1992**, *38*, 747–770. [CrossRef]
30. ECCLES, J. Who Am I and What Am I Going to Do with My Life? Personal and Collective Identities as Motivators of Action. *Educ. Psychol.* **2009**, *44*, 78–89. [CrossRef]
31. Mauk, V. *Einflussfaktoren der Studienwahl und des Studienverbleibs in MINT-Studienrichtungen an österreichischen Universitäten*; FB2 Biologie/Chemie, Universität Bremen: Bremen, Germany, 2016.
32. Baur, N.; Blasius, J. (Eds.) . *Handbuch Methoden der empirischen Sozialforschung*; Springer Fachmedien Wiesbaden: Wiesbaden, Germany, 2014.
33. Haselmeier, K.; Humbert, L.; Killich, K.; Müller, D. Interesse an Informatik und Informatikselbstkonzept zu Beginn der Sekundarstufe I des Gymnasiums. In *Informatik für Alle*; Pasternak, A., Ed.; Gesellschaft für Informatik: Bonn, Germany, 2019; pp. 99–108.
34. Holstermann, N.; Bögeholz, S. Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I. *Z. Für Didakt. Der Nat.* **2007**, *13*, 71–86.
35. Daniels, Z. *Entwicklung schulischer Interessen im Jugendalter*; Waxmann Verlag: Münster, Germany, 2008; ISBN 9783830970224.
36. Lazarides, R.; Ittel, A.; Juang, L. Wahrgenommene Unterrichtsgestaltung und Interesse im Fach Mathematik von Schülerinnen und Schülern. *Unterrichtswissenschaft* **2015**, *43*, 67–82.
37. Bergmann, A. *Mathematisch-Naturwissenschaftliches Fachinteresse durch Profilunterricht Fördern—Theoriebasierte Evaluation Eines Thüringer Schulversuchs in der Sekundarstufe I.*; Universitätsbibliothek Leipzig: Leipzig, Germany, 2020.
38. Krapp, A.; Prenzel, M. Research on Interest in Science: Theories, methods, and findings. *Int. J. Sci. Educ.* **2011**, *33*, 27–50. [CrossRef]
39. Mokhonko, S.; Nickolaus, R.; Windaus, A. Förderung von Mädchen in Naturwissenschaften: Schülerlabore und ihre Effekte. *Z. Für Didakt. Der Nat.* **2014**, *20*, 143–159. [CrossRef]
40. Diethelm, I.; Schneider, N.; Matzner, M.; Brückmann, M.; Zeising, A. Investigation of the Informatics-Based Self-Concept of Primary school Children. In Proceedings of the 15th Workshop on Primary and Secondary Computing Education, Virtual, 28–30 October 2020.

41. Leifheit, L. *The Role of Self-Concept and Motivation within the "Computational Thinking" Approach to Early Computer Science Education*; Universität Tübingen: Tübingen, Germany, 2021.
42. Förtsch, S.; Schmid, U. *The Influence of Academic Self-Concept on the Study Program Choice of Computer Scientists*; Verlag Barbara Budrich: Leverkusen, 2018; pp. 275–292. [CrossRef]
43. Janneck, M.; Vincent-Höper, S.; Ehrhardt, J. The Computer-Related Self Concept. *Int. J. Soc. Organ. Dyn. IT* **2013**, *3*, 1–16. [CrossRef]
44. Nagy, G.; Trautwein, U.; Baumert, J.; Köller, O.; Garrett, J. Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educ. Res. Eval.* **2006**, *12*, 323–345. [CrossRef]
45. Saß, S.; Kampa, N. Self-Concept Profiles in Lower Secondary Level—An Explanation for Gender Differences in Science Course Selection? *Front. Psychol.* **2019**, *10*, 836. [CrossRef] [PubMed]
46. Young, D.M.; Rudman, L.A.; Buettner, H.M.; McLean, M.C. The Influence of Female Role Models on Women's Implicit Science Cognitions. *Psychol. Women Q.* **2013**, *37*, 283–292. [CrossRef]
47. Shin, J.E.L.; Levy, S.R.; London, B. Effects of role model exposure on STEM and non-STEM student engagement. *J. Appl. Soc. Psychol.* **2016**, *46*, 410–427. [CrossRef]
48. Cheryan, S.; Drury, B.J.; Vichayapai, M. Enduring Influence of Stereotypical Computer Science Role Models on Women's Academic Aspirations. *Psychol. Women Q.* **2013**, *37*, 72–79. [CrossRef]
49. Boston, J.S.; Cimpian, A. How Do We Encourage Gifted Girls to Pursue and Succeed in Science and Engineering? *Gift. Child Today* **2018**, *41*, 196–207. [CrossRef]
50. Lawner, E.K.; Quinn, D.M.; Camacho, G.; Johnson, B.T.; Pan-Weisz, B. Ingroup role models and underrepresented students' performance and interest in STEM: A meta-analysis of lab and field studies. *Soc. Psychol. Educ.* **2019**, *22*, 1169–1195. [CrossRef]
51. Artelt, C.; Baumert, J.; Julius-McElvany, N.; Peschar, L.J. *Das Lernen Lernen: Voraussetzungen für Lebensbegleitendes Lernen*; Ergebnisse von PISA 2000; OECD: Paris, France, 2003.
52. Zinn, B.; Latzel, M. *Abschlussbericht zum Projekt "Evaluation des Schulversuchs NwT in den Jahrgangsstufen—Zweistündig (NwT-K2)"*; University of Stuttgart: Stuttgart, Germany, 2017. Available online: https://www.ife.uni-stuttgart.de/dokumente/bpt/bpt_forschung/nwtk2_Abschlussbericht-NwT-K2.pdf (accessed on 9 September 2022).
53. Waligora, K.; Schöne, C.; Dickhäuser, O.; Spinath, B.; Stiensmeier-Pelster, J. Skalen zur Erfassung des schulischen Selbstkonzepts (SESSKO). Göttingen: Hogrefe. *Prax. Der Kinderpsychol. Und Kinderpsychiatr.* **2003**, *52*, 465–467. [CrossRef]
54. Krapp, A.; Schiefele, U.; Wild, K.P.; Winteler, A. Der "Fragebogen zum Studieninteresse" (FSI). 1993. Available online: https://publishup.uni-potsdam.de/frontdoor/index/index/docId/3179 (accessed on 9 September 2022).
55. Eitemüller, C.; Walpuski, M. Wahl- und Abwahlprofile im Fach Chemie: Ergebnisse einer Clusteranalyse zur Charakterisierung von Lernenden am Ende der Sekundarstufe I. *Z. Für Didakt. Der Nat.* **2018**, *24*, 251–263. [CrossRef]
56. Taber, K.S. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Res. Sci. Educ.* **2018**, *48*, 1273–1296. [CrossRef]
57. Mayring, P. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*; 12., überarb. Aufl.; Beltz: Weinheim, Germany, 2015; ISBN 3407293933.
58. MacPhail, C.; Khoza, N.; Abler, L.; Ranganathan, M. Process guidelines for establishing Intercoder Reliability in qualitative studies. *Qual. Res.* **2016**, *16*, 198–212. [CrossRef]
59. Burla, L.; Knierim, B.; Barth, J.; Liewald, K.; Duetz, M.; Abel, T. From text to codings: Intercoder reliability assessment in qualitative content analysis. *Nurs. Res.* **2008**, *57*, 113–117. [CrossRef]
60. Fritz, C.O.; Morris, P.E.; Richler, J.J. Effect size estimates: Current use, calculations, and interpretation. *J. Exp. Psychol. Gen.* **2012**, *141*, 2–18. [CrossRef]
61. Tomczak, M.; Tomczak, E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends Sport Sci.* **2014**, *21*, 19–25.
62. Mann, H.B.; Whitney, D.R. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Statist.* **1947**, *18*, 50–60. [CrossRef]
63. Stonehouse, J.M.; Forrester, G.J. Robustness of the t and U tests under combined assumption violations. *J. Appl. Stat.* **1998**, *25*, 63–74. [CrossRef]
64. Kühnel, S.; Krebs, D. *Statistik für die Sozialwissenschaften: Grundlagen, Methoden, Anwendungen, Orig.-Ausg*; Rowohlt-Taschenbuch-Verl.: Reinbek bei Hamburg, Germany, 2001; ISBN 978-3-499-55639-5.
65. Dengel, A.; Heuer, U. Motivation, Fachinteresse und Schulleistung in Informatik. In *INFOS 2021-19. GI-Fachtagung Informatik und Schule*; Humbert, L., Ed.; Gesellschaft für Informatik: Bonn, Germany, 2021; pp. 113–122.
66. Heyder, A.; Weidinger, A.F.; Steinmayr, R. Only a Burden for Females in Math? Gender and Domain Differences in the Relation Between Adolescents' Fixed Mindsets and Motivation. *J. Youth Adolesc.* **2021**, *50*, 177–188. [CrossRef] [PubMed]
67. Watt, H.M.G. The role of motivation in gendered educational and occupational trajectories related to maths. *Educ. Res. Eval.* **2006**, *12*, 305–322. [CrossRef]
68. Barbara, S. Mehr Frauen in die Informatik: Den Gender-Gap in der Informatik schließen! Was zeigt Wirkung? In *INFORMATIK 2022*; Demmler, D., Krupka, D., Federrath, H., Eds.; Gesellschaft für Informatik: Hamburg, Germany, 28 September 2022.
69. Pearson, K.X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1900**, *50*, 157–175. [CrossRef]

70. Gao, Y. *Deconstructing the "Leaky Pipeline": Three Studies to Describe and Explain STEM Career Trajectories with Gender Differences*; UC Irvine: Irvine, CA, USA, 2022.
71. Nash, J. Understanding How to Interest Girls in Stem Education: A Look at How Lego® Education Ambassador Teachers Engage Female Students in Stem Learning. Ph.D. Thesis, University of Florida, Gainesville, FL, USA, 2017.
72. Main, J.B.; Schimpf, C. The Underrepresentation of Women in Computing Fields: A Synthesis of Literature Using a Life Course Perspective. *IEEE Trans. Educ.* **2017**, *60*, 296–304. [CrossRef]
73. Annis, B.; Nesbitt, R. *Results at the Top: Using Gender Intelligence to Create Breakthrough Growth*; 1. Auflage; John Wiley & Sons: New York, NY, USA, 2017; ISBN 1119384087.
74. García-Peñalvo, F.J.; Reimann, D.; Tuul, M.; Rees, A.; Jormanainen, I. *An Overview of the Most Relevant Literature on Coding and Computational Thinking with Emphasis on the Relevant Issues for Teachers*; TACCLE3 Consortium: Hasselt, Belgium, 2016. [CrossRef]
75. Labudde, P. Fächer übergreifender Unterricht in und mit Physik: Eine zu wenig genutzte Chance. *PhyDid A-Phys. Und Didakt. Sch. Und Hochsch.* **2003**, *1*, 48–66.
76. Seidel, T. Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma: Utilization-of-learning-opportunities models in the psychology of Instruction: Integration of the paradigms of structure and of process. *Z. Pädagogik* **2014**, *60*, 850–866.
77. Mackenzie, E.; Berger, N.; Holmes, K. Predicting adolescent girls' intentions to study science in senior high school. *Issues Educ. Res.* **2021**, *31*, 2021.
78. Häußler, P.; Hoffmann, L. Chancengleichheit für Mädchen im Physikunterricht—Ergebnisse eines erweiterten BLK-Modellversuchs. *Z. Didakt. Nat.* **1998**, *4*, 51–67.

education sciences

Article

Leveraging Cultural Wealth, Identities and Motivation: How Diverse Intersectional Groups of Low-Income Undergraduate STEM Students Persist in Collegiate STEM Environments

Raeshan D. Davis [1,*] and Zakiya Wilson-Kennedy [1,2]

[1] Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803, USA; zwilson@lsu.edu
[2] College of Science, Louisiana State University, Baton Rouge, LA 70803, USA
* Correspondence: raeshand@lsu.edu

Abstract: Grounded in a conceptual framework incorporating intersectionality, motivation, self-determination, and self-efficacy, this empirical study investigated how individuals' identities, mindsets, and resources in educational environments intentionally cultivated to support their decision-making, development, and connections in the science community, can significantly increase the recruitment, persistence, and success of low-income, academically talented science students from diverse backgrounds. Several factors—academic performance in coursework, self-image, self-agency, financial support, and social integration in the science culture—continue to significantly impact student retention and persistence in STEM disciplines. Many of these factors are negatively affected based on a students' intersecting identities, which can be detrimental to their academic success if not addressed. We found that additional considerations to factor in concerning low-income students from diverse backgrounds that is pertinent to supporting their persistence and success in the postsecondary STEM educational context.

Keywords: low-income students; intersectionality; motivation; self-efficacy; self-determination; social agency; higher education; STEM education

Citation: Davis, R.D.; Wilson-Kennedy, Z. Leveraging Cultural Wealth, Identities and Motivation: How Diverse Intersectional Groups of Low-Income Undergraduate STEM Students Persist in Collegiate STEM Environments. *Educ. Sci.* **2023**, *13*, 888. https://doi.org/10.3390/educsci13090888

Academic Editors: Erin Mackenzie, Bernhard Ertl, Christine R. Starr, Silke Luttenberger and Manuela Paechter

Received: 15 June 2023
Revised: 17 August 2023
Accepted: 28 August 2023
Published: 1 September 2023

1. Introduction

The national goal of increasing the proportion of students from diverse low-income families succeeding in science, technology, engineering, and mathematics (STEM) education and research enterprise remains paramount to advancing STEM fields in the United States. However, more than simply having the opportunity to engage in postsecondary STEM education is required. Students must consider the costs and the resources needed to enroll and persist in postsecondary education. According to The National Postsecondary Student Aid Study, there are many equity gaps in postsecondary education, including with regard to students from low-income households [1]. Many college aspirants, particularly individuals from low-income families, identified lack of sufficient financial resources as one of the most significant hurdles affecting their ability to pursue postsecondary education. Further, statistics show that approximately 43% of students from all racial and ethnic groups identified as low-income [1]. In particular, students identifying as Hispanic, Black, and American Indian or Alaska Native had the highest numbers identifying as low-income at 52.9, 54 and 59.8 percent, respectively. These jarring statistics affirmed by existing studies illuminate the truth that educational journeys and social mobility attainment can vary significantly by race, ethnicity, gender, and financial resources for various student groups [2–4]. Students from low-income backgrounds and who identify as Black or Hispanic are more likely to leave college without obtaining a credential compared to their White and Asian counterparts.

Layered with the context of socioeconomic status impacting educational attainment, the continued trend of underrepresentation of individuals from diverse racial and ethnic backgrounds in certain career fields continues to bring into focus the layered challenges impacting these groups. In particular, science, technology, engineering, and mathematics (STEM) disciplines show a continued trend of underrepresentation of women and individuals from diverse racial and ethnic backgrounds [5–7]. The 2023 Diversity and STEM report released by the National Center for Science and Engineering Statistics (NCSES) illuminated the disparities based on gender and race and ethnicity within the U.S. STEM workforce (see Figure 1) [8]. Certain racial and ethnic and gender groups are significantly underrepresented in specific STEM fields in science and engineering, while other groups are significantly overrepresented. For instance, women are significantly underrepresented in S&E occupations, which are typically jobs that require a bachelor's degree, in five major categories: (1) computer and mathematical scientists, (2) biological, agricultural, and environmental life scientists, (3) physical scientists, (4) social scientists, and (5) engineers. Women are overrepresented in S&E-related occupations, such as health care workers, S&E managers, S&E precollege teachers, technologists, and technicians, which are a subset of STEM occupations that require STEM skills and expertise, but are not classified in the five main S&E categories [8]. When investigating the statistics in Figure 1 through the lens of race and ethnicity, notably, white individuals represent over 60% of the S&E workforce in all occupation categories. In comparison, Black and Hispanic individuals represent between 8 and 10% of the S&E workforce, which is also a lower percentage than their representation in the general population. Such statistics and reports beg the question of why some groups are successful in some STEM occupations while others struggle to increase their representation. This concerning landscape assessment of science and engineering fields presents educational leaders with the task of evaluating the policies, values, pedagogies, and resource distribution in their institutional environments that continue to be inequitable against groups vulnerable to attrition.

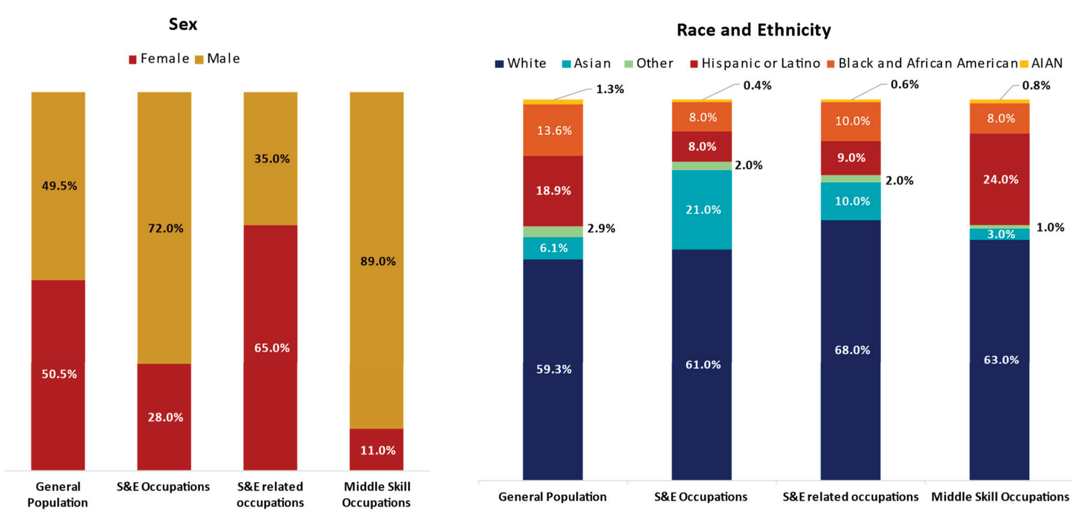


Figure 1. Characteristics of the U.S. STEM workforce ages 18–74, by occupation: 2021.

Earlier studies investigating the disparity between population groups and their participation in STEM disciplines identified several significant barriers to the participation of historically underrepresented groups in STEM fields, particularly a lack of a sense of belonging, a lack of science identity, and a lack of career pathways identified for the prescribed major [9–12]. However, a longstanding notion foundational to the discourse regarding the disparity between groups has perpetuated the view that the deficit depends solely on the student's abilities, without factoring in the role that STEM culture and environments can play in their access to resources and community. This traditional thinking and these attitudes towards understanding who participates in STEM and who exits STEM do not provide a complete picture of the educational experience of individuals from diverse backgrounds. Centering on the individual's knowledge and abilities without factoring in the systematic and structural powers at play severely underestimates the impact of the educational environment and the exclusionary undertones implied by traditional thinking and approaches to addressing these challenges [13,14].

In recent years, there has been an increase in empirical research studies and critical scholarship that has critiqued the traditional notions and dominant narratives of STEM culture. In particular, scholars have challenged the notion of the deficit solely lying with the individual and have provided substantial evidence that many of these challenges are perpetuated by systematic and structural policies, practices, and ideologies interwoven into the fabric of the STEM enterprise [15–19]. Further, scholars have posited that to broaden STEM participation for marginalized groups, we must critically disrupt the dominant STEM culture and narrative rooted in white and male dominance [20–22]. There are structural and systematic biases that preferentially encourage students from dominant groups and those who are well-resourced, i.e., in terms of finances, positions of power, and influence. Among the studies discussing the need for disrupting deficit discourses in STEM educational contexts, Castro has highlighted that many existing STEM recruitment and retention programs frame their targeted student population with deficit descriptions such as "underprepared" and "at risk" and typically focus solely on the student's deficit [15]. This study brings to light a critical, foundational step for educational equity, starting with challenging the language and labels we use to describe marginalized students. Like many others, this study shows how inequities are foundational and interwoven in educational contexts and must be redressed to support all students interested in pursuing a career in STEM.

Throughout the higher education literature centered on STEM disciplines, scholars have shown that science identity, sense of belonging, access to role models and mentors, social capital and agency, and other factors contribute to undergraduate success and the development of a science identity [9,23–29]. Within STEM disciplines, it is increasingly critical to acknowledge the importance of cultivating educational spaces that promote inclusivity for all individuals, particularly persons that have traditionally been excluded. These studies suggest that meaningful interventions can support the persistence of students with talent and the motivation to learn science. Scholars have urged application of the collective power of the STEM faculty and education discipline researchers to work together to advance the STEM enterprise. Within the discourse on collaboration and change for STEM education, Wilson-Kennedy underscores that "the cultivation of talent in the STEM community belongs to and is the responsibility of the community. . . we must collectively work together to address systemic barriers" [30]. To support this scholarship and sense of belonging and to broaden participation in STEM fields, the National Science Foundation and many other federal, state, and non-profit organizations have offered their support, resources, and advocacy for investigating and improving the experiences of vulnerable populations to broaden their participation in the STEM enterprise, specifically examining the impact of the STEM culture on their persistence and graduation outcomes to shed light on the structural and systemic inequities stunting the growth of these populations [31,32].

The Present Study

Our study is centered on understanding how diverse intersectional groups of undergraduate students pursuing science and mathematics majors utilize and leverage their identities, mindsets, skills, agency, and resources in STEM environments to establish connections to their scientific community that encourage them to remain in their STEM field. The central research question guiding this study is: How do diverse intersectional groups of low-income undergraduate STEM students describe their experience navigating and persisting in their postsecondary STEM environment?

2. Conceptual Framework

Fundamentally, this empirical study investigates how an individual's identities, motivation and self-efficacy can be supported by intentionally reimagining STEM postsecondary spaces to foster a sense of belonging and to support science identity development for low-income populations and marginalized and historically excluded backgrounds. The conceptual framework utilized to situate this study was adapted from integrating intersectionality theory and three strands of motivation theory (self-determination theory, self-efficacy, and intrinsic and extrinsic motivation) (See Figure 2). Utilizing these theoretical underpinnings strengthens our ability to investigate and understand how diverse groups of low-income students leverage their identities, skillsets, and resources to develop their science identity and social agency in the scientific community. Further, we couple these theories to depict the positive impact of high-impact educational practices when situating an individual's abilities in environments that are intentionally designed to foster and support their science identity and social agency development.

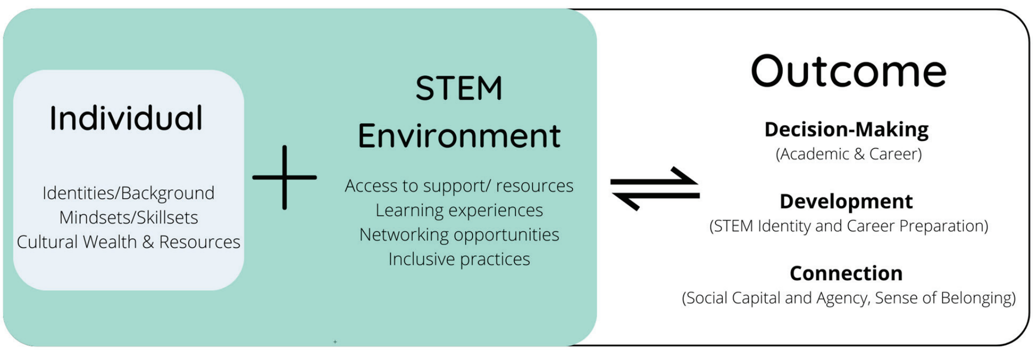


Figure 2. A Conceptual Model Integrating Individual and Environmental Conditions Necessary to Yield Desired Outcomes for Undergraduate Science Students.

The first element of the equation in the conceptual model focuses on the elements an individual possesses prior to entering the academic environment. In particular, this model highlights identities and background (demographic characteristics—race, ethnicity, gender, socioeconomic status, religious affiliation, first-generation status), mindsets and skillsets (motivation, self-efficacy, self-determination), and cultural wealth and resources. Salient to each individual is their unique background and identities, which, intertwined, shape their experiences—also known as intersectionality. Coined by Kimberly Crenshaw, mainly focusing on African American women, the term intersectionality was initially used to describe the combination of a person's social identities which overlap, interplay, and combine to produce different forms of oppression and empowerment [33,34]. Creshaw's work has since expanded to explore how individuals experience multiple oppressive forces based on their different, overlapping identities, such as class, race, gender, ability, sexual orientation, etc. This study of the interplay of overlapping identities enables us to better understand how to combat interwoven injustices in the human experience. In STEM education, intersectionality theory is used within research studies as a frame for understanding the complex lived experiences of individuals from marginalized groups in order to facilitate equity in education and the overall STEM enterprise [35–40].

Further, studies investigating power and oppression in social identities have shown that individuals from dominant identity groups seldom understand or are aware of the experiences of those individuals from marginalized identity groups [41–44]. Individuals in dominant groups can consciously or unconsciously invalidate the experiences of individuals from non-dominant groups due to a myriad of reasons, including inability to relate because of a lack of experience with oppression and lack of recognition of differences, among others. For instance, STEM fields such as engineering, physics, and the computer sciences have been known to be, and remain, traditionally white and male-dominated. Individuals from non-dominant groups, such as women and racial and ethnic marginalized groups, encounter numerous systematic and structural barriers when attempting to participate in such fields [45–48]. As such, our understanding of the human experience requires more thoughtful discussion and understanding of the complexities and intersections of power and oppression experienced by an individual, particularly in educational contexts that are often binary and traditional and do not always account for the nuanced nature of the human experience, which can severely discredit and inhibit their ability to persist.

In addition to the identities and backgrounds of individuals, this conceptual model brings into focus the mindsets and skillsets developed to support individuals' goals and aspirations. Our research utilizes three strands of motivation theory (self-determination theory, self-efficacy, and intrinsic and extrinsic motivation) as elements of our conceptual framework, particularly for its focus on understanding and centering the individual [28,49–52]. These elements provide a vital context for elevating the individual or student voice. Motivation in the dominant STEM culture is ascribed to individuals based on interpretations of persons in power. Our usage of motivation theory with cultural wealth disrupts traditional notions in disciplines and frames the student voice to highlight the goals, mindsets, and assets that students bring to their experiences in STEM.

Intrinsic and extrinsic motivation are established predictors of personal action [50,53–56]. Intrinsic motivation is doing an activity for its inherent satisfaction rather than for some separable consequence. As Deci and Ryan articulated, "when intrinsically motivated, a person is moved to act for the fun or challenge rather than because of external products, pressures, or rewards. . . in contrast, extrinsic motivation is a construct that pertains whenever an activity is done in order to attain some separable outcome" [53] (p. 56). Studies have shown that intrinsically motivated individuals are more likely to expend copious amounts of effort on an activity they enjoy, whereas extrinsically motivated individuals exert energy in a task or activity linearly according to what they perceive the return on investment will yield. While the perceived value gained differs, intrinsic and extrinsic motivation ultimately foster investment in the task at hand.

Self-determination theory (SDT) seeks to articulate the connection between motivation and one's ability to be self-determined. This broad framework presents a meta-theory of how social and cultural factors support or inhibit an individual's sense of initiative and volition. Existing research has posited that the highest quality forms of motivation and engagement are determined based on the individual's autonomy (the need to feel ownership of one's behavior), competence (the need to produce desired outcomes and to experience mastery), and relatedness (the need to feel connected to others) [50–52]. The optimal occurrence of these three conditions promotes an individual's creativity, persistence, and performance in an activity. For instance, research has shown that an individual's perceived competency in work-related activities increases their motivation and engagement [50,53].

Self-efficacy, a foundational element of theories of motivation, refers to an individual's belief in their own ability, knowledge, and skills to accomplish personal goals. Their self-efficacy is developed through personal performance accomplishments, vicarious learning, social influences, and physiological and affective states [54–56]. Studies have highlighted self-efficacy as a predictor of a student's motivation, engagement, and persistence in STEM scholarly activities [57–59].

The second element of the equation in the conceptual model is the collegiate and STEM environment context. In particular, the environment context highlights the role that learning experiences, opportunities to build community with peers, knowledge and access to resources, and inclusive STEM environments play in shaping an individual's science identity, sense of belonging and, ultimately, their persistence in their STEM discipline. The climate and culture of STEM disciplines play a significant role in students' sense of belonging and science identity development. Notably, the incorporation of recognized high-impact educational practices and learning experiences in STEM curricula and culture have been proven to aid in cultivating spaces that foster a sense of belonging and science identity [26,60–62].

In addition to understanding the optimal conditions of the educational environment, it is also important to understand the ways in which the individual and environment interplay together in a unique way to produce desired outcomes. Studies have shown that high-impact educational practices provide students with the opportunity to engage in a variety of ways to bolster their social connections, skills development, and self-efficacy, among other qualities [57,63–68]. Several additive educational practices can be incorporated to reimagine curricular and co-curricular experiences, such as living learning communities, communication training, summer intensive bridge experiences, and exposure to professionals in their specific fields, among others, that can promote the persistence of science students, particularly those from vulnerable groups [65,69,70].

We posit that an individual's background and motivations for STEM are supported and encouraged in STEM learning environments that foster a sense of belonging and science identity, so that there are expected outcomes and results affirming their persistence in STEM. In particular, our conceptual model highlights the outcomes achieved are decision-making (academic choices supporting the actualization of a career in the STEM field), development (science identity development, social agency development, research and technical skills development) and connection (within the institution and the science community). The outcomes achieved show how situating an individual's abilities, skills, and resources in environments intentionally cultivated to support their growth significantly affects their willingness and ability to persist in STEM.

2.1. Key Concepts

In addition to presenting the conceptual framework, this section will define several key terms fundamental to situating this study.

2.1.1. Community Cultural Wealth

We utilize the term community cultural wealth (CCW), which refers to the six assets that individuals from marginalized populations utilize to navigate educational environments [71]. Tara Yosso expanded on Bourdieu's work on cultural capital to provide a more inclusionary approach to understanding the cultural assets possessed by marginalized groups [71]. We chose to utilize the term community cultural wealth as our scholarship actively seeks to decenter ideologies that promote deficit approaches, maintain the status quo, and perpetuate inequities impacting individuals from marginalized populations.

2.1.2. Science Identity

As defined by Carlone and Johnson, science identity refers to the factors that affirm an individual's belief that they are a scientist. The three factors are their performance (engagement in relevant scientific practices), recognition (recognizing oneself and being recognized by others as a science person), and competence (knowledge and understanding of science content) [9].

2.1.3. Social Agency

Social agency is an individual's capacity to have the power and resources to fulfill their potential in their community [72]. In the context of this study, we posit that the combination of the individual in a supportive STEM environment can assist in cultivating their social agency in the scientific community. As a result, individuals experience a sense of ownership and the power to engage and navigate the science community successfully.

3. Methods

3.1. Participants and Research Setting

This study employed a qualitative case study approach to explore the formation of science identity and social agency among low-income, academically talented undergraduate science students from diverse backgrounds. Case study research allows researchers to comprehensively understand complex issues within their real-world context [73,74]. Case study research explores real-life, bounded cases through multiple forms of data. The setting of the research study was a large, research-intensive public university in the Deep South region of the United States. Within this college of science, approximately 1500 undergraduate science majors are enrolled across seven academic programs. About 16% of the undergraduates identify as members of a historically underrepresented racial and ethnic group.

The target population of this research study was science and math students who currently participate in or who participated in a scholarship program for students with significant academic talent and financial need from Fall 2020 to Spring 2023. The rationale for engaging this specific participant pool was the diversity in the group and the opportunity for access to collect interviews and other data sources for collection and triangulation. The term diverse group in this study is defined as individuals with differences in race, ethnicity, socioeconomic status, religious belief, sexual orientation, academic pursuits, and life experience, among others.

Table 1 summarizes the profiles and demographics of the participants in this research study. The participants included five women, four men, and one non-binary individual. The academic classifications of the group ranged from sophomores to graduating seniors. The racial makeup of the group was Asian, Black/African American, and White. All the participants identified as low to moderate income in their socioeconomic status. Four participants identified as first-generation students—neither parent had attended college or a university. The students actively participated in several high-impact educational practices and engagement opportunities that holistically supported and encouraged their academic and social integration into scientific and university culture.

Table 1. Participants' profiles.

Pseudonym	Classification	STEM Discipline	Gender	Race/Ethnicity	First Generation
Marie	Junior	Biological Sciences	Woman	Black or African American	N
William	Junior	Physics-Astronomy, Mechanical Engineering	Man	White	N
Alex	Junior	English Literature (previously Physics)	Non-binary	Two or more races (White and Asian)	Y
Irene	Junior	Biological Sciences-Marine Biology	Woman	Asian or Pacific Islander	N
Louis	Sophomore	Microbiology	Man	White	Y
Ruth	Sophomore	Microbiology	Woman	Black or African American	Y
Mary	Sophomore	Mathematical Statistics	Woman	White	N
Rachael	Sophomore	Biological Sciences (Pre-Med)	Woman	White	N
Malcolm	Senior	Biochemistry (Pre-Med)	Man	Black or African American	Y
Davis	Senior	Biological Sciences	Man	Asian	N

3.2. Data Collection and Analysis

Aligned with the case study methodology, we collected multiple forms of data, inclusive of documents and interviews [75]. We posit that the collective evaluation of these data points paints a detailed picture of our scholars' experiences, particularly giving a more in-depth understanding of how they developed and utilized their science identity and social agency in connection with their involvement in the mentoring scholarship program to integrate into their scientific community. The primary sources of data collected in this study were one-on-one, semi-structured interviews. The interviews lasted approximately 45 min and were conducted through the Zoom video platform. The interview protocol for this study can be found in the Supplementary Materials. Each participant provided verbal informed consent prior to participating in the research study. This study was approved by the Institutional Review Board at a research university in the Deep South region of the U.S. Participants were assigned pseudonyms to ensure their identity remained confidential and to provide a means of reference in the findings and discussion portion. The interview protocol questions for the research study were developed from the conceptual framework. In particular, the questions explored participants' transition to college, social agency development, science identity formation, participation in curricular and co-curricular experiences, and their perceptions of race and gender within the science community. The interview protocol questions are provided in the Supplementary Materials.

In addition to interviews, the researchers evaluated each participant's success metrics collected yearly. The success metrics included grade point average (G.P.A), professional development engagement, awards/honors, and individual development plan development and progress. This study's third source of data was annual evaluation reports provided by the external program evaluator. The annual evaluation report provided an overall summation of the program's efficacy in relation to outcomes, student and staff feedback and suggestions for areas of improvement. We posit that the collective review of these data points paints a detailed picture of our scholars' experiences, particularly giving a more in-depth understanding of how they developed and utilized their science identity and social agency in connection to their involvement in the mentoring scholarship program to integrate into their scientific community.

The data analysis process was grounded in an inductive coding strategy that allowed the researchers to identify patterns and assign codes to concepts of interest in the interview data [76,77]. The first step in this analysis process was reading through the transcripts to ensure accuracy, to correct grammatical errors, and to gain an initial understanding of each participant in the case study. Next, each transcript was uploaded into the Dedoose qualitative analysis software for coding. Each transcript was open-coded using an inductive strategy: reading and interpreting raw data to develop themes and concepts via interpretations based on data [74]. In this phase, the codes produced are tentative and subject to evolve and change as the analysis continues. After initial open-coding of all transcripts, the first iteration of the codebook was completed. In subsequent rounds of axial coding and refining the codebook, the final codebook was completed, which included 76 codes.

The reporting of the data analysis was two-fold. We conducted a narrative analysis to understand how participants interpreted their own lives through their identities as they navigated the academic and social environment in college, utilizing multiple data sources, including student metrics, evaluation reports and interview data, to deepen our understanding of each participant and the context of their environment [78,79]. Secondly, we examined the data across the group by identifying relationships and patterns across the ten participants. Utilizing these groupings of relationships and patterns, we labeled the emerging themes shared by the participants [80]. The final themes presented in the results section are the key outputs produced by the identified patterns or trends between the participants' experiences. We utilized peer debriefing sessions and concept mapping throughout the study development, implementation, and execution phases to ensure trustworthiness [81,82].

3.3. Researchers' Positionality

Our individual research agendas and collective interests focus on justice, equity, diversity, and inclusion (JEDI) systemic change, and education models in chemistry education. We investigate topics regarding faculty and student recruitment, retention, and success in STEM. The first author is a Black woman, a postdoctoral fellow in chemistry education at a public PWI. As a trained forensic scientist and qualitative educational researcher, her lived experiences as a student and professional have informed her scholarship and practice, which examines the narratives and lived experiences of historically excluded and marginalized populations, particularly investigating critical points in their transition to and navigation of the STEM educational pipeline. The second author is a faculty member in chemical education research and practice and an administrator within the College of Science at a PWI. As a leader on almost $30 million in extramural support from NSF, NIH, USDoEd, and philanthropic agencies, she has designed and implemented over 20 education projects, which have employed mentoring models to create and test development structures that cultivate self-efficacy and agency, particularly for groups historically underrepresented in STEM. Her research centers on studies of the persistence of individuals from all backgrounds in STEM higher education and careers, with a primary focus on faculty and student recruitment, retention, and success. Her lived experience as a woman of color in STEM influences her passion for studying students from diverse backgrounds and their pathways in STEM higher education and careers.

4. Results

The data analysis revealed three salient themes that emerged across the ten participants in relation to the central research question. We found that our participants discussed the ways in which their intersecting identities significantly shaped how they were navigating their collegiate STEM experience. Additionally, we found that our participants attributed their persistence in their STEM discipline to their motivation sources (intrinsic and extrinsic), support from their community (peers and mentors), and intentional engagement in academic activity and access to well-resourced environments (academic and social opportunities).

4.1. Intersecting Identities Significantly Shaping Their Collegiate Experience

As participants discussed navigating their educational experience, many of them discussed an awareness of the realities of the privilege and oppression of their intersecting identities as they moved through academic and social spaces. In particular, each participant described the most salient identities they were most conscious of as they navigated their educational spaces. In our discussions about the ways in which their identities were shaping their collegiate experience, Ruth, a Black woman and first-generation college student in her sophomore year of the microbiology program, shared her experience of developing awareness when she participated in a panel discussion during her freshman year about being Black at a predominantly white institution. She explains:

> *I was on a panel last year about being black and the PWI, and how you felt about that... after that panel I started seeing things I didn't really notice before... like the fact that an average class has about two hundred people and there is usually only maybe ten to fifty of us [Black people]... if there's ten of us here, we tend to group together based on that [race]... but then sometimes I think about this underlying thing which is in a job force only one of us can succeed.*

Ruth also discussed the dichotomy she experienced when she sought out a Black science-based organization:

> *I am a member of a Black science-based organization. I went to a couple of events, and it was really cool to see so many people in my major, who looked like me ... however, when I talked to my aunt, she said try not to get pigeon hole into just going to black organizations, because whenever you graduate, or you go into the workforce. Everyone's*

> *not going to be black and you want to have experiences outside of that [being Black], and not just be so uncomfortable in branching out because of that [race].*

Conversely, Malcolm, a Black man and first-generation college student in his senior year of the biochemistry program, shared how the representation of Black and Brown students has increased since his freshman year at this institution. For him, the growth in these specific student populations has become a form of motivation for his persistence.

> *From freshman year to now, there is a lot more color around campus…Sometimes I'd be the only black or minority person in a classroom… I would notice it, but it wouldn't be that big of an issue. Now, when I go into class, there's a bunch of other black and brown people, and it makes me feel good… even against odds and what anyone else thinks they're doing this too, which drives me. I don't want them to feel like they're alone in this, so I'm going to keep going too.*

In addition to the discussion about race, other participants focused on the dichotomy of their identities in academic and social spaces and how their academic spaces seemed neutral and erased their identities and human essence. Irene and Alex discussed how their experience navigating their identities was much more challenging socially than academically. Irene, a South Asian woman in her junior year of a biological sciences, marine biology concentration program, discussed how identifying as South Asian and Hindu was most apparent in the social settings of her collegiate experience. Particularly, she shared how she was negatively impacted by being in an institution with a dominant Christian student population:

> *There are a lot of people here identifying as Christian…so my religious identity and growing up Hindu has definitely impacted my social experience here… there are a lot of Christian-centered organizations and a lot of Christian-centered talk in some classes… it affected it a little negatively, just because I grew up in a very diverse community…Here everyone is either Catholic or Protestant and if you don't like fit that, it can be a little isolating… but in terms of feeling supported in my academic life, that's never affected me… people don't bring up religion in my Bio lab.*

While Irene appreciates the neutrality of her STEM academic space, Alex, a White and Asian non-binary person and first-generation college student in their junior year of the English literature program and formerly in the physics program, interpreted the neutral approach of their STEM academic environment as an erasure of their identities and put more of an emphasis on their ability to produce quality work as the determinant of their value. Alex shared: "*The physics department faculty were very supportive of my identity and my transition… however I always say being queer in STEM is a double-edged sword because, on the one hand, your identity doesn't really matter that much, and it just depends on the quality of the work that you put out…But, on the other hand, there isn't necessarily an explicit validation of identity…I think that both helped and hurt.*

Louis, a first-generation, White queer man with a speech disability in his sophomore year of the microbiology program, eloquently discussed the intersectionality of privilege and oppression experienced through his identities.

> *To start off with being first-gen. I don't feel like I'm good enough sometimes, and like I don't deserve a spot at college… I come from a disadvantage background, and it puts me in a spot where I feel like they have advantages. I'm going to kind of let them keep having those, and I'm going to sit back… As a queer person… I feel isolated sometimes because it's hard to judge someone's acceptance of me… I don't bring up certain things about myself, because I don't know how they'll react and it's a safety thing… which kind of leads me to not having any full interactions, because that's such a big part of me. And I'm a white man. I acknowledge that I'm a white man and the privilege that comes with that… I also have a speech impediment… It's kind of hard for me to communicate. I also hold back sometimes if someone's not taking the time to understand or if people are going to make fun of it.*

Conversely, students from dominant identity groups did not explicitly view their experience through their own identities of race, ethnicity, gender, and religion but rather through the lens of being low-income or being accepting of others' differences. Particularly, Davis, an Asian man in his senior year in the biological sciences program, shared his sentiments about the work ethic as the main factor influencing his educational experience:

> *I've interacted with a lot of people of different races, genders, genders, and like classes as me... I've learned to just respect people. For me, my race, class or gender do not limit me from what I want to be because it's about mindset, you know. I feel like it's about your grind. The harder you work, the more opportunities you have.*

Similarly, William, a White man in his junior year of physics and mechanical engineering programs, shared that his focus was not on identities but instead on connecting with peers over the shared student experience in the collegiate environment:

> *Every time I walk into an environment on campus, there's a wide variety of different types of races, genders and religions and a lot of different types of people. I think that's a great thing. I don't really feel I get that effect of there's too little or too much of this one specific group... it kind of blends to the point where I look past that stuff [identities]... I think we're all students trying to make it through our days so how can we help each other, or how can we share experiences, or get to know each other, and our differences, and all the stuff that makes us the same.*

Rachael and Mary, White women in their sophomore year in the biological sciences and mathematical statistics programs, respectively, discussed identities from the perspective of accepting others without acknowledging how they view their education journey through their own identities. For instance, when asked how her identities or perceptions of race, class, or gender have shaped her experience, Rachael said, *"I wouldn't say it has influenced it as much... I'll talk to anybody. It doesn't bother me. It's nothing wrong with your identity... everybody's unique."*

4.2. Motivation Powering Their Pursuits

In a discussion about their educational experience in their STEM academic program, participants explored and shared several topics and concepts rooted in actively remembering their motivation to pursue science. All the students discussed some degree of intrinsic motivation and self-reliance as the primary driving force for their direction and success academically. For all the students, their intrinsic motivation was cultivated from a belief in self, curiosity, and interest in the specific science discipline and achieving their definition of success. When asked to share their definition of success that motivated them to pursue their goals, three common themes emerged from their definitions of success: (a) accomplishing their goals; (b) being able to comfortably live the life that they envisioned for themselves; and (c) progression in developing their skills, abilities, and knowledge. Louis and several others brought into focus how their progression to becoming better was a demarcation of their success on their journey to their career. Specifically, Louis reflected on how he viewed his academic progression as a source of affirmation and motivation:

> *I reaffirm myself . . . that's why a lot of my validation comes from academics because there's an active scale . . . it's a double edge because sometimes, for example, I was struggling with organic chemistry, but for me, I'm succeeding because I understand more than I thought I would. In reality, I may not be my best, but I'm succeeding because I'm passing the class.*

In addition to their shared attributes of motivation and success, several participants discussed another layer to their motivation to pursue their career path: their commitment to serve and support their communities and society at large. Malcolm, Irene, and several others, stated that their motivations were rooted in their desire for civic responsibility to their communities. It is this connection to something greater than self that these participants found a source of fuel to their motivation to pursue their science degrees. Malcolm shared that his motivation was rooted in creating access for individuals from underserved and economically disadvantaged backgrounds and encouraged him to be persistent in his pursuit of becoming a neurosurgeon with experience in hospital administration.

> *I want to create more access for folks... people shouldn't have to risk the quality of their life even more so getting surgery... I definitely do want to go into the policy and procedures... I find there is a large disconnect between the admin and the people that are actually practicing medicine when the admin has not been in the doctor's shoes, and they're trying to implement things that they think could work instead of things that they've seen in practice... I want to limit the need for spinal surgery to begin with, and then, if someone does need spinal surgery, I want to develop methods that are going to reduce the risk associated with it.*

Similar to Malcolm, Irene conveyed that one of her motivations to pursue a career in environmental science was being able to connect science to society through educational resources. In particular, she discussed her love for documentaries and the ways in which documentaries are used to educate the general population about complex science matters. "*Film (documentaries) is what made me want to go into environmental science... it was seeing really good documentaries or seeing TV personalities like Steve Erwin... That's what I think connected me to nature. When documentaries like plastic ocean came out... these environmental documentaries really attracted me. I've always really liked what they did in terms of public outreach about bringing the science to large-scale audiences.*"

For others, their motivation to persist in their STEM programs was fueled by an initial curiosity in their specific discipline and motivated their initial pursuit, while the opportunities, freedom, and encouragement to explore and refine what their career path could be fueled their motivation and enthusiasm to continue to persist and navigate their educational experience. William, Louis, and Ruth all discussed the ways in which their persistence in their specific discipline had been sparked by their initial curiosity and sustained by their ability to refine their path to their own niche. For instance, William's childhood love for sci-fi movies piqued his interest in physics. Once he came to start his education journey, he added an engineering major to allow him to bring his theoretical physics training to life through an engineering design background. He explained, "*When I was a lot younger, I would always watch the sci-fi movies... I set my path initially for just purely physics, but then I started realizing that I wanted to go into a field where I could do more design and try to apply these concepts... realized that in order to get that design background, you would ideally, you'd want to have an engineering background*".

4.3. Community and Connection Matters

A notable commonality among all participants was their active engagement in various academic and social opportunities that supported their sense of belonging in the science and university community. As the students shared about their engagement outside of classes, they discussed participating in at least one student organization and one academic experience (i.e., conferences, undergraduate research, shadowing, etc.). Table 2 represents data from the student metrics collected on their participation in undergraduate research, academic organizations, social organizations, and academic co-curricular experiences for the ten students in this study. Utilizing a ranking scale from the activity with the highest participation to the lowest participation, science-related academic organizations, undergraduate research, and social organizations were the top three engagement activities in which most participants engaged. Notably, all the participants reported participating in at least one science-related academic organization.

Table 2. Engagement in Specific Academic and Social Activities.

Engagement Activity	Total Participation	Percentage
STEM-Related Academic Organizations	10	100%
Undergraduate Research	7	70%
Social Organizations	7	70%
STEM Conference Participation	6	60%
STEM Conference Presentations	5	50%
Shadowing/Internships	3	30%

When asked about their engagement experiences, many participants discussed how their participation in various engagement activities had supported their personal and professional growth, provided a community of like-minded peers, and affirmed their sense of belonging in the science community. Davis, Mary, and Rachael discussed at length the positive impact that their engagement outside of the classroom had in supporting them. For Davis, his engagement in academic and social opportunities had been the linchpin in his persistence in his discipline. He expounded, "*These experiences give me like a buffer between science classes because I feel sometimes science can be a little bit overwhelming no matter how interested you are in the subject. . . It gives me an opportunity to destress and come back better than ever, ready to focus, ready to take on science.*" Further, when discussing his social connections, he shared his appreciation for his peers, "*It's so necessary for the college experience. I don't think I would have made it. . . The people in the program are already accomplished people. I talk to them about stuff, and conversational pieces are in the back of my head. . . I wanted to be around like-minded people and to be able to engage in that type of space*".

Davis, Rachael, and Mary shared how engaging with peers in social settings had been one of the highlights of their collegiate experience to help them balance the rigor of being a science major. Rachael explained, "*. . . doing fun activities definitely helps lighten the load and makes this experience feel a little better. For instance, we went bowling, and I had an exam the next day, but I like planning ahead, and I was like, I'm still going to this event because It's going to be a nice relaxation before having to go take an exam.*" Mary shared how her engagement with her peers had served as extrinsic motivation to persist in STEM. "*Having people doing similar things like me or making sure that I'm staying on track has definitely motivated me to keep working at this degree and to keep up with my other friends in STEM. Also, to not let people down, you know. . . Its motivation being around similar like-minded people*".

In addition to being in a community with like-minded peers, students who participated in undergraduate research discussed the ways in which their ability to understand the course content and growth in their research skills and academic decision-making were significant benefits of this opportunity. For instance, William participated in a summer undergraduate research experience, the summer before his sophomore year, and shared: "*I did my research after my freshman year. . .everything I had to learn in the research project over that summer was stuff that I was going to be learning in courses down the road. . . I'm still coming across content I learned in that research experience that's coming up in the courses I'm taking now. . . it's extremely helpful getting involved in research.*" For Irene, her undergraduate research experience had strengthened her research skills and supported her career plans and actualization in her STEM field.

> *Everything I've done outside of the classroom has been pushing me in this direction that I think I've always been headed, which is like just wanting to do something either service-oriented, or for climate change, or something that feels urgent. . . Everything is set to give me skills that I will apply in the future. . . It's really been good to know that I can see how this was applied in the lab.*

5. Discussion

The purpose of this present study was to explore the ways in which low-income undergraduate science students from diverse backgrounds describe navigating their post-secondary educational experience, particularly the contexts, skills, and supports that have encouraged their persistence to remain in their desired STEM field. Particularly for our ten participants in this study, their educational journey was investigated with a focus on their intersecting identities and how these intersecting identities shape how they connect with peers, access resources, engage in activities, develop skills, and actualize their place within the science community. Considering the experiences shared by the ten students, several streams of understanding emerged to answer the central research question guiding this study.

Foundational in this study, our participants shared the identity of being classified as low-income in terms of socioeconomic status. Through conversations with the students, we found that many of them shared similar work-ethic philosophies that their hard work can pay off and their goals would be met. Their sentiments affirmed the common societal belief that individuals from low-income backgrounds can achieve upward economic and social mobility with their hard work [2–4,83]. While this belief may be a common assumption, research has shown that individuals from low-income backgrounds face and consistently contend with prolonged challenges that impact their ability to achieve economic stability, particularly those from marginalized racial and ethnic groups.

The existing literature asserts that the dominant culture and narratives in STEM disciplines prefer individuals who identify as either white, Asian, male, cisgender, or a combination [20–22]. As such, individuals from these identity backgrounds are more likely to achieve their desired goals and careers in STEM. Conversely, those individuals departing STEM disciplines are the students from non-dominant groups in STEM. In this present study on low-income science students, we, too, found evidence of the dominant culture and narratives in several of our participants. Also, we noted stark differences in the participants' navigation of their academic environments through the lens of race, ethnicity, gender, and religion, and how the intersection of these identities shaped their ability to navigate academic and social spaces and actively shaped their development, connection, and persistence in their STEM discipline. For instance, the Black women in this study discussed how operating in academic spaces with their Black identity at the forefront produced feelings of inadequacy, lack of belonging, and unequal performance expectations. The White women in this study did not even acknowledge their whiteness as a salient identity when navigating their educational spaces and often took the approach of blindness to diversity when discussing and engaging with diverse groups of people in their academic and social settings. Aligned with existing studies on student perceptions of their identities in the science community and specific disciplines, we know that students from historically underrepresented groups can experience feelings of inferiority, tokenism, unequal performance expectations, and not belonging at higher rates than their counterparts [62–66]. As students navigate their educational experience through the lens of their intersectional identities, they often navigate oppressive systems, policies, and practices that perpetuate inequities and barriers for marginalized groups. As such, it is imperative to account for nuanced considerations and, in some cases, the layering effect of their intersectional identities, when supporting students in STEM disciplines.

Regarding motivation powering their pursuits, Deci and Ryan suggested that motivation for one's goals should be self-directed and grounded in one's reasoning as the most satisfying and successful approach to success and accomplishment [53]. Aligned with existing studies and theoretical frameworks centered on motivation, self-efficacy, and self-determination [51,53,58,59,84], our participants affirmed the importance of one's motivation and self-efficacy to pursue and actualize their educational and career goals regardless of the challenges one encountered. Several participants discussed extrinsic motivation connected to civic responsibility—a calling more significant than themselves—usually to their communities, families, and the betterment of the human experience. In line with extrinsic motivation, our participants' self-defined success goals and additional external elements encouraged and supported their success to date and, in the long-term, influenced their desired end goal of success [50]. Motivation in this context not only supported the students in focusing their attention on their personal goals but encouraged several of them through the inspiration of their communities. These varied sources of intrinsic and extrinsic motivation were notable factors that supported and encouraged their persistence in STEM.

This study also brings to the forefront how consequential inclusive, affirming, and welcoming environments are in supporting the persistence of low-income, diverse student groups in STEM. The existing literature clearly articulates the importance of connection to peers and faculty and how engagement in academic and social activities outside of classes is valuable for developing their social agency and strengthening their relationship with the science community [16,19,85–88]. When exploring the impact of peer support and motivation, many of the participants discussed the value and benefits of being in a community with peers with similar goals and pursuits. Many of the students in this study discussed the benefits of having financial support for tuition and engagement opportunities and how access to such opportunities was critical to developing their social and professional networks that enriched their educational experience. The students highlighted that being in a community of peers with similar goals provided a source of motivation, increased their sense of belonging, and expanded their access to social networks, opportunities, and resources. Their participation in various engagement activities also supported their career and academic decision-making, career preparation, research skills development, and sense of belonging.

Across our participants' extracurricular engagement experiences, we noted that several of them discussed their engagement in STEM academic organizations connected to one or more of their salient identities. The desire to be in identity-centered educational spaces was four-fold: representation of professionals in their field, expansion of their social network, supporting their academic and career decision-making outcomes, and being in a space that affirmed an identity or intersecting identities seen as a barrier in the larger educational setting. For instance, the National Organization for the Professional Advancement of Black Chemists and Chemical Engineers (NOBCChE) is an organization with the expressed purpose of supporting the development and training of Black scientists in academic, development, leadership, and philanthropic endeavors throughout college, and careers in chemistry and chemical engineering. Our findings further affirm the importance for STEM environments to be intentionally cultivated to foster a sense of belonging and to promote inclusivity for students from diverse backgrounds, linking increased engagement in the STEM curricular and extracurricular experience to increased persistence in the STEM discipline [11,12,67,89,90].

Limitations

The participants' sensemaking of intersectionality varied significantly across the group. While most of them could list their salient identities, we recognize that our participants varied in their understanding of how their identities shaped the navigation of their STEM academic experience. As such, including a secondary interview in the research design would be beneficial to investigate the concept of intersectionality and to support our participants' sensemaking process more comprehensively.

6. Conclusions

Our study explored how diverse low-income students leverage their identities, skillsets, and resources within supportive STEM academic environments. Students enter postsecondary educational spaces with unique identities and backgrounds, mindsets and skillsets, cultural wealth, and resources. Our research shows the importance of the relationship between low-income diverse student groups and their academic environment for their persistence in their STEM discipline.

Looking deeper at low-income populations, we acknowledge that it is not a "one size fits all" approach. Much of the rhetoric surrounding low-income populations focuses on their work ethic as the sole determinant of success. However, our participants' interviews and narratives illustrate clearly that additional considerations must be considered when supporting the success of diverse low-income students, particularly regarding how their intersectional identities, skillsets, and resources are acknowledged and engaged in the collegiate environment. Their stories illuminate that their persistence in their STEM discipline is primarily attributed to their motivation sources (intrinsic and extrinsic), support from their community (peers and mentors), and intentional engagement in academic activities and access to well-resourced environments (academic and social opportunities). Thus, one implication of this study is for leaders and professionals supporting diverse students in STEM to actively consider how they are accounting for and factoring in students' identities, skillsets, and resources to best support them in developing a sustainable STEM ecosystem to encourage their persistence in STEM collegiate environments through to the STEM workforce.

As scholars, practitioners, and leaders in the STEM enterprise addressing the call for broadening participation, we encourage consideration of the realities of low-income students and other marginalized groups when establishing initiatives and engagement opportunities. For instance, participation in supplemental science-based activities has been long touted as beneficial to strengthen credentials and connection to the science community at little to no cost. However, we present two critical considerations for low-income populations. The first is that several immersive opportunities, such as conference participation and international undergraduate research experiences, are typically available to students with the time and financial resources to participate. The second is that many low-income students support themselves financially through college and, therefore, do not have the time to engage in supplemental activities. With this context in mind, we implore STEM program leaders to consider how their initiatives can be equity-minded. An example is the provision of dedicated funding sources for economically disadvantaged students to support their engagement in professional development and science-based activities. This approach removes the financial burden and potential stigma low-income students may experience regarding finances and encourages their involvement in learning experiences similar to their well-resourced counterparts.

Finally, our research encourages using high-impact educational practices adapted to account for the diverse student populations our institutions serve. In addition to the guidance on high-impact educational practices, we also posit that consideration can be given to program operation, logistics, cultural competence, and accessibility, among others. For instance, undergraduate research experiences are considered a high-impact educational practice. When designing such experiences for students with marginalized identities, we offer the following considerations as examples: paying undergraduate researchers, provision of culturally relevant training approaches, and connection to a mentor. It is incumbent that educators and leaders consider how we can intentionally and thoughtfully engage students who have the desire and willingness to pursue their desired STEM career path.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci13090888/s1.

Author Contributions: Conceptualization, R.D.D. and Z.W.-K.; methodology, R.D.D. and Z.W.-K.; investigation, R.D.D. and Z.W.-K.; resources, R.D.D. and Z.W.-K.; data curation, R.D.D.; writing—original draft preparation, R.D.D.; writing—review and editing, R.D.D. and Z.W.-K.; visualization, R.D.D. and Z.W.-K.; supervision, Z.W.-K.; project administration, R.D.D. and Z.W.-K.; funding acquisition, Z.W.-K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Science Foundation, grant number 1930474.

Institutional Review Board Statement: This study was approved by the Institutional Review Board at Louisiana State University, IR-BAM-22-0129.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data and analysis codes for this study are available by emailing the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Taylor, M.; Turk, J. *Race and Ethnicity in Higher Education: A Look at Low-Income Undergraduates*; American Council on Education: Washington, DC, USA, 2019.
2. Bastedo, M.N.; Jaquette, O. Running in place: Low-income students and the dynamics of higher education stratification. *Educ. Eval. Policy Anal.* **2011**, *33*, 318–339. [CrossRef]
3. Berg, G.A. *Low-Income Students and the Perpetuation of Inequality: Higher Education in America*; Routledge: London, UK, 2016.
4. Jack, A.A. The privileged poor. In *The Privileged Poor*; Harvard University Press: Cambridge, MA, USA, 2019.
5. Alfred, M.V.; Ray, S.M.; Johnson, M.A. Advancing women of color in STEM: An imperative for US global competitiveness. *Adv. Dev. Hum. Resour.* **2019**, *21*, 114–132. [CrossRef]
6. Betancur, L.; Votruba-Drzal, E.; Schunn, C. Socioeconomic gaps in science achievement. *Int. J. STEM Educ.* **2018**, *5*, 38. [CrossRef]
7. Blackburn, H.; Heppler, J. Women in STEM in Higher Education: A Citation Analysis of the Current Literature. *Sci. Technol. Libr.* **2019**, *38*, 261–271. [CrossRef]
8. National Center for Science and Engineering Statistics (NCSES). *Diversity and STEM: Women, Minorities, and Persons with Disabilities 2023*; Special Report NSF 23-315; National Science Foundation: Alexandria, VA, USA, 2023.
9. Carlone, H.B.; Johnson, A. Understanding the science experiences of successful women of color: Science identity as an analytic lens. *J. Res. Sci. Teach.* **2007**, *44*, 1187–1218. [CrossRef]
10. Graham, M.J.; Frederick, J.; Byars-Winston, A.; Hunter, A.-B.; Handelsman, J. Increasing persistence of college students in STEM. *Science* **2013**, *341*, 1455–1456. [CrossRef]
11. O'Keeffe, P. A sense of belonging: Improving student retention. *Coll. Stud. J.* **2013**, *47*, 605–613.
12. Rainey, K.; Dancy, M.; Mickelson, R.; Stearns, E.; Moller, S. Race and gender differences in how sense of belonging influences decisions to major in STEM. *Int. J. STEM Educ.* **2018**, *5*, 10. [CrossRef]
13. Hatfield, N.; Brown, N.; Topaz, C.M. Do introductory courses disproportionately drive minoritized students out of STEM pathways? *PNAS Nexus* **2022**, *1*, pgac167. [CrossRef]
14. Asai, D.J.; Bauerle, C. From HHMI: Doubling Down on Diversity. *CBE-Life Sci. Educ.* **2016**, *15*, fe6. [CrossRef]
15. Castro, E.L. "Underprepared" and "at-risk": Disrupting deficit discourses in undergraduate STEM recruitment and retention programming. *J. Stud. Aff. Res. Pract.* **2014**, *51*, 407–419. [CrossRef]
16. Martin, J.P.; Stefl, S.K.; Cain, L.W.; Pfirman, A.L. Understanding first-generation undergraduate engineering students' entry and persistence through social capital theory. *Int. J. STEM Educ.* **2020**, *7*, 37. [CrossRef]
17. Ortiz, N.A.; Morton, T.R.; Miles, M.L.; Roby, R.S. What about us? Exploring the challenges and sources of support influencing black students' STEM identity development in postsecondary education. *J. Negro Educ.* **2019**, *88*, 311–326. [CrossRef]
18. Peck, F. Towards anti-deficit education in undergraduate mathematics education: How deficit perspectives work to structure inequality and what can be done about it. *Primus* **2021**, *31*, 940–961. [CrossRef]
19. Rincón, B.E.; Rodriguez, S. Latinx students charting their own STEM pathways: How community cultural wealth informs their STEM identities. *J. Hisp. High. Educ.* **2021**, *20*, 149–163. [CrossRef]
20. Barton, A.C.; Menezes, S.; Mayas, R.; Ambrogio, O.; Ballard, M. What Are the Cultural Norms of STEM and Why Do They Matter. Center for Advancement of Informal Science Education. 2018. Available online: https://resources.informalscience.org//what-are-cultural-norms-stem-and-why-do-they-matter (accessed on 1 June 2023).
21. Kanny, M.A.; Sax, L.J.; Riggers-Piehl, T.A. Investigating forty years of STEM research: How explanations for the gender gap have evolved over time. *J. Women Minor. Sci. Eng.* **2014**, *20*, 127–148. [CrossRef]
22. Watkins, J. "That is Still STEM": Appropriating the Engineering Design Process to Challenge Dominant Narratives of Engineering and STEM. *Cogn. Instr.* **2023**, *41*, 1–31. [CrossRef]
23. Campbell, C.M.; Smith, M.; Dugan, J.P.; Komives, S.R. Mentors and college student leadership outcomes: The importance of position and process. *Rev. High. Educ.* **2012**, *35*, 595–625. [CrossRef]
24. Killpack, T.L.; Melón, L.C. Toward inclusive STEM classrooms: What personal role do faculty play? *CBE—Life Sci. Educ.* **2016**, *15*, es3. [CrossRef]
25. Metzner, B.S. Perceived quality of academic advising: The effect on freshman attrition. *Am. Educ. Res. J.* **1989**, *26*, 422–442. [CrossRef]
26. Peters, A.W.; Tisdale, V.A.; Swinton, D.J. High-impact educational practices that promote student achievement in STEM. In *Broadening Participation in STEM*; Emerald Publishing Limited: Leeds, UK, 2019.
27. Robinson, K.A.; Perez, T.; Carmel, J.H.; Linnenbrink-Garcia, L. Science identity development trajectories in a gateway college chemistry course: Predictors and relations to achievement and STEM pursuit. *Contemp. Educ. Psychol.* **2019**, *56*, 180–192. [CrossRef]
28. Trujillo, G.; Tanner, K.D. Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity. *CBE—Life Sci. Educ.* **2014**, *13*, 6–15. [CrossRef] [PubMed]
29. Strayhorn, T.L. Bridging the pipeline: Increasing underrepresented students' preparation for college through a summer bridge program. *Am. Behav. Sci.* **2011**, *55*, 142–159. [CrossRef]
30. Wilson-Kennedy, Z. *Omwana Ni Wa Bhone . . . It Takes a Village*; American Association for the Advancement of Science: Washington, DC, USA, 2023.

31. National Center for Science and Engineering Statistics. *Women, Minorities, and Persons with Disabilities in Science and Engineering*; National Center for Science and Engineering Statistics: Alexandria, VA, USA, 2017.
32. National Science Board. *Higher Education in Science and Engineering*; National Science Board: Alexandria, VA, USA, 2019.
33. Crenshaw, K. Beyond racism and misogyny: Black feminism and 2 Live Crew. In *Feminist Social Thought: A Reader*; Meyer, D.T., Ed.; Routledge: New York, NY, USA, 1997; pp. 245–263.
34. Crenshaw, K. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. *Univ. Chic. Leg. Forum* **1989**, *140*, 139–167.
35. Cochran, G.L.; Boveda, M. A framework for improving diversity work in physics. In Proceedings of the Physics Education Research Conference 2020, Virtual Conference, 22–23 July 2020; p. 2021.
36. Cochran, G.L.; Boveda, M.; Prescod-Weinstein, C. Intersectionality in STEM education research. In *Handbook of Research on STEM Education*; Routledge: London, UK, 2020; pp. 257–266.
37. Charleston, L.J.; Adserias, R.P.; Lang, N.M.; Jackson, J.F. Intersectionality and STEM: The role of race and gender in the academic pursuits of African American women in STEM. *J. Progress. Policy Pract.* **2014**, *2*, 273–293.
38. Armstrong, M.A.; Jovanovic, J. The intersectional matrix: Rethinking institutional change for URM women in STEM. *J. Divers. High. Educ.* **2017**, *10*, 216. [CrossRef]
39. Ibourk, A.; Hughes, R.; Mathis, C. "It is what it is": Using Storied-Identity and intersectionality lenses to understand the trajectory of a young Black woman's science and math identities. *J. Res. Sci. Teach.* **2022**, *59*, 1099–1133. [CrossRef]
40. Sparks, D.M.; Przymus, S.D.; Silveus, A.; De La Fuente, Y.; Cartmill, C. Navigating the intersectionality of race/ethnicity, culture, and gender identity as an aspiring Latina STEM student. *J. Lat. Educ.* **2021**, *22*, 1355–1371. [CrossRef]
41. Kozlowski, D.; Larivière, V.; Sugimoto, C.R.; Monroe-White, T. Intersectional inequalities in science. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2113067119. [CrossRef]
42. Bhopal, K. Confronting White privilege: The importance of intersectionality in the sociology of education. *Br. J. Sociol. Educ.* **2020**, *41*, 807–816. [CrossRef]
43. DeMarrais, E.; Castillo, L.J.; Earle, T. Ideology, materialization, and power strategies. *Curr. Anthropol.* **1996**, *37*, 15–31. [CrossRef]
44. Allen, B.J. *Difference Matters: Communicating Social Identity*; Waveland Press: Long Grove, IL, USA, 2023.
45. Mim, S.A. Women Missing in STEM Careers: A Critical Review through the Gender Lens. *J. Res. Sci. Math. Technol. Educ.* **2019**, *2*, 59–70. [CrossRef]
46. Harris, C.M. Quitting science: Factors that influence exit from the STEM workforce. *J. Women Minor. Sci. Eng.* **2019**, *25*, 93–118. [CrossRef]
47. Cheryan, S.; Ziegler, S.A.; Montoya, A.K.; Jiang, L. Why are some STEM fields more gender balanced than others? *Psychol. Bull.* **2017**, *143*, 1–35. [CrossRef]
48. Kricorian, K.; Seu, M.; Lopez, D.; Ureta, E.; Equils, O. Factors influencing participation of underrepresented students in STEM fields: Matched mentors and mindsets. *Int. J. STEM Educ.* **2020**, *7*, 16. [CrossRef]
49. Abbott, L.E.; Andes, A.; Pattani, A.C.; Mabrouk, P.A. Authorship Not Taught and Not Caught in Undergraduate Research Experiences at a Research University. *Sci. Eng. Ethics* **2020**, *26*, 2555–2599. [CrossRef]
50. Ryan, R.M.; Deci, E.L. Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemp. Educ. Psychol.* **2020**, *61*, 101860. [CrossRef]
51. Deci, E.L.; Ryan, R.M. Self-determination theory. In *Handbook of Theories of Social Psychology*; Sage Publications: Thousand Oaks, CA, USA, 2012.
52. Deci, E.L.; Olafsen, A.H.; Ryan, R.M. Self-determination theory in work organizations: The state of a science. *Annu. Rev. Organ. Psychol. Organ. Behav.* **2017**, *4*, 19–43. [CrossRef]
53. Ryan, R.M.; Deci, E.L. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemp. Educ. Psychol.* **2000**, *25*, 54–67. [CrossRef]
54. Bandura, A. *Self-Efficacy: The Exercise of Control*; Worth Publishers: New York, NY, USA, 1997.
55. Bandura, A. *Social Foundations of Thought and Action: A Social Cognitive Theory*; Prentice Hall: Englewood Cliffs, NJ, USA, 1986.
56. Lent, R.W.; Brown, S.D. Social cognitive model of career self-management: Toward a unifying view of adaptive career behavior across the life span. *J. Couns. Psychol.* **2013**, *60*, 557. [CrossRef]
57. McCoy, W. Black Girls Accepting the Grand Challenge: A Qualitative Exploration of a Summer Engineering Program's Influence on Black Girls' Racial Identity, Engineering Identity, and STEM Self-Efficacy. Ph.D. Thesis, North Carolina State University, Raleigh, NC, USA, 2020.
58. Tellhed, U.; Bäckström, M.; Björklund, F. Will I Fit in and Do Well? The Importance of Social Belongingness and Self-Efficacy for Explaining Gender Differences in Interest in STEM and HEED Majors. *Sex Roles* **2017**, *77*, 86–96. [CrossRef] [PubMed]
59. Macphee, D.; Farro, S.; Canetto, S.S. Academic self-efficacy and performance of underrepresented STEM majors: Gender, ethnic, and social class patterns. *Anal. Soc. Issues Public Policy* **2013**, *13*, 347–369. [CrossRef]
60. Brownell, J.E.; Swaner, L.E. *Five High-Impact Practices: Research on Learning Outcomes, Completion and Quality*; Association of American Colleges and Universities: Washington, DC, USA, 2010.
61. Carter, D.F.; Ro, H.K.; Alcott, B.; Lattuca, L.R. Co-Curricular Connections: The Role of Undergraduate Research Experiences in Promoting Engineering Students' Communication, Teamwork, and Leadership Skills. *Res. High. Educ.* **2016**, *57*, 363–393. [CrossRef]

62. Kuh, G.D. Excerpt from high-impact educational practices: What they are, who has access to them, and why they matter. *Assoc. Am. Coll. Univ.* **2008**, *14*, 28–29.
63. Carpi, A.; Ronan, D.M.; Falconer, H.M.; Lents, N.H. Cultivating minority scientists: Undergraduate research increases self-efficacy and career ambitions for underrepresented students in STEM. *J. Res. Sci. Teach.* **2017**, *54*, 169–194. [CrossRef]
64. Chen, S.; Binning, K.R.; Manke, K.J.; Brady, S.T.; McGreevy, E.M.; Betancur, L.; Limeri, L.B.; Kaufmann, N. Am I a science person? A strong science identity bolsters minority students' sense of belonging and performance in college. *Personal. Soc. Psychol. Bull.* **2021**, *47*, 593–606. [CrossRef] [PubMed]
65. Leary, M.; Tylka, A.; Corsi, V.; Bryner, R. The effect of first-year seminar classroom design on social integration and retention of stem first-time, full-time college freshmen. *Educ. Res. Int.* **2021**, *2021*, 426290. [CrossRef]
66. Kim, A.Y.; Sinatra, G.M.; Seyranian, V. Developing a STEM Identity Among Young Women: A Social Identity Perspective. *Rev. Educ. Res.* **2018**, *88*, 589–625. [CrossRef]
67. Ong, M.; Smith, J.M.; Ko, L.T. Counterspaces for women of color in STEM higher education: Marginal and central spaces for persistence and success. *J. Res. Sci. Teach.* **2018**, *55*, 206–245. [CrossRef]
68. Davis, R.D.; Wilson-Kennedy, Z.S.; Spivak, D. International Research Experiences in the Development of Minority Scientists. *Front. Educ.* **2021**, *6*, 674673. [CrossRef]
69. Cabrera, N.L.; Miner, D.D.; Milem, J.F. Can a summer bridge program impact first-year persistence and performance?: A case study of the New Start Summer Program. *Res. High. Educ.* **2013**, *54*, 481–498. [CrossRef]
70. Indiana University–Purdue University Indianapolis. *Program Review for Summer Bridge, First-Year Seminars, and Themed Learning Communities*, 2014th ed.; Indiana University–Purdue University Indianapolis: Office of Institutional Research & Decision Support: Indianapolis, IN, USA, 2014; pp. 1–123.
71. Yosso, T.J. Whose culture has capital? *A critical race theory discussion of community cultural wealth. Race Ethn. Educ.* **2005**, *8*, 69–91.
72. Silver, C.A.; Tatler, B.W.; Chakravarthi, R.; Timmermans, B. Social Agency as a continuum. *Psychon. Bull. Rev.* **2021**, *28*, 434–453. [CrossRef] [PubMed]
73. Yin, R.K. Case study methods. In *APA Handbook of Research Methods in Psychology*; Cooper, H., Camic, P.M., Long, D.L., Panter, A.T., Rindskopf, D., Sher, K.J., Eds.; American Psychological Association: Washington, DC, USA, 2012.
74. Yin, R.K. *Case Study Research: Design and Methods*, 5th ed.; SAGE Publications: Los Angeles, CA, USA, 2014.
75. Stake, R.E. *The art of Case Study Research*; Sage: Los Angeles, CA, USA, 1995.
76. Strauss, A.; Corbin, J. *Basics of Qualitative Research Techniques*; Citeseer: University Park, PA, USA, 1998.
77. Yin, R.K. *Case Study Research and Applications: DESIGN and Methods*; Sage Publications: Los Angeles, CA, USA, 2017.
78. Nie, Y. Combining narrative analysis, grounded theory and qualitative data analysis software to develop a case study research. *J. Manag. Res.* **2017**, *9*, 53–70. [CrossRef]
79. Thorne, S. Data analysis in qualitative research. *Evid.-Based Nurs.* **2000**, *3*, 68–70. [CrossRef]
80. Creswell, J.W.; Poth, C.N. *Qualitative Inquiry and Research Design: Choosing among Five Approaches*; Sage Publications: Los Angeles, CA, USA, 2016.
81. Denzin, N.K.; Lincoln, Y.S. *The Sage Handbook of Qualitative Research*; Sage Publications: Los Angeles, CA, USA, 2011.
82. Lincoln, Y.S.; Guba, E.G. Establishing trustworthiness. *Nat. Inq.* **1985**, *289*, 289–327.
83. Yarrow, A. *Hard Work, Hard Lives*; Oxfam: Washington, DC, USA, 2013.
84. Green, A.; Sanderson, D. The Roots of STEM Achievement: An Analysis of Persistence and Attainment in STEM Majors. *Am. Econ.* **2018**, *63*, 79–93. [CrossRef]
85. Fries-Britt, S.L.; Onuma, F.J. The Role of Family, Race, and Community as Sources of Motivation for Black Students in STEM. *J. Minor. Achiev. Creat. Leadersh.* **2020**, *1*, 151–187. [CrossRef]
86. Dagley, M.; Georgiopoulos, M.; Reece, A.; Young, C. Increasing Retention and Graduation Rates Through a STEM Learning Community. *J. Coll. Stud. Retent. Res. Theory Pract.* **2016**, *18*, 167–182. [CrossRef]
87. Palmer, R.; Gasman, M. "It takes a village to raise a child": The role of social capital in promoting academic success for African American men at a Black college. *J. Coll. Stud. Dev.* **2008**, *49*, 52–70. [CrossRef]
88. Nolan, J.R.; McConville, K.S.; Addona, V.; Tintle, N.L.; Pearl, D.K. Mentoring Undergraduate Research in Statistics: Reaping the Benefits and Overcoming the Barriers. *J. Stat. Educ.* **2020**, *28*, 140–153. [CrossRef]
89. Dennehy, T.C.; Dasgupta, N. Female peer mentors early in college increase women's positive academic experiences and retention in engineering. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 5964–5969. [CrossRef] [PubMed]
90. Dasgupta, N.; Stout, J.G. Girls and Women in Science, Technology, Engineering, and Mathematics: STEMing the Tide and Broadening Participation in STEM Careers. *Policy Insights Behav. Brain Sci.* **2014**, *1*, 21–29. [CrossRef]

education sciences

Article

A Person-Centered Approach toward Balanced Gender Identity in Emerging Adults: Associations with Self-Esteem and Attitudes about Education

Joyce J. Endendijk

Child and Adolescent Studies, Faculty of Social and Behavioural Sciences, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, The Netherlands; j.j.endendijk@uu.nl

Abstract: Balanced identity theory (BIT) has played an important role in research examining women's underrepresentation in science, technology, engineering, and mathematics (STEM). Yet, BIT's main balanced-congruity principle has not been tested specifically for gender-science cognitions. Additionally, BIT's predictions have been tested primarily from a variable-centered approach. The current study therefore examined whether (1) gender-science cognitions form a balanced identity configuration; (2) different identity profiles can be distinguished; (3) identity profiles differ in background characteristics, study motivation, and self-esteem. Dutch emerging adults (18–25 years old) enrolled in education (N = 318, 51% female) completed a gender-science Implicit Association Test (gender-science stereotypes) and questionnaires assessing felt similarity to males and females (gender identity), interest in science and liberal arts occupations (occupational self-concept), self-esteem, and study motivation and engagement. Hierarchical regression analyses revealed multiplicative interactions between gender-science stereotypes, gender identity, and occupational self-concept, providing evidence for a balanced identity configuration. Furthermore, latent profile analyses revealed three balanced identity profiles and two unbalanced profiles. Unbalanced identity profiles were characterized by non-Dutch ethnicity, lower educational level, and living independently without parents. The identity profiles did not differ in self-esteem and study motivation. Future research should examine the longer term consequences of unbalanced identity for academic and career pursuits.

Keywords: balanced identity theory; gender-science stereotypes; gender identity; self-concept; occupational interest

Citation: Endendijk, J.J. A Person-Centered Approach toward Balanced Gender Identity in Emerging Adults: Associations with Self-Esteem and Attitudes about Education. *Educ. Sci.* **2023**, *13*, 424. https://doi.org/10.3390/educsci13040424

Academic Editors: Erin Mackenzie, Bernhard Ertl, Christine R. Starr, Silke Luttenberger and Manuela Paechter

Received: 9 March 2023
Revised: 13 April 2023
Accepted: 19 April 2023
Published: 21 April 2023

1. Introduction

Even though gender roles for men and women have become more equal over the past decades, gender segregation in the occupational domain is still clearly visible and declines in occupational gender segregation may have stagnated in the past two decades [1]. For instance, women are underrepresented in some of the STEM fields (i.e., science, technology, engineering, mathematics), whereas men are underrepresented in fields such as health care and education [2,3]. More specifically, in OECD countries (Organisation for Economic Co-operation and Development) only 11.6% of employed women worked in the industry sector compared to 32.6% of men [3]. Additionally, 18% of primary school teachers are male, and only 10% of long-term (elder) care providers are male [3]. In the US, women make up less than a quarter of workers in computing and engineering, a proportion that remained stable or even decreased slightly since 1990 [4]. These gender disparities are also apparent in tertiary degree enrolments, with less than 20% of students enrolled in engineering and computer science programs being female [2]. In contrast, about 20% of students enrolled in programs related to education, health and welfare were male, a percentage that remained low and stable over time [2,5]. More gender equality in the occupational domain is crucial because occupational gender segregation has been associated with the gender pay gap [3,6], and because more gender-diversity in work teams improves team collaboration and performance [7].

There is ample evidence that, rather than gender differences in ability, gender stereotypes, self-concept, and gender identity, as well as other gender attitudes and beliefs, are important explanatory factors for women's underrepresentation in STEM [8,9]. Yet, relatively less is known about how these gender cognitions together play a role in occupational gender segregation. The first goal of this study was therefore to examine the interplay between gender-science stereotypes, gender identity, and occupational self-concept in emerging adults. Gender-science stereotypes are conceptualized as the degree to which people (implicitly) associate gender with science (e.g., chemistry, physics, engineering, mathematics, and astronomy) and liberal arts (e.g., philosophy, humanities, arts, languages, music, and history) [10]. Gender identity refers to associations between the self and gender [10], and is conceptualized as the degree to which one feels similar to the male and female gender group [11]. Importantly, even though gender similarity and gender identity are used interchangeably throughout the manuscript, gender similarity represents only a part of the multidimensional gender identity construct. Occupational self-concept refers to associations between the self and science and liberal arts occupations [10], and is conceptualized as one's interest in science and liberal arts occupations.

1.1. Balanced Identity Theory

According to balanced identity theory, the principle of consistency organizes associations between the three gender cognitions into a balanced configuration [12]. This schematic consistency is also central to gender schema theories that propose that people will act and think consistently with their gender schemas in order to avoid feelings of discomfort or anxiety [13]. Consider, for example, a man who identifies as male (gender identity: me = male) and strongly associates science with men (gender stereotype: science = male). To avoid threats to his identity, this man is likely to see himself as being interested in science occupations (self-concept: me = science). Previous research demonstrated this balance-congruity principle for math cognitions in adults as well as children [12,14]. In the case of strong associations of the self with female and math with male, there was little association of the self with math.

Recently, a meta-analysis including 36 experiments provided proof of the balanced-congruity principle for implicit measures as well as for self-report measures across a range of different social cognition domains (e.g., race, gender, age, math/language, good/bad, work/family) [15]. The current study extends this research by examining whether balanced-congruity can also be demonstrated with a combination of self-report (i.e., explicit) and implicit measures. Also, none of the included studies in the meta-analysis specifically focused on the gender-science domain.

There are some studies that provide partial or indirect evidence for the balanced-congruity principle regarding gender-science stereotypes. For instance, US high-school students' stronger explicit STEM stereotypes (STEM is for geniuses) were related to less STEM motivation, but only for students who did not explicitly identify as nerd-genius (low nerd-genius self-concept) [16]. Additionally, in a sample of US university students, men and women who explicitly identified with science differed considerably in the strength of their explicit gender-science stereotypes [17]. Specifically, for male students, stronger science-with-self associations were linked with stronger gender-science stereotypes and stronger science career aspirations. For female students, stronger science-with-self associations were linked with weaker gender-science stereotypes and weaker science career aspirations. However, neither of these studies assessed gender identity, only the participant's sex/gender. Lane et al. [18] did assess three gender-science cognitions, and found that among female students who strongly identified as female (implicitly), implicit associations between female and science predicted greater intent to pursue science over the humanities in university. For men, the association between stronger gender-science stereotypes and the greater likelihood of pursuing science rather than the humanities in university was not moderated by their gender identity.

Lane et al. [18], however, only examined one element or direction of the balanced-congruity principle, i.e., stereotype-emulation. Stereotype-emulation entails that the more people identify with a particular gender, the more they incorporate attributes they perceive as associated with that gender into their self-concept [19]. Lane et al. [18] did not investigate stereotype-construction, i.e., whether people who identify more with a particular gender also project attributes they perceive in themselves more onto the gender collective (i.e., stereotypes) [19]. Neither did they examine identity-construction, i.e., whether people who perceive that their own attributes match with their stereotypes for a particular gender, also identify more with that gender [19]. The current study will test all three elements of the balanced-congruity principle for gender-science cognitions.

1.2. A Person-Centered Approach toward Balanced Identity Theory

It seems likely to assume that there might be individual differences in the degree to which people have achieved a balanced configuration of gender cognitions. Balanced identity theory assumes that "when two unlinked or weakly linked nodes share a first-order link, the association between these two should strengthen" [10] (p. 6). This might indicate that achieving a balance between social cognitions is an ongoing process [10] and when one cognition changes (i.e., because of external pressure, maturation or development), one of the other two cognitions changes as well in order to restore balance. Disbalance (or dissonance, inconsistency) between cognitions is thought to provide a source of discomfort and stress, providing pressure toward change and to renew balance [20,21].

Temporary disbalance between gender identity, gender-science stereotypes, and occupational self-concept might be particularly likely in emerging adulthood, as this developmental period is considered as highly volitional, offering the most opportunity for identity exploration in the areas of love, work, and worldviews [22]. Additionally, there is evidence for developmental changes in gender stereotypes [23], gender identity [24,25], and occupational self-concept [26] into emerging adulthood. Changes in each gender cognition would require ongoing adjustment to the other gender cognitions to restore balance.

A useful method to capture individual differences in the balanced configuration of gender cognitions is a person-centered approach. This approach allows for the distinguishing of identity profiles based on individual variation in scores on multiple gender cognitions. This approach is common in the identity development literature, demonstrating that individuals may adopt various identity configurations [27,28], but a person-centered approach has, to the best of my knowledge, not yet been applied to balanced identity theory. Based on balanced identity theory, a concordance between the three gender cognitions is expected. Therefore, one may find overrepresentations of emerging adults in specific types of balanced identity configurations. See Table 1 for examples of balanced identity configurations. As people can also hold egalitarian gender-science stereotypes (science = male/female) or feel similar to both genders (me = male/female), less straightforward configurations of balanced identity are possible as well, such as: science = male/female, me = male/female, me = science/liberal arts. Yet, because emerging adults are still actively exploring their occupational identity and world views [22], I also expect to identify unbalanced identity configurations, such as: science = female, me = female, science $\neq$ me.

Table 1. Examples of Balanced Configurations of Gender-Science Stereotypes, Gender Identity, and Occupational Self-Concept.

Identity Configuration	Gender-Science Stereotypes	Gender Identity	Occupational Self-Concept
1. Balanced	science = male	me = male	science = me
2. Balanced	science = female	me = female	science = me
3. Balanced	science = male	me = female	science $\neq$ me
4. Balanced	science = female	me = male	science $\neq$ me

Note. The balanced identity profiles in the table are developed by the author based on previous research examining balanced identity [10,12].

1.3. Correlates of Different Profiles of Balanced and Unbalanced Identity

Once different types of balanced and unbalanced identity configurations have been identified, an important next step is to examine the correlates of identity profile membership. This may yield valuable insights for both theory development and clinical practice, because it allows for the identification of which emerging adults are most likely to exhibit balanced or unbalanced identity profiles. There is hardly any research on the predictors of individual differences in balanced identity, so my investigation of correlates of balanced identity profiles was mainly explorative.

First, age might be associated with the identity profiles, as there is some developmental evidence that the strength of balanced identity was positively associated with age in childhood [14,29]. Relatedly, increases in cognitive flexibility in emerging adulthood could be reflected in more flexible views of one's own identity, self-concept, and stereotypes [30,31], which might either increase or decrease the likelihood of achieving a balanced identity. Second, balanced identity might differ between genders, as an unpublished Master's thesis showed that female STEM students had a less balanced configuration of gender-STEM cognitions than male STEM students [32]. Third, balanced identity might also differ between students from different study majors. For instance, in the same unpublished Master's thesis, students with a biological or life-science major showed more pronounced balanced identity than students from engineering and computer science majors [32]. This difference might be due to a more equal gender representation in the biological and life science majors. Fourth, emerging adults' educational level might be associated with the identity profiles, as a higher educational level is associated with more egalitarian views about gender [33,34] and less gender-typical identity [35], which might increase the likelihood of achieving balanced identity. Finally, balanced identity might also differ between emerging adults who still live with their parents and those who live by themselves. Emerging adults who live with their parents might experience pressure from parents as well as peers towards gender conformity [36,37], whereas their counterparts who live by themselves might experience pressure from parents to a lesser extent. Higher pressure towards gender conformity is associated with more gender-identity typicality (strong associations between me and male or me and female) [36,37], which might introduce disbalance in the associations between gender cognitions.

1.4. Outcomes of Different Profiles of Balanced and Unbalanced Identity

A final aim of this study was to examine how the different identity profiles differ in terms of study motivation and engagement and self-esteem. Study motivation is defined as the extent to which students are motivated to do well in education [38]. Study engagement entails students' positive (and negative) affective reactions to school work, as well as a psychological investment in schoolwork and a preference for challenge [39]. Self-esteem entails an association between the self and positive valence attributes, hence a positive view of oneself [10].

Balanced identity theory [40] suggests that people who achieve a balanced identity are more likely to persist in their academic and career pursuits. Similarly, Adams and Marshall [41] theorized that a stable identity provides people with a sense of consistency and harmony among one's beliefs; a future orientation, including goals and direction; and a sense of personal control and agency that together enables people to take an active role in the process of pursuing academic and career goals. Therefore, balanced identity is expected to be associated with increased study motivation and engagement [42].

In addition, self-esteem plays a central role in balanced identity theory [10], therefore the identity profiles might also differ with regard to self-esteem. According to cognitive-developmental theory [31], people's motivation to match their behavior and self-concept to the stereotypes for the gender they identify with is considered to be an intrinsic desire for cognitive consistency and the enhancement of self-esteem. Therefore, I expect self-esteem to be higher for emerging adults with a balanced identity profile than for emerging adults with unbalanced identity profiles.

1.5. Current Study

In sum, the goals of this study were as follows: (1) to examine whether gender-science stereotypes, gender similarity, and occupational self-concept (i.e., interest) form a balanced configuration of gender-science cognitions; (2) explore whether different profiles of balanced or unbalanced gender-science cognitions can be distinguished; (3) explore how different identity profiles are associated with background variables; and (4) examine how different identity profiles differ in terms of study motivation and engagement and self-esteem.

Regarding the first aim, it is expected that: (1) the more people feel similar to a particular gender, the more they incorporate attributes they associate with that gender (science or liberal arts) into their occupational self-concept (the stereotype-emulation hypothesis); (2) the more people feel similar to a particular gender, the more they project attributes they associate with themselves (science or liberal arts) onto the whole gender group (the stereotype-construction hypothesis); (3) the more people perceive that the attributes they associate with themselves (science or liberal arts) match with their gender-science stereotypes for a particular gender, the more similar they feel to that gender (the identity-construction hypothesis) [10,19].

Regarding the second aim, it is expected that both balanced (e.g., science = me; me = male, science = male) and unbalanced identity profiles could be discerned (e.g., science = me; me = female, science = male), but the balanced identity profiles would be more prevalent [10,22]. Regarding the third aim, the hypothesis was explorative, and I just explored associations between identity profiles and age, gender, study major, educational level, ethnicity, and living with parents. Regarding the fourth aim, I expected that students in balanced identity profiles would report higher study motivation and engagement and self-esteem compared to students with unbalanced identity profiles [10,31,41].

I specifically studied these aims in the Dutch context. The Netherlands generally scores high on worldwide indices of gender equality [43]. However, at the same time adults in the Netherlands scored highest on the gender-science stereotypes of 66 countries [44]. These seemingly contrasting findings have been explained by the clear domain-specific occupational gender segregation that is visible in the Netherlands. For instance, the female share of graduates in STEM fields is less than 30%, and among the lowest of Western countries, whereas more than 75% of graduates in health care and education are female [2,3]. It might be particularly compelling to examine the presence of balanced identity in a context with clear occupational gender segregation and strong domain-specific gender stereotypes.

2. Methods

2.1. Participants

This study made use of the data from a larger project on the role of love and friendship in the psychosocial functioning and gender development of emerging adults [24]. Dutch emerging adults between 18 and 25 years of age were recruited via the personal networks of 29 students that were writing their Bachelor's or Master's thesis under supervision of the author. Using information leaflets (provided in-person, via email, or social media), each student recruited 10–20 participants currently enrolled in education. The focus on emerging adults that are enrolled in education is important because they are still in the middle of the career-decision making process [45]. Furthermore, emerging adulthood is an important period for gender identity formation and consolidation [46]. These developments make emerging adulthood an optimal period for studying the role of balanced identity in the career-decision making process.

The initial sample consisted of 409 participants. Samples >300 are generally recommended for studies with person-centered designs [47–49]. Of these participants, 28 were excluded, as they did not complete the implicit association task (see Measures), which was used as an attention check. The final sample thus consisted of 381 emerging adults across the three educational levels available in the Netherlands: lower vocational level (preparation for an associate's degree, e.g., clerk, plumber, $n = 101$), higher vocational level (preparation for a vocational bachelor's degree, e.g., secondary school teacher, real-estate agent, $n = 119$), and university level (preparation for a master's degree, $n = 161$) (see Table 2 for sample characteristics). About half of the sample was enrolled in a science or liberal arts major. The ethnic diversity of the sample was similar to that of the Dutch population.

Table 2. Participant demographics.

Variable	
N	381
Females, *n* %	196 (51)
Age, *M* (*SD*)	21.73 (2.02)
Ethnicity, %	
Dutch	81
Moroccan	1
Turkish	2
Surinam	7
Asian	1
Indonesian	2
Other	5
Education level, %	
Lower Vocational	26.5
Higher Vocational	31.2
University	42.3
Science major, %	20
Liberal arts major, %	31
Living with parents, %	49

2.2. Procedure and Measures

Participants completed an online survey (duration: approximately 45 min) including questions about background characteristics, gender identity, gender stereotypes, occupational interests, friendships, romantic relationships, and social-emotional adjustment. At the beginning of the survey they provided informed consent. Approximately half of the lower vocational students completed the questionnaires in class under the supervision of a student assistant. The other half of the lower vocational students as well as the higher vocational students and academic students completed the questionnaire by themselves at their convenience. Controlling for this difference in procedure in the analyses did not change the results. Participants did not receive any financial compensation for their participation. This research was approved by the Ethical Review Board of the Faculty of Social and Behavioural Sciences of the author's university. The study was not pre-registered.

2.3. Gender-Science Stereotypes

Gender-science stereotypes were measured with a Dutch translation of the gender-science Implicit Association Test (IAT; described in detail in [50]). This computer task measures the strength of associations between the concepts of male (i.e., man, boy, father, male, grandpa, husband, son, uncle) and female (i.e., girl, female, aunt, daughter, wife, woman, mother, grandma) with the attributes of science (i.e., biology, physics, chemistry, astronomy, engineering, math, geology, biophysics, computer science, electrotechnics) and liberal arts (i.e., philosophy, humanities, arts, English, music, history, literature, theology, cultural sciences, social sciences). During the IAT, participants were requested to sort words into groups by pressing keys. In congruent blocks, female concepts (e.g., 'girl') and liberal arts attributes (e.g., 'humanities') needed to be sorted under the 'Female & Liberal arts' category, and male concepts (e.g., 'boy') and science attributes (e.g., 'math') needed to be sorted under the 'Male & Science' category. In incongruent blocks, female concepts and liberal arts attributes needed to be sorted under the 'Female & Science' category and male concepts and science words needed to be sorted under the 'Male & Liberal arts' category.

The improved scoring algorithm of Greenwald and colleagues [51] was used to determine the level of gender-science stereotypes of the participant. In short, the gender stereotype score calculated with this algorithm reflects the difference in response latencies between stereotype-incongruent and stereotype-congruent blocks (divided by the pooled SD of response latencies across all trials). Positive scores indicated male-with-science and female-with-liberal arts associations, whereas negative scores represented female-with-science and male-with-liberal arts associations.

2.4. Gender Similarity

Similarity to male and female peers was assessed with a measure developed by Martin and colleagues [11]. Students answered 10 questions regarding how similar they felt to both men (e.g., "How similar do you feel to men?") and women (e.g., "How similar do you feel to women?"). A graphical response scale was used with two circles, one representing oneself and one representing either men or women [11]. The closer the two circles were together, the greater the perceived similarity. Responses ranged from 0 (two circles farthest apart) to 4 (two overlapping circles). Participant responses on the five male and five female items were averaged separately, with higher scores representing more similarity to respectively the male gender and the female gender (male similarity: $\alpha = 0.84$, female similarity: $\alpha = 0.78$). Mirroring the response scale of the gender-science stereotypes, the female similarity scale was subsequently subtracted from the male similarity scale, leading to a gender similarity variable with positive scores representing more similarity to males compared to females (me = male), and negative scores representing more similarity to females compared to males (me = female).

2.5. Occupational Interests in Science and Liberal Arts

Participants reported their interest in a range of occupations related to science and liberal arts. Occupations were selected from the Strong Interest Inventory [52] that were most closely aligned with the science and liberal arts domains used in the gender-science IAT (see Supplementary Materials, p.1 for a complete list of occupations). Participants rated their interest in each of 25 jobs by indicating how much they would like to be in that job on a 5-point scale (1 = not at all, 2 = not really, 3 = a little, 4 = moderately, 5 = very much). Participant responses on the 14 science and 11 liberal arts occupations were averaged, with higher scores representing more interest in occupations in science ($\alpha = 0.87$) and liberal arts ($\alpha = 0.83$), respectively. Mirroring the response scale of the gender-science stereotypes, the liberal arts interest scale was subsequently subtracted from the science interest scale, leading to an occupational interest variable with positive scores representing more interest in science compared to liberal arts (me = science), and negative scores representing more interest in liberal arts compared to science (me = liberal arts).

2.6. Self-Esteem

Self-esteem was assessed with the Rosenberg Self-Esteem Scale [53]. Participants indicated feelings of value and self-worth on 10 items (e.g., "I am satisfied with myself") on a scale from 0 = strongly disagree to 3 = strongly agree. Item scores were averaged to create a self-esteem scale (α = 0.89), with higher scores representing higher self-esteem.

2.7. Study Motivation and Engagement

Study motivation was assessed by a five-item questionnaire developed by Field et al. [38]. Participants answered questions such as "While you are doing school work, how often do you feel . . . [like you want to try hard]?" on a 5-point scale, ranging from 1 (never/almost never) to 5 (always/almost). The responses were averaged to create a study motivation scale (α = 0.72).

Study engagement was measured with an adapted questionnaire developed by Eccles et al. [54]. Participants answered nine questions, such as "While you are doing school work how often do you feel... [excited and challenged]?" on a 5-point scale, ranging from 1 (never/almost never) to 5 (always/almost). The responses were averaged to create a study engagement scale (α = 0.67).

3. Results

3.1. Descriptive Statistics

First, descriptive statistics and Pearson correlations were computed in SPSS (version 28). Means, standard deviations, and intercorrelations of all study variables can be found in Table 3. The significant positive correlation between occupational interest (science vs. liberal arts) and gender similarity (to males vs. females) indicates that more interest in science occupations is associated with more felt similarity to males. More interest in science occupations was also significantly associated with more self-esteem and less study motivation. Similarly, more felt similarity to males was significantly associated with more self-esteem and less study motivation. Finally, higher study engagement was significantly related to higher self-esteem and study motivation. Gender-science stereotypes were not associated with any of the study variables.

Table 3. Descriptive Statistics and Intercorrelations for all Study Variables.

Variable	1.	2.	3.	4.	5.	*M* (*SD*)
1. Occupational interest						−0.37 (0.86)
2. Gender-science stereotypes	−0.01					0.33 (0.37)
3. Gender similarity	0.43 ***	−0.06				0.12 (1.63)
4. Self-esteem	0.12 *	−0.07	0.15 **			3.13 (0.55)
5. Study motivation	−0.13 *	−0.01	−0.24 **	0.07		3.26 (0.69)
6. Study engagement	0.11	−0.03	0.05	0.35 **	0.34 **	3.12 (0.52)

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

3.2. Hierarchical Regressions Testing for Balanced Identity

Second, in order to test for the presence of balanced identity, three hierarchical regression analyses were conducted in SPSS. In each regression, one of the gender cognitions (e.g., occupational interest) was the outcome variable, the other two gender cognitions were entered as predictors in the first step of the regression (e.g., gender-science stereotypes, gender similarity), and the interaction between the two predictors was entered in the second step. As there is discussion in the literature on how to best test for balanced identity [55,56], I also applied a second method, known as the 4-test method [56] (for more detail, see Supplemental Materials, p. 5), in which the same regression analyses as described above were conducted, but with a reversed order of entering predictors and the interaction term (i.e., step 1: interaction term only; step 2: predictors added). This 4-test method arguably provides a more pure test of balanced identity [56].

Table 4 presents the results of the 3 hierarchical regression analyses testing for balanced identity. First, in the regression with occupational interest as the outcome, a significant main effect of gender similarity was found, indicating that more felt similarity to males (compared to females) was associated with more interest in science occupations (compared to liberal arts occupations). The interaction between gender-science stereotypes and gender similarity was also significant. In Figure 1A, the interaction is decomposed via simple slopes (Mean gender similarity +/− 1*SD*), revealing that for emerging adults who feel more similar to males (me = male), stronger gender-science stereotypes (science = male) were associated with more interest in science occupations (me = science). For emerging adults who feel more similar to females (me = female), stronger gender-science stereotypes (science = male, liberal arts = female) were associated with more interest in liberal arts occupations (me = liberal arts).

Table 4. Hierarchical Regression Analyses Testing Interactions Between Occupational Interests (Science vs. Liberal arts), Gender-Science Stereotypes, and Gender Similarity (to Males vs. Females. * $p < 0.05$. *** $p < 0.001$.

Variable	*B*	*SE B*	β	R^2	ΔR^2
Outcome = Occupational interest					
Step 1					
Constant	−0.37 ***	0.04			
Gender-science stereotypes	0.04	0.11	0.02		
Gender similarity	0.23 ***	0.02	0.44		
Step 2				0.22	0.04 ***
Constant	−0.36 ***	0.04			
Gender-science stereotypes	0.02	0.10	0.01		
Gender similarity	0.23 ***	0.02	0.44		
Gender stereotypes × Gender similarity	0.26 ***	0.06	0.19		
Outcome = Gender-science stereotypes					
Step 1				<0.01	<0.01
Constant	0.33 ***	0.02			
Occupational interests	0.01	0.03	0.02		
Gender similarity	−0.02	0.01	−0.07		
Step 2				0.05	0.05 ***
Constant	0.29	0.04			
Occupational interests	−0.02	0.03	−0.04		
Gender similarity	−0.02	0.01	−0.07		
Occupational interests × Gender similarity	0.07 ***	0.02	0.22		
Outcome = Gender similarity					
Step 1				0.19	0.19 ***
Constant	0.12	0.08			
Occupational interests	0.82 ***	0.09	0.43		
Gender-Science stereotypes	−0.23	0.20	−0.05		
Step 2				0.20	0.01 *
Constant	0.12	0.08			
Occupational interests	0.79 ***	0.09	0.42		
Gender-Science stereotypes	−0.26	0.20	−0.06		
Occupational interests × Gender stereotypes	0.55 *	0.23	0.11		

Second, in the regression with gender-science stereotypes as outcome, only a significant interaction between occupational interests and gender similarity was found. Figure 1B shows that for emerging adults who feel more similar to males (me = male), more interest in science occupations (me = science) was associated with stronger gender-science stereotypes (science = male). For emerging adults who feel more similar to females (me = female), more interest in science occupations (me = science) was associated with less strong gender-science stereotypes (science = female).

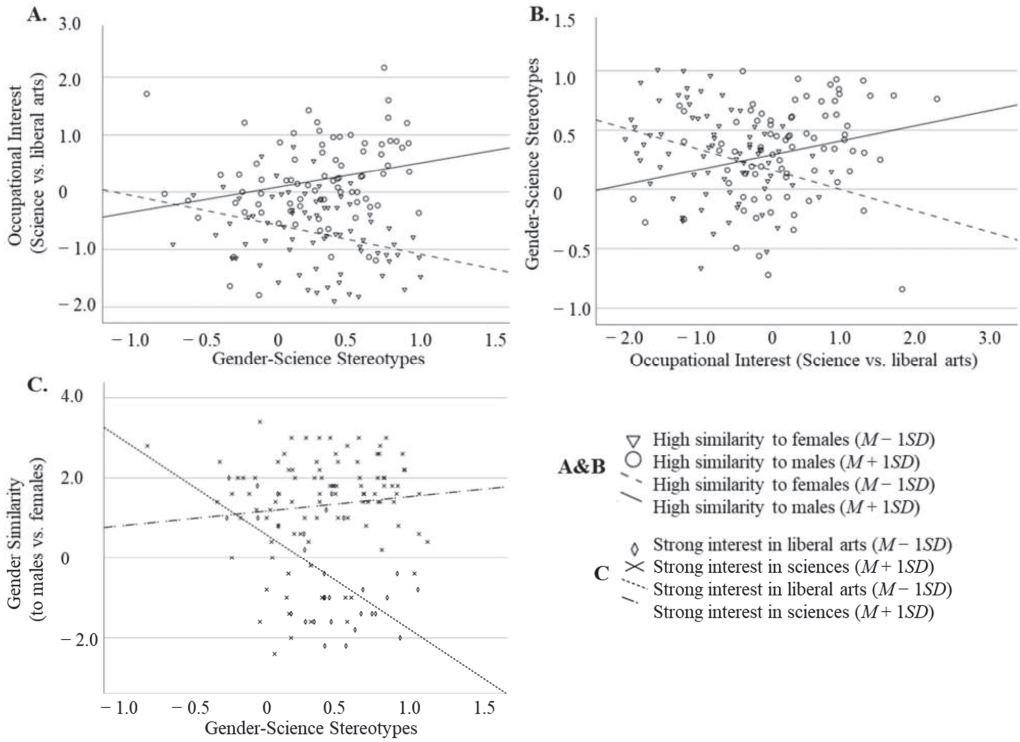


Figure 1. Simple Slopes for Interactions Between Gender-Science Stereotypes, Gender Similarity, and Occupational Interests. Part (**A**) represents the interaction between gender-science stereotypes and gender similarity. Part (**B**) displays the interaction between occupational interests and gender similarity. Part (**C**) displays the interaction between gender-science stereotypes and occupational interest.

Third, in the regression with gender similarity as outcome, a significant main effect of occupational interest was found, indicating that more interest in science occupations (compared to liberal arts occupations) was associated with more felt similarity to males (compared to females). The interaction between gender-science stereotypes and occupational interest was also significant. Figure 1C shows that for people with more interest in the liberal arts (me = liberal arts), stronger gender-science stereotypes (liberal arts = female) were associated with feeling more similar to females (me = female).

I also tested whether the interaction effects were found over and above the effect of study major (liberal arts or science). The results remained the same with the inclusion of the study major as a covariate (see Supplementary Materials, Tables S2–S4). Relatedly, hierarchical regression analyses testing for pure balanced identity revealed that eight of the 12 steps of Greenwald's test for balanced identity were met, and the four steps that were not met might be attributable to the scaling of the variables "gender-science stereotypes" and "occupational interest" (see Supplementary Materials, Tables S5–S7).

3.3. Latent Profile Analyses to Identify Different Types of Balanced Identity

Third, to test for the presence of different types of balanced identity, Latent Profile Analyses (LPA) were conducted in MPlus (version 8.7) [57,58] with gender-science stereotypes, gender similarity, and occupational interest as indicator variables. LPA identifies distinct subgroups in the sample and fits individuals to the most likely class based on their responses to the indicator variables. The default MLR estimator was used to fit a series of models to identify the optimal number of classes from one- to six-class options. To address the potential problem of local maxima, models were estimated with 1000 random starts and 250 iterations per random start. To identify the optimal number of classes, several fit indices were compared between the models, and preference was given to the model with superior statistics and theoretical interpretability. The indices of model fit included the Bayesian information criterion (BIC) and sample size adjusted Bayesian information criterion (SABIC), the Lo-Mendel-Rubin (LMR) likelihood ratio test, and the bootstrap likelihood ratio test (BLRT) [47,48]. A smaller BIC value indicates better model fit than models with larger BIC values, and significant LMR and BLRT tests indicate that the model significantly improved model fit compared to the previous model with one class less. Entropy was also used to reflect the accuracy of class assignments, with cutoff values above 0.8 deemed accurate, though no definitive cutoff value is suggested [49]. Finally, models with very small classes (with a size of <5% of the total sample) are unfavorable for model performance [59].

See Table 5 for class solutions of one to six classes from the LPA. The best fit was a solution with five classes, as evidenced by decreases in AIC, SABIC, sufficient entropy, significant LMR and BLRT tests, and sufficient emerging adults in each group. Figure 2 shows the mean profiles of the five classes of participants (see Supplementary Materials Table S8 for ANOVA results and post-hoc comparisons between the five classes on the three gender cognitions). Emerging adults in class 1 (n = 32, 8%) associated science with males and liberal arts with females, and although they felt similar to males, they had a strong interest in liberal arts occupations. They therefore represent an unbalanced gender identity configuration. The largest class 2 (n = 124, 33%) represents a balanced configuration, with people associating science with males and liberal arts with females, feeling similar to females and having interest in liberal arts. People in class 3 (n = 98, 26%) associated science with males and liberal arts with females, and although they felt similar to males, they had a somewhat stronger interest in liberal arts occupations than in science occupations. Class 3 was less unbalanced than class 1, as the interest in liberal arts in class 3 was less pronounced than in class 1. People in class 4 (n = 68, 18%) displayed a balanced identity by strongly associating science with males and liberal arts with females, feeling similar to females, and being interested in liberal arts. Although classes 4 and 2 were similar, class 4 displayed stronger gender-science stereotypes and a more pronounced interest in liberal arts. Finally, people in class 5 (n = 59, 15%) also displayed a balanced identity by strongly associating science with males and liberal arts with females, feeling similar to males and being interested in science.

Table 5. Model Fit Indices of LPA's for Deciding the Number of Classes.

	Number of Classes					
Fit Indices	**1**	**2**	**3**	**4**	**5**	**6**
−2LL	−1620.35	−1510.50	−1493.31	−1488.80	**−1476.14**	−1469.83
AIC	3252.69	3041.00	3014.61	3013.96	**2996.29**	2991.65
BIC	3276.35	3080.43	3069.81	3084.93	**3083.03**	3094.17
SABIC	3257.31	3048.70	3025.39	3027.82	**3013.23**	3011.67
LMR	-	<0.001	0.020	0.659	**0.013**	0.392
BLRT	-	<0.001	<0.001	0.235	**<0.001**	0.065
Entropy	1.00	0.92	0.81	0.67	**0.80**	0.82
Class n's						
1	381	195	192	126	**32**	124
2		186	135	96	**124**	95
3			54	63	**98**	68
4				96	**68**	55
5					**59**	32
6						7

Note. The selected model is indicated in bold. −2LL = −2 log likelihood; AIC = Akaike information criteria; BIC = Bayesian information criteria; SABIC = sample-size adjusted BIC; LMR = p-value of Lo–Mendell–Rubin likelihood ratio test; BLRT = p-value of the Bootstrapped Likelihood Ratio Test.

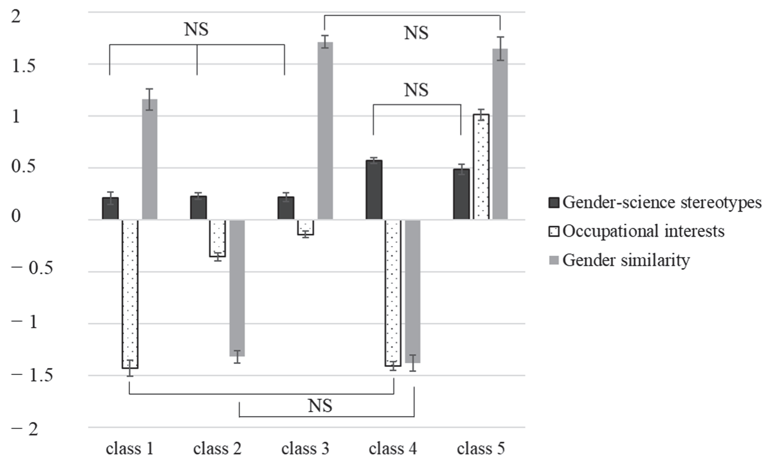


Figure 2. Mean Scores of the 5 Classes on Gender-Science Stereotypes, Occupational Interests, and Gender Similarity. Note. Raw scores are displayed instead of standardized scores for ease of interpretation. Error bars represent standard errors of the mean. Only non-significant differences (NS) between classes on the variables are highlighted; all other differences are significant. Positive scores on gender-science stereotypes represent science = male and liberal arts = female associations. Positive scores on occupational interest represent interests in science over liberal arts, whereas negative scores represent interest in liberal arts over science. Positive scores on gender similarity represent more felt similarity to males, whereas negative scores represent more felt similarity to females.

3.4. Characterization of Class Membership

Fifth, to examine differences between classes on the background variables, the following background variables were inserted into the model as covariates: gender, age, education level, liberal arts major, Dutch ethnicity, and living with parents. Science major was not entered as a covariate, as it lead to inflated estimates and convergence issues in the models. An automated three-step approach (r3step) was taken [60], including multinomial logistic regressions of participants' most likely class membership (as nominal variable) on the background variables.

Table 6 shows the results of the multinomial logistic regressions for the effects of the background variables on class membership. Regarding gender, emerging adults in class 4 were most likely to be female, followed by class 2, and then followed by classes 1 and 5, and finally by class 3. This fits with their reported gender similarity. Regarding the effect of a liberal arts major, there was a greater likelihood that emerging adults in class 1 and 4 followed a liberal arts major compared to class 2 and 3. Emerging adults in class 5 had the lowest likelihood of all classes to have a liberal arts major. This fits with their reported occupational interest. Regarding living with parents, emerging adults in class 2 and 4 (balanced) were more likely to live with their parents compared to class 5 (balanced). Emerging adults in class 1 and 3 (unbalanced) were the least likely of all classes to live with their parents. Regarding educational level, emerging adults in class 5 (balanced) were more likely to be higher educated than participants in class 3 (unbalanced). Finally, participants in class 4 (balanced) were more likely to be Dutch than participants in class 3 (unbalanced). No effect of age was found on class membership.

Table 6. Multinomial Logistic Regressions for the Effects of Background Variables on Class Membership.

Model		Reference Classes							
		Class 1 *		Class 2		Class 3		Class 4	
Non-Reference Class		**Est. (SE)**	***p***	**Est. (SE)**	***p***	**Est. (SE)**	***p***	**Est. (SE)**	***p***
Class 2	Female gender	**44.63 (1.69)**	**<0.001**						
Science = male	Liberal arts major	−2.46 (1.36)	0.071						
Me = female	Age	−0.58 (0.64)	0.371						
Me = liberal arts	Living with parents	**19.79 (0.90)**	**<0.001**						
	Educational level	1.58 (0.96)	0.100						
	Dutch ethnicity	0.57 (1.19)	0.634						
Class 3	Female gender	**−27.00 (1.69)**	**<0.001**	**−71.62 (<0.01)**	**<0.001**				
Science = male	Liberal arts major	**−4.27 (1.42)**	**0.003**	−1.81 (1.94)	0.350				
Me = male	Age	0.08 (0.26)	0.759	0.66 (0.65)	0.313				
Me = liberal arts	Living with parents	1.22 (0.75)	0.105	**−18.58 (0.60)**	**<0.001**				
	Educational level	0.66 (0.68)	0.335	−0.92 (1.00)	0.359				
	Dutch ethnicity	−1.92 (1.33)	0.150	−2.48 (1.47)	0.092				
Class 4	Female gender	**48.59 (1.69)**	**<0.001**	**3.96 (<0.01)**	**<0.001**	**75.58 (<0.01)**	**<0.001**		
Science = male	Liberal arts major	−0.14 (1.30)	0.916	**2.32 (0.72)**	**0.001**	**4.13 (1.90)**	**0.030**		
Me = female	Age	−0.27 (0.67)	0.683	0.30 (0.22)	0.174	−0.35 (0.67)	0.600		
Me = liberal arts	Living with parents	**20.19 (1.22)**	**<0.001**	0.40 (0.67)	0.550	**18.98 (0.90)**	**<0.001**		
	Educational level	1.15 (1.04)	0.268	−0.43 (0.47)	0.365	0.49 (1.08)	0.650		
	Dutch ethnicity	1.80 (1.22)	0.141	1.23 (0.73)	0.092	**3.72 (1.51)**	**0.014**		
Class 5	Female gender	−0.56 (1.76)	0.752	**−46.71 (<0.01)**	**<0.001**	**24.91 (<0.01)**	**<0.001**	**−50.67 (<0.01)**	**<0.001**
Science = male	Liberal arts major	**−5.74 (1.77)**	**0.001**	**−4.59 (2.19)**	**0.036**	**−2.78 (1.38)**	**0.044**	**−6.91 (2.16)**	**0.001**
Me = male	Age	0.17 (0.27)	0.532	0.72 (0.65)	0.272	0.06 (0.10)	0.545	0.42 (0.68)	0.540
Me = science	Living with parents	**3.22 (1.02)**	**0.002**	**−16.98 (<0.01)**	**<0.001**	**1.60 (0.60)**	**0.013**	**−17.38 (0.67)**	**<0.001**
	Educational level	1.21 (0.79)	0.124	−0.16 (0.98)	0.874	**0.76 (0.37)**	**0.040**	0.27 (1.07)	0.798
	Dutch ethnicity	−0.28 (1.40)	0.842	−0.66 (1.52)	0.662	1.82 (1.26)	0.149	−1.90 (1.56)	0.224

Note. Est. = logistic regression coefficient, adjusted for all other variables in the model. Significant effects are highlighted in bold. * Identity configuration of class 1: science = male, me = male, me = liberal arts.

3.5. Associations between Class Membership and Self-esteem and Study Motivation

Finally, to examine how class membership predicted participants' self-esteem and study motivation and engagement, self-esteem and study motivation and engagement were standardized and added to the LPA model as distal outcomes of each class. An automated three-step approach (du3step) was taken, and Wald tests were performed to compare class differences on the distal outcomes [61,62].

Figure 3 shows differences between classes on self-esteem and study motivation. Self-esteem differed significantly between classes (χ^2 = 10.50, p = 0.033). Specifically, participants in class 4 (balanced: me = female, me = liberal arts) reported lower self-esteem than those in class 5 (balanced: me = male, me = science) (χ^2 = 5.49, p = 0.019) or in class 3 (unbalanced: me = male, me = liberal arts) (χ^2 = 5.49, p = 0.019) (χ^2 = 3.89, p = 0.048). The other classes did not differ significantly in level of self-esteem. Study motivation differed significantly between classes as well (χ^2 = 34.11, p < 0.001). Specifically, class 2 reported more study motivation than class 1 (χ^2 = 10.99, p = 0.001), class 3 (χ^2 = 9.87, p = 0.002), and class 5 (χ^2 = 9.99, p = 0.002). Similarly, class 4 reported more study motivation than class 1 (χ^2 = 8.99, p = 0.003), class 3 (χ^2 = 8.71, p = 0.003), and class 5 (χ^2 = 8.43, p = 0.004). Remember that classes 2 and 4 were balanced and mostly included women who felt similar to females, whereas classes 1 (unbalanced), 3 (unbalanced), and 5 (balanced) least likely included women and were characterized by felt similarity to males. The other class comparisons on study motivation were not significant. Study engagement did not differ between classes (χ^2 = 4.24, p = 0.375).

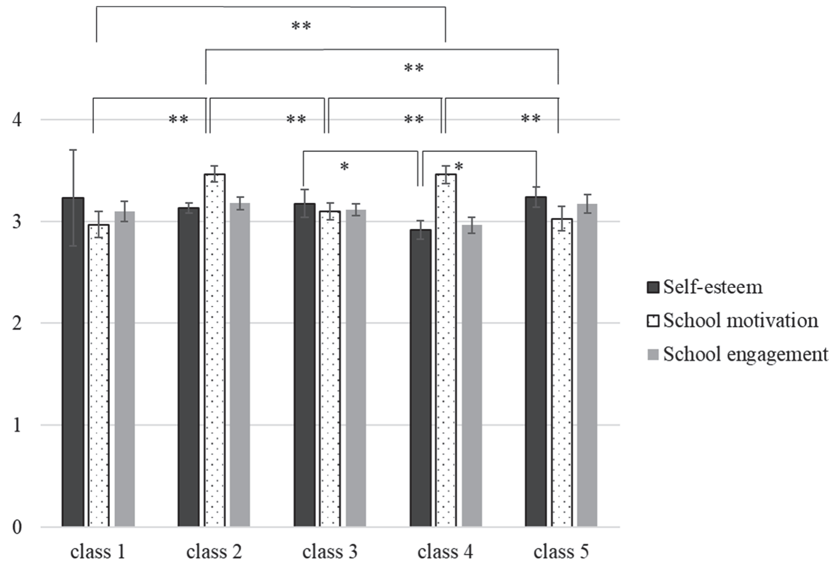


Figure 3. Mean Scores of the Five Classes on Self-Esteem, Study Motivation, and Engagement. Note. * $p < 0.05$. ** $p < 0.01$. Error bars represent standard errors of the mean.

4. Discussion

The first aim of this study was to examine whether three gender-science cognitions (i.e., implicit gender-science stereotypes, explicit gender identity, and explicit occupational self-concept) form a balanced identity configuration. Evidence was found for the stereotype-emulation hypothesis, the stereotype-construction hypothesis, and the identity-construction hypothesis [19]. First, in line with stereotype-emulation [19] there was a cross-over interaction between gender-science stereotypes and gender similarity. Emerging adults who felt similar to males and implicitly associated science with males reported more interest in science occupations, whereas people who felt similar to females and associated science with males reported more interest in liberal arts occupations. This indicates that in order to increase emerging adults' interest in occupational domains in which the gender they feel similar to is underrepresented, intervention programs could foster more egalitarian gender-science stereotypes or promote people's felt similarity with the other gender. Second, in line with stereotype-construction [19] a cross-over interaction was found between gender similarity and occupational self-concept. Emerging adults who felt similar to males and were interested in science occupations showed stronger science-with-male associations, whereas people who felt similar to females and were interested in science occupations showed stronger science-with-female associations. Finally, in line with identity-construction [19] there was an interaction between gender-science stereotypes and occupational self-concept. However, stronger science-with-male associations were associated with feeling more similar to females only for emerging adults with an interest in liberal arts occupations. For emerging adults with an interest in science occupations, gender-science stereotypes were not associated with felt similarity to males over females. This might indicate that other processes than identity construction contribute to the perceived gender similarity of people interested in science occupations, such as social pressures to conform to gender norms [36].

Overall, these three interaction effects showed that gender-science stereotypes, gender identity, and occupational self-concept were organized into a balanced identity configuration [10]. These findings extend previous research that found evidence for balanced identity across a range of different social cognition domains (e.g., race, gender, age, math/language, good/bad, work/family) [12,14,15], by showing that balanced identity is also present for gender-science cognitions and with a combination of explicit (self-report) and implicit measures. Furthermore, these findings suggest that the interplay between gender-science stereotypes, gender identity, and occupational self-concept might play an important role in educational and occupational gender segregation in STEM fields and the liberal arts. Programs aimed at increasing women's representation in STEM and men's representation in health care and education could focus on this triad of interrelated gender cognitions. Specifically, to foster perceived similarity with members of the other gender stimulating positive contact with the other-gender group in education might be a fruitful direction to take [63]. In addition, the use of counter-stereotypical role models might be effective in reducing gender-science stereotypes or increasing self-identification with occupations and studies dominated by the other gender [64].

This study did not find evidence for pure balanced identity, which asserts that the interaction between two social cognitions is the sole predictor of a third social cognition. This conclusion was supported by the additional analyses I conducted as proposed by Greenwald et al. [10,56] to test for pure balanced identity (see Supplementary Materials). Only eight of the 12 tests for pure balanced identity were met. A reason for the absence of a pure balance could be the combination of explicit and implicit measures to assess gender-science cognitions in the current study. Evidence for the balanced-congruity principle has been found more consistently with implicit measures than with explicit measures [65], and is stronger for implicit compared to explicit measures, possibly because of the presence of more error variance in self-report measures [15].

Regarding the second aim of the study, different profiles of balanced and unbalanced identity were identified. Two clearly balanced identity profiles were present, characterized by the following configurations of gender cognitions: (1) science = male, me = male, me = science; (2) science = male, me = female, me $\neq$ science. Another moderately balanced profile had the same configuration of gender cognitions as the second strongly balanced profile, but emerging adults' occupational interests and gender-science stereotypes were less pronounced. One strongly unbalanced identity profile was evident with emerging adults associating males-with-science, feeling similar to males, but being interested in liberal arts occupations. Another moderately disbalanced profile had the same configuration of gender-science cognitions as the strongly unbalanced profile, but emerging adults' occupational interest in liberal arts was less pronounced. As expected, the balanced identity profiles were more prevalent, including 66 percent of emerging adults. Yet, not all emerging adults achieved a balance in their social cognitions, which might be because emerging adults are still actively exploring their occupational identities and world views [22]. These findings further highlight the importance of focusing on individual differences or taking a person-centered approach in the study of balanced identity, as well as studying the developmental or temporal processes of achieving balanced identity with longitudinal designs.

Regarding the third study aim, the profiles of balanced and unbalanced identity that were identified differed meaningfully on several background variables. Not surprisingly, the profiles in which emerging adults strongly identified with females more likely included women compared to profiles characterized by strong identification with males. Similarly, profiles in which emerging adults had a strong interest in the liberal arts, being enrolled in a liberal arts major was more likely than in profiles characterized by a strong interest in science. The only differences between balanced and unbalanced identity profiles were found on the background variables "living with parents", "ethnicity" and "educational level". Specifically, emerging adults in the two unbalanced profiles more often lived by themselves compared to emerging adults with balanced identity profiles. Thus, different factors might contribute to an imbalance in social cognitions. Moving out of the parental home is a major life transition [22], and is often accompanied by a reconsideration or further exploration of one's identity [66], as well as exposure to a wider range of world views than in the nuclear family [67]. These changes may have led to a temporary disbalance in emerging adults' gender cognitions. Furthermore, emerging adults in the moderately unbalanced identity profile were more likely to be less educated and of non-Dutch ethnicity compared to the balanced identity profiles. Both lower educational level and a non-Western cultural background are associated with less egalitarian views about gender [33,34,68] and more gender-typical identity [24,35], which might decrease the likelihood of achieving a balanced identity. It should be noted that the interpretation of these differences is speculative, as there is little research yet on this topic to support these arguments. Future research is needed to examine underlying processes that can explain why certain groups of emerging adults develop unbalanced identity profiles.

Finally, regarding the fourth study aim, the identity profiles differed in level of study motivation and self-esteem. However, these differences were more attributable to gender identity differences than to differences between balanced and unbalanced identity profiles. Actually, profiles with emerging adults who felt similar to females scored higher on study motivation and lower on self-esteem than profiles with emerging adults who felt similar to males. The gender difference in self-esteem is well established (for meta-analyses, see [69]). For study motivation and engagement, evidence is only now emerging that girls score higher than boys [70,71]. The finding that balanced and unbalanced identity profiles did not differ in meaningful ways on self-esteem and study motivation and engagement might indicate that a momentary disbalance in one's gender cognitions is not necessarily detrimental for self-esteem or attitudes about education. Future research could examine whether a failure to restore balance in a triad of gender cognitions in the longer term is detrimental for people's self-esteem and persistence in academic and career pursuits, consistent with predictions from different identity theories [10,31,41].

4.1. Limitations and Future Directions

Some limitations of this study need to be considered. First, the design of this study was correlational. Therefore, no conclusions could be drawn about the direction of effects in the associations between the gender-science cognitions, as well as in the associations between the identity profiles and self-esteem and study motivation. Future longitudinal studies could examine the balanced-congruity principle more as a developmental process to provide further clarity on the direction of effects. Such research can also shed light on the consequences of temporary versus longer-term disbalance in social cognitions as well as yield insight into when each part of a balanced identity configuration emerges.

Second, the sample size might have been too small to detect less prevalent identity profiles. The number of possible balanced and unbalanced configurations of the triad of gender-science cognitions is larger than five. Identity profiles found in the current study therefore need to be confirmed in larger studies taking a person-centered approach to balanced identity.

Third, the measures used to assess self-esteem and study motivation were general in nature, even though these constructs appear to be domain- and course-specific [72] and might be better predictors of future academic and career pursuits than general measures [73,74].

Finally, a person-centered approach might provide difficulties with labelling profiles as balanced or unbalanced, because individuals within that profile, although similar in scores on the social cognitions, still show some within-profile variation. Future research could develop other ways to assess individual variation in balanced identity. In this way the consequences of balanced identity for future gender-typical academic and career pursuits could be studied more fully.

4.2. Conclusions

In sum, the current study found evidence for a balanced-identity configuration of gender-science stereotypes, gender identity, and occupational self-concept in emerging adults. Not all emerging adults showed this balanced configuration, as several balanced and unbalanced identity profiles could be distinguished. Emerging adults in unbalanced identity profiles were more likely to have left the parental home, be of non-Dutch ethnicity, and to have a lower educational level. Although balanced and unbalanced identity profiles did not differ in self-esteem and study motivation and engagement, it remains to be studied whether unbalanced identity profiles might have detrimental effects on people's career and academic pursuits in the longer term.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/educsci13040424/s1, Table S1: Hierarchical Regression Results for Occupational Interests (Science vs. Humanities), Predicted From Gender-Science Stereotypes and Gender Similarity; Table S2: Hierarchical Regression Results for Gender-Science Stereotypes, Predicted From Occupational Interests (Science vs. Humanities) and Gender Similarity; Table S3: Hierarchical Regression Results for Gender Similarity, Predicted From Occupational Interests (Science vs. Humanities) and Gender-Science Stereotypes; Table S4: Hierarchical Regression Results for Occupational Interests (Science vs. Humanities), Predicted From Gender-Science Stereotypes and Gender Similarity; Table S5: Hierarchical Regression Results for Gender-Science Stereotypes, Predicted From Occupational Interests (Science vs. Humanities) and Gender Similarity; Table S6: Hierarchical Regression Results for Gender Similarity, Predicted From Occupational Interests (Science vs. Humanities) and Gender-Science Stereotypes; Table S7: ANOVA Results and Post Hoc Comparisons Across the 5 Classes.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of the Faculty of Social and Behavioural Sciences of Utrecht University (FETC15-048, 18 September 2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data is publicly available on https://osf.io/c86uj/.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Cortes, P.; Pan, J. *Occupation and Gender*; IZA Discussion Papers, No. 10672; Institute of Labor Economics (IZA): Bonn, Germany, 2017. Available online: http://hdl.handle.net/10419/161295 (accessed on 1 July 2022).
2. OECD. *Education at a Glance 2017: OECD Indicators*; OECD Publishing: Paris, France, 2017. [CrossRef]
3. OECD. *The Pursuit of Gender Equality: An Uphill Battle*; OECD Publishing: Paris, France, 2017. [CrossRef]
4. Fry, R.; Kennedy, B.; Funk, C. *STEM Jobs See Uneven Progress in Increasing Gender, Racial and Ethnic Diversity*; Pew Research Center: Washington, DC, USA, 2021.
5. Croft, A.; Schmader, T.; Block, K. An underexamined inequality: Cultural and psychological barriers to men's engagement with communal roles. *Personal. Soc. Psychol. Rev.* **2015**, *19*, 343–370. [CrossRef] [PubMed]
6. International Labour Organization. Female Share of Employment in Managerial Positions. 2019. Available online: https://www.ilo.org/shinyapps/bulkexplorer3/?lang=en&segment=indicator&id=SDG_0552_OCU_RT_A (accessed on 1 September 2021).
7. Bear, J.B.; Woolley, A.W. The role of gender in team collaboration and performance. *Interdiscip. Sci. Rev.* **2011**, *36*, 146–153. [CrossRef]
8. Diekman, A.B.; Clark, E.K.; Belanger, A.L. Finding common ground: Synthesizing divergent theoretical views to promote women's stem pursuits. *Soc. Issues Policy Rev.* **2019**, *13*, 182–210. [CrossRef]
9. Wang, M.T.; Degol, J. Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* **2013**, *33*, 304–340. [CrossRef] [PubMed]
10. Greenwald, A.G.; Banaji, M.R.; Rudman, L.A.; Farnham, S.D.; Nosek, B.A.; Mellott, D.S. A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychol. Rev.* **2002**, *109*, 3–25. [CrossRef]
11. Martin, C.L.; Andrews, N.C.; England, D.E.; Zosuls, K.; Ruble, D.N. A dual identity approach for conceptualizing and measuring children's gender identity. *Child Dev.* **2017**, *88*, 167–182. [CrossRef]
12. Nosek, B.A.; Banaji, M.R.; Greenwald, A.G. Math= male, me= female, therefore math≠ me. *J. Personal. Soc. Psychol.* **2002**, *83*, 44–59. [CrossRef]
13. Martin, C.L.; Halverson, C.F. The roles of cognition in sex role acquisition. In *Current Conceptions of Sex Roles and Sex Typing: Theory and Research*; Carter, D.B., Ed.; Praeger: New York, NY, USA, 1987; pp. 123–137.
14. Cvencek, D.; Meltzoff, A.N.; Greenwald, A.G. Math-gender stereotypes in elementary school children. *Child Dev.* **2011**, *82*, 766–779. [CrossRef] [PubMed]
15. Cvencek, D.; Meltzoff, A.N.; Maddox, C.D.; Nosek, B.A.; Rudman, L.A.; Devos, T.; Dunham, Y.; Baron, A.S.; Steffens, M.C.; Lane, K.; et al. Meta-analytic use of balanced identity theory to validate the Implicit Association Test. *Personal. Soc. Psychol. Bull.* **2021**, *47*, 185–200. [CrossRef]

16. Starr, C.R.; Leaper, C. Do adolescents' self-concepts moderate the relationship between STEM stereotypes and motivation? *Soc. Psychol. Educ.* **2019**, *22*, 1109–1129. [CrossRef]
17. Cundiff, J.L.; Vescio, T.K.; Loken, E.; Lo, L. Do gender–science stereotypes predict science identification and science career aspirations among undergraduate science majors? *Soc. Psychol. Educ.* **2013**, *16*, 541–554. [CrossRef]
18. Lane, K.A.; Goh, J.X.; Driver-Linn, E. Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles* **2012**, *66*, 220–234. [CrossRef]
19. Tobin, D.D.; Menon, M.; Menon, M.; Spatta, B.C.; Hodges, E.V.; Perry, D.G. The intrapsychics of gender: A model of self-socialization. *Psychol. Rev.* **2010**, *117*, 601–622. [CrossRef]
20. Festinger, L. *A Theory of Cognitive Dissonance*; Stanford University Press: Stanford, CA, USA, 1957; Volume 2.
21. Heider, F. *The Psychology of Interpersonal Relations*; Wiley: New York, NY, USA, 1958.
22. Arnett, J.J. Emerging adulthood: A theory of development from the late teens through the twenties. *Am. Psychol.* **2000**, *55*, 469–480. [CrossRef] [PubMed]
23. Endendijk, J.J.; Derks, B.; Mesman, J. Does parenthood change implicit gender-role stereotypes and behaviors? *J. Marriage Fam.* **2018**, *80*, 61–79. [CrossRef]
24. Endendijk, J.J.; Andrews, N.C.; England, D.E.; Martin, C.L. Gender-identity typologies are related to gender-typing, friendships, and social-emotional adjustment in Dutch emerging adults. *Int. J. Behav. Dev.* **2019**, *43*, 322–333. [CrossRef]
25. Hsu, N.; Badura, K.L.; Newman, D.A.; Speach ME, P. Gender, "masculinity", and "femininity": A meta-analytic review of gender differences in agency and communion. *Psychol. Bull.* **2021**, *147*, 987–1011. [CrossRef]
26. Hoff, K.A.; Briley, D.A.; Wee, C.J.; Rounds, J. Normative changes in interests from adolescence to adulthood: A meta-analysis of longitudinal studies. *Psychol. Bull.* **2018**, *144*, 426–451. [CrossRef]
27. Meca, A.; Ritchie, R.A.; Beyers, W.; Schwartz, S.J.; Picariello, S.; Zamboanga, B.L.; Hardy, S.A.; Luyckx, K.; Kim, S.Y.; Whitbourne, S.K.; et al. Identity centrality and psychosocial functioning: A person-centered approach. *Emerg. Adulthood* **2015**, *3*, 327–339. [CrossRef]
28. Schachter, E.P. Identity configurations: A new perspective on identity formation in contemporary society. *J. Personal.* **2004**, *72*, 167–200. [CrossRef]
29. Cvencek, D.; Meltzoff, A.N.; Kapur, M. Cognitive consistency and math–gender stereotypes in Singaporean children. *J. Exp. Child Psychol.* **2014**, *117*, 73–91. [CrossRef] [PubMed]
30. Erikson, E. *Youth: Identity and Crisis*; Norton: New York, NY, USA, 1968. [CrossRef]
31. Kohlberg, L. A cognitive-developmental analysis of children's sex-role concepts and attitudes. In *The Development of Sex Differences*; Maccoby, E.E., Ed.; Stanford University Press: Stanford, CA, USA, 1966; pp. 82–173.
32. Pedersen, R.M. The Relationship between Gender and Implicit and Explicit Balanced STEM Identity Profiles. Master's Thesis, Texas A&M University, College Station, TX, USA, 2021. Available online: https://srw.tamu.edu/wp-content/uploads/2021/03/2021-SRW-Poster-Major_Gender-Differences-Implicit_Explicit-Balance.pdf (accessed on 1 March 2022).
33. Dodson, T.A.; Borders, L.D. Men in traditional and non-traditional careers: Gender role attitudes, gender role conflict, and job satisfaction. *Career Dev. Q.* **2006**, *54*, 283–296. [CrossRef]
34. Harris, R.J.; Firestone, J.M. Changes in predictors of gender role ideologies among women: A multivariate analysis. *Sex Roles* **1998**, *38*, 239–252. [CrossRef]
35. Marcell, A.V.; Eftim, S.E.; Sonenstein, F.L.; Pleck, J.H. Associations of family and peer experiences with masculinity attitude trajectories at the individual and group level in adolescent and young adult males. *Men Masc.* **2011**, *14*, 565–587. [CrossRef]
36. Cook, R.E.; Nielson, M.G.; Martin, C.L.; DeLay, D. Early adolescent gender development: The differential effects of felt pressure from parents, peers, and the self. *J. Youth Adolesc.* **2019**, *48*, 1912–1923. [CrossRef]
37. Nielson, M.G.; Schroeder, K.M.; Martin, C.L.; Cook, R.E. Investigating the relation between gender typicality and pressure to conform to gender norms. *Sex Roles* **2020**, *83*, 523–535. [CrossRef]
38. Field, R.D.; Martin, C.L.; Andrews, N.C.; England, D.E.; Zosuls, K.M. The influence of gender-based relationship efficacy on attitudes toward school. *Merrill-Palmer Q.* **2017**, *63*, 396–422. [CrossRef]
39. Fredricks, J.A.; Blumenfeld, P.; Friedel, J.; Paris, A. School Engagement. In *What Do Children Need to Flourish? The Search Institute Series on Developmentally Attentive Community and Society*; Moore, K.A., Lippman, L.H., Eds.; Springer: Boston, MA, USA, 2005; Volume 3. [CrossRef]
40. Greenwald, A.G.; McGhee, D.E.; Schwartz, J.L. Measuring individual differences in implicit cognition: The implicit association test. *J. Personal. Soc. Psychol.* **1998**, *74*, 1464–1480. [CrossRef]
41. Adams, G.R.; Marshall, S.K. A developmental social psychology of identity: Understanding the person-in-context. *J. Adolesc.* **1996**, *19*, 429–442. [CrossRef]
42. Crocetti, E.; Avanzi, L.; Hawk, S.T.; Fraccaroli, F.; Meeus, W. Personal and social facets of job identity: A person-centered approach. *J. Bus. Psychol.* **2014**, *29*, 281–300. [CrossRef]
43. World Economic Forum. Global Gender Gap Report 2021. 2021. Available online: http://www3.weforum.org/docs/WEF_GGGR_2021.pdf (accessed on 1 September 2022).
44. Miller, D.I.; Eagly, A.H.; Linn, M.C. Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *J. Educ. Psychol.* **2015**, *107*, 631–644. [CrossRef]

45. Kvitkovičová, L.; Umemura, T.; Macek, P. Roles of attachment relationships in emerging adults' career decision-making process: A two-year longitudinal research design. *J. Vocat. Behav.* **2017**, *101*, 119–132. [CrossRef]
46. Barrett, A.E.; White, H.R. Trajectories of gender role orientations in adolescence and early adulthood: A prospective study of the mental health effects of masculinity and femininity. *J. Health Soc. Behav.* **2002**, *43*, 451–468. [CrossRef] [PubMed]
47. Nylund, K.L.; Asparouhov, T.; Muthén, B.O. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Struct. Equ. Model.* **2007**, *14*, 535–569. [CrossRef]
48. Tein, J.Y.; Coxe, S.; Cham, H. Statistical power to detect the correct number of classes in latent profile analysis. *Struct. Equ. Model.* **2013**, *20*, 640–657. [CrossRef]
49. Weller, B.E.; Bowen, N.K.; Faubert, S.J. Latent class analysis: A guide to best practice. *J. Black Psychol.* **2020**, *46*, 287–311. [CrossRef]
50. Nosek, B.A.; Smyth, F.L.; Sriram, N.; Lindner, N.M.; Devos, T.; Ayala, A.; Bar-Anan, Y.; Bergh, R.; Cai, H.; Gonsalkorale, K.; et al. National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10593–10597. [CrossRef]
51. Greenwald, A.G.; Nosek, B.A.; Banaji, M.R. Understanding and using the implicit association test: I. An improved scoring algorithm. *J. Personal. Soc. Psychol.* **2003**, *85*, 197–216. [CrossRef]
52. Harmon, L.W.; DeWitt, D.W.; Campbell, D.P.; Hansen, J.I.C. *Strong Interest Inventory: Applications and Technical Guide: Form T317 of the Strong Vocational Interest Blanks*; Stanford University Press: Stanford, CA, USA, 1994.
53. Rosenberg, M. *Rosenberg Self-Esteem Scale*; Basic Books: New York, NY, USA, 1979.
54. Eccles, J.; Wigfield, A.; Harold, R.D.; Blumenfeld, P. Age and gender differences in children's self- and task perceptions during elementary school. *Child Dev.* **1993**, *64*, 830–847. [CrossRef]
55. Blanton, H.; Jaccard, J. Tests of multiplicative models in psychology: A case study using the unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychol. Rev.* **2006**, *113*, 155–166. [CrossRef]
56. Greenwald, A.G.; Rudman, L.A.; Nosek, B.A.; Zayas, V. Why so little faith? A reply to Blanton and Jaccard's (2006) skeptical view of testing pure multiplicative theories. *Psychol. Rev.* **2006**, *113*, 170–180. [CrossRef]
57. Muthén, L.K.; Muthén, B.O. *Mplus User's Guide*, 8th ed.; Muthén & Muthén: Los Angeles, CA, USA, 2017; Available online: https://www.statmodel.com/HTML_UG/introV8.htm (accessed on 1 February 2020).
58. Oberski, D.L. Mixture models: Latent profile and latent class analysis. In *Modern statistical methods for HCI*; Robertson, J., Kaptein, M., Eds.; Human-Computer Interaction Series; Springer: Berlin/Heidelberg, Germany, 2016; pp. 275–287. [CrossRef]
59. Tueller, S.; Lubke, G. Evaluation of structural equation mixture models: Parameter estimates and correct class assignment. *Struct. Equ. Model.* **2010**, *17*, 165–192. [CrossRef] [PubMed]
60. Vermunt, J.K. Latent class modeling with covariates: Two improved three-step approaches. *Political Anal.* **2010**, *18*, 450–469. [CrossRef]
61. Asparouhov, T.; Muthén, B. Auxiliary variables in mixture modeling: Three-step approaches using M plus. *Struct. Equ. Model. Multidiscip. J.* **2014**, *21*, 329–341. [CrossRef]
62. Bakk, Z.; Vermunt, J.K. Robustness of stepwise latent class modeling with continuous distal outcomes. *Struct. Equ. Model.* **2016**, *23*, 20–31. [CrossRef]
63. Halim ML, D.; Martin, C.L.; Andrews, N.C.; Zosuls, K.M.; Ruble, D.N. Enjoying each other's company: Gaining other-gender friendships promotes positive gender attitudes among ethnically diverse children. *Personal. Soc. Psychol. Bull.* **2021**, *47*, 1635–1653. [CrossRef]
64. Olsson, M.; Martiny, S.E. Does exposure to counterstereotypical role models influence girls' and women's gender stereotypes and career choices? A review of social psychological research. *Front. Psychol.* **2018**, *9*, 2264. [CrossRef]
65. Cvencek, D.; Greenwald, A.G.; Meltzoff, A.N. Balanced identity theory: Evidence for implicit consistency in social cognition. In *Cognitive Consistency: A Unifying Concept in Social Psychology*; Gawronski, B., Strack, F., Eds.; Guilford Press: New York, NY, USA, 2012; pp. 157–177.
66. Jones, G.W. A risky business: Experiences of leaving home among young rural women. *J. Youth Stud.* **2004**, *7*, 209–220. [CrossRef]
67. Gutierrez, I.A.; Park, C.L. Emerging adulthood, evolving worldviews: How life events impact college students' developing belief systems. *Emerg. Adulthood* **2015**, *3*, 85–97. [CrossRef]
68. Kavli, H.C. Adapting to the dual earner family norm? The case of immigrants and immigrant descendants in Norway. *J. Ethn. Migr. Stud.* **2015**, *41*, 835–856. [CrossRef]
69. Zuckerman, M.; Li, C.; Hall, J.A. When men and women differ in self-esteem and when they don't: A meta-analysis. *J. Res. Personal.* **2016**, *64*, 34–51. [CrossRef]
70. Bugler, M.; McGeown, S.P.; St Clair-Thompson, H. Gender differences in adolescents' academic motivation and classroom behaviour. *Educ. Psychol.* **2015**, *35*, 541–556. [CrossRef]
71. Martin, A.J. Examining a multidimensional model of student motivation and engagement using a construct validation approach. *Br. J. Educ. Psychol.* **2007**, *77*, 413–440. [CrossRef]
72. Green, J.; Martin, A.J.; Marsh, H.W. Academic Motivation and Engagement: A Domain Specific Approach. In Proceedings of the Australian Association for Research in Education Annual Conference, Paramatta, Australia, 27 November–1 December 2005; Available online: http://www.aare.edu.au/data/publications/2005/gre05384.pdf (accessed on 1 March 2022).

73. Durik, A.M.; Vida, M.; Eccles, J.S. Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *J. Educ. Psychol.* **2006**, *98*, 382–393. [CrossRef]
74. Lauermann, F.; Chow, A.; Eccles, J.S. Differential effects of adolescents' expectancy and value beliefs about math and English on math/science-related and human services-related career plans. *Int. J. Gend. Sci. Technol.* **2015**, *7*, 205–228.

education sciences

MDPI

Article

Academic Self-Efficacy and Value Beliefs of International STEM and Non-STEM University Students in Germany from an Intersectional Perspective

Judith Sarah Preuß *, Julia Zimmermann and Kathrin Jonkmann

Department of Educational Psychology, Faculty of Psychology, FernUniversität in Hagen, 58097 Hagen, Germany; julia.zimmermann@fernuni-hagen.de (J.Z.); kathrin.jonkmann@fernuni-hagen.de (K.J.)
* Correspondence: judith.preuss@fernuni-hagen.de

Abstract: The expectancy–value theory (EVT) positions expectations and value beliefs as important predictors of academic success. We, thus, investigated the prevalence of academic self-efficacy and intrinsic, attainment, utility, and cost values amongst international STEM students in Germany, as well as their associations with gender, parental academic background, cultural characteristics, and their interplay. We also compared STEM to non-STEM students. Analyses with hierarchical multi-group regression models using data from 1590 international bachelor students (n_{STEM} = 882, $n_{\text{non-STEM}}$ = 708) revealed high levels of academic self-efficacy, attainment, intrinsic, and utility values but also high costs. International STEM students indicated lower levels of academic self-efficacy than non-STEM students; all other results were similar in both subject-groups. There were no direct associations between gender and the expectancy–value components but continuous-generation students showed higher academic self-efficacy than first-generation students. Significant associations between cultural background and all expectancy–value components were identified, most of them applied to costs. In some cases, the associations differed by gender. Study-related language skills were related to all expectancy–value components whereas host- and home-culture orientations were distinctly associated with attainment, intrinsic, and utility values. Implications of the results for interventions supporting the academic success of international STEM students and future research needs are discussed.

Keywords: international students; academic self-efficacy and value beliefs; (situated) expectancy–value theory (SEVT); STEM students; intersectionality; multi-group regression analyses

Citation: Preuß, J.S.; Zimmermann, J.; Jonkmann, K. Academic Self-Efficacy and Value Beliefs of International STEM and Non-STEM University Students in Germany from an Intersectional Perspective. *Educ. Sci.* **2023**, *13*, 786. https://doi.org/10.3390/educsci13080786

Academic Editors: Erin Mackenzie, Bernhard Ertl, Christine R. Starr, Silke Luttenberger and Manuela Paechter

Received: 12 May 2023
Revised: 4 July 2023
Accepted: 27 July 2023
Published: 2 August 2023

1. Introduction

A comprehensive answer to the Special Issue's leading question of who is sticking with STEM requires the consideration of many aspects. For example, it is important to consider who comes to study STEM subjects at German universities and to explore these students' (psychological) prerequisites for study success that may influence who stays and who goes in due course.

An important group amongst the STEM students at German universities constitutes international students. These are foreign students who obtained their university entrance qualification outside of their study country [1,2]. In 2020, approximately 4.4 million international students were enrolled in the OECD which represents about 7% of the higher education students in these states [3]. About 369,000 international students studied at German universities, i.e., 11% of the student body [3]. Nearly 52% of the international students in Germany were enrolled in STEM subjects [4]. Hence, international students could help mitigate the shortage of skilled workers in STEM fields (e.g., in Germany; see Hoffmeyer-Zlotnik and Grote [5]). Yet, unfortunately, many international students struggle with their studies as the high dropout rates of 41% in the bachelor's and 28% in master's programs (compared to 28% and 21% among domestic students) show [6]. Against this background, understanding the (psychological) prerequisites of international STEM students' academic success is essential.

A prominent psychological approach to explain interindividual differences in academic choices and success is the (*situated*) *expectancy–value theory* (EVT [7]; SEVT [8]). It suggests that study-related expectations and value beliefs are building blocks in explaining why some individuals succeed whilst others struggle or drop out of their studies. Yet, previous meta-analytic research showed that learners differ in their expressions of these psychological constructs by their demographic characteristics such as gender, parental academic background, cultural background, and their interplay [9]. Accordingly, a thorough assessment of individual differences in international STEM students' study-related expectations and value beliefs at the beginning of their studies in Germany is an important endeavor in order to identify student groups who might be at risk for (academic) struggles throughout the further study course.

In the present study, we addressed the manifestation of expectations and value beliefs amongst international STEM students. In doing so, we assessed how demographic (gender and parental academic background) and cultural characteristics (cultural background, study-related language skills, previous residence in Germany, and acculturation orientations), as well as the intersectional interplay (e.g., [10,11]) between gender and parental academic and cultural background, relate to these motivational constructs at the beginning of the students' degree courses in Germany. In doing so, we aim to contribute to a better understanding of the heterogeneity within the (international) student body with regard to their (psychological) prerequisites for STEM success. This may help to identify target groups for tailored interventions to prevent study dropout and, thus, shed light on the question "Sticking with STEM: Who comes, who stays, who goes, and why?".

1.1. Expectancy and Value Beliefs

The EVT [7] explains behavior and behavioral intentions using expectancy and value components which are influenced by one's own experiences, cultural norms, and the environment, e.g., family and peers [12]. The expectancies (for success) and subjective task values of individuals are postulated as the most relevant psychological predictors of task choice, performance, and effort in the chosen tasks [7]. The EVT was extended to the situated expectancy–value theory (SEVT) [8]. Eccles and Wigfield [8] described this extension as a situated and cultural view on the EVT model, pointing to the relevance of the specific situation, i.e., environmental conditions and the cultural background, for the development of expectancy and value hierarchies.

The first central factor of the EVT, expectancy, incorporates different theoretical conceptions of self-belief such as (academic) self-concepts [13,14] and self-efficacy [15,16]. For both constructs. there is evidence that they adequately represent the expectancy factor (i.e., Eccles and Wigfield [17]). In the present study, we focused on self-efficacy. Bandura [15,18] defined general self-efficacy within the social cognitive learning theory as a person's assessment of the extent to which he or she expects to be able to master tasks and challenges or to achieve goals. Self-efficacy can be understood as a generalized construct, but can also refer to specific domains, e.g., academic self-efficacy.

Academic self-efficacy describes students' beliefs that they can control and reflect their study behaviors, understand and successfully process learning material, and meet academic requirements [19]. The *Social Cognitive Career Theory* (SCCT) [20] drew on this social cognitive framework to explain three closely related aspects of education and career development: (a) development of career-relevant interests, (b) selection of study and career choices, and (c) performance and persistence in education and occupation. According to the SCCT, the development of STEM interests before the study starts, the choice of a STEM subject, and the academic success in STEM are predicted by individuals' self-efficacy beliefs in their interplay with the environment [19,20].

Pertaining to the second factor of EVT, value beliefs, four value components are differentiated: attainment value (importance of success for self-image, identification with the subject and school in general), intrinsic value (interest and enjoyment of the task), utility value (utility for short- and long-term personal goals), and cost value (potential costs of investing time in this activity rather than another).

Both expectations and value beliefs guide individuals, their educational achievements, and related decisions both consciously and unconsciously [21–23]. Wigfield and Eccles [23] reported accumulated findings from more than 35 years of research which support these assumptions. Results from the secondary school context revealed that expectations and value beliefs, as well as their interaction, predicted (math) achievement [24,25]. Former studies corroborated that individuals are more likely to choose STEM studies if they show higher STEM-related self-efficacy and ability self-concepts [26] (for a general review, see [27]). Furthermore, positive value beliefs with regard to mathematics and science among adolescents at the beginning of high school were associated with STEM achievement, as well as STEM subject choices, not only in high school, but also 7 years later in college [28].

Yet, even if the decision to study a STEM subject has already been made, differences in STEM students' self-efficacy and value beliefs, e.g., due to their perceived minority status as a female or first-generation student, occur and should not be neglected, as they were shown to be relevant for study success [9,29,30]. For example, a study with a sample of first-year engineering undergraduates showed that higher levels of engineering self-efficacy and highly valuing engineering were related to higher engineering persistence [31].

To conclude, empirical findings corroborate the importance of expectations and value beliefs (as well as their interplay) in the prediction of academic choices and success in STEM fields and beyond. As a consequence, investigating both the level of and the interindividual differences in these constructs amongst international STEM students at the beginning of their studies is important to identify student groups who might be at risk of struggling in or dropping out of their degree programs.

1.2. Differences in Expectations and Value Beliefs by Gender and Parental Academic Background

Gender. One of the purposes of the EVT [7] is to explain gender differences in STEM-related academic attitudes and decisions such as subject choices.

Indeed, in the university context, female students remain to be underrepresented in many STEM subjects. In the US—despite differences between subjects (e.g., 19% female students in engineering-technology and computer science vs. 39% in physics)—the overall proportion of 35% female students in STEM corroborates the gender imbalance [2]. Similarly, in Germany, female students remain to be underrepresented in STEM subjects [32] with, for example, only 16% females in engineering and 21% in computer science [32,33]. We may, thus, assume that the female international STEM students in Germany represent a selective sample who chose to study STEM despite the unpopularity of these subjects amongst their same gender fellows. This may have implications with regard to gender differences in their study-related expectations and value beliefs. Yet, to the best of our knowledge, there is no systematic research on gender differences in expectations and value beliefs amongst international STEM students in Germany and beyond, although many studies addressed these issues in different stages of the educational career. For example, a meta-analysis of 187 studies by Huang [29] identified gender differences in academic self-efficacy that were moderated by subject area. Males showed higher levels of self-efficacy in mathematics, computer science, and science than females. Importantly, gender differences in academic self-efficacy also varied with age. The largest effect size occurred for respondents who were over 23 years old, which supports their importance in higher education.

A more recent meta-analysis (n = 176) revealed gender differences in all expectancy–value components but costs (mostly because there were not enough studies available) across different subject fields [9]. In particular, males showed higher levels of expectancy for success in math, science, computing, engineering, and physical sciences (small to medium effect sizes). In biological sciences, there were no gender differences in expectancy for success. With regard to the value components, the results showed higher levels of intrinsic value for males in math, science, computing, physical sciences, and engineering (medium effect sizes), whereas females reported higher intrinsic values with regard to biological sciences. Yet, for utility and attainment value, gender differences were less obvious; males only indicated higher levels of utility value in computing. Moreover, males showed higher overall task value in computing and math, whereas girls showed higher overall task value in physical sciences (all small effect sizes). All other gender differences were negligible. With regard to the cost component, the low number of studies and effect sizes limited the results. The only substantial effect sizes were observed for physical science, where females reported higher costs (medium effect size). Overall, gender differences were weaker for general science expectancy/value than they were in more precisely defined areas of STEM (e.g., physical or biological sciences). With regard to the interaction of gender and other demographic characteristics, only one significant interaction, for gender and age (categorized into elementary school, middle school, high school, and (young) adult samples) was observed, i.e., the gender gap in expectancy for success was larger for older samples in more advanced stages of the educational career [9].

To conclude, both meta-analyses [9,29] suggested that gender differences in STEM-related expectancies and values in the favor of males are prevalent in all educational stages. Surprisingly, the findings suggest that gender differences gain importance in more advanced stages of the educational career despite potential self-selection mechanisms with regard to university subject choices. These considerations led to our investigation of the relationship between gender and expectancy–value components at the beginning of international STEM students study time abroad.

Parental academic background. Previous research on the role of the parental academic background with regard to academic decisions and success suggested that students' generational status, i.e., being a first-generation or continuous-generation student, was decisive [34,35]. Being the first in the family who is studying at a higher education institution contains several barriers with regard to academic success, e.g., due to missing role models, knowledge about the campus life, or a lack of university belonging. Several studies substantiated advantages of being a continuous-generation student with regard to the academic success of students with and without an immigrant background in Germany [36–38].

Yet, few studies considered the association between parental academic background and study-related expectancies and values. Findings from the school context by Gaspard et al. [39] showed associations of parental academic background with ability self-concept and task value (attainment) in math; high school students with at least one parent who holds an academic degree reported a higher academic self-concept and higher task (attainment) values. Former research using a sample of secondary education students in Germany showed significant correlations between socioeconomic status (which included parental academic background) and all expectancy and value components. In particular, positive associations for expectancy, attainment, intrinsic, and utility value (r = .15 to .11) and a negative one for cost value ($r = -.09$) were identified [40]. Accordingly, Goldman et al. [41] revealed that first-generation college students indicated higher levels of cost value and that these more strongly increased during the semester as compared to continuous-generation students.

Overall, these few findings suggest that coming from a non-academic household may have negative consequences with regard to international students' (STEM) expectations and values. Yet, further empirical research to evaluate these tentative assumptions is needed.

1.3. Expectations and Value Beliefs and Cultural Characteristics

The (S)EVT and related research postulated that cultural norms, such as individualism and collectivism, influence expectations and value beliefs [8,12,21,42,43]. Accordingly, previous studies showed that gender differences in STEM attitudes and success varied between European countries [44,45]. The authors explained these differences with references to cultural norms. Furthermore, Donohue [46] showed that cultural norms predicted the intrinsic value of undergraduate university students in the United States. Mok et al. [47] reported associations between cultural norms and the utility value of study programs amongst a sample of international students in Germany. Hence, we deemed it was important to consider students' cultural background as an external predictor of the EVT factors and included students' countries of origin (summarized to country groups) into our analyses. Former studies also showed a strong relation between language skills and international students' study success and emphasized their importance for the academic (and social) integration (e.g., Wisniewski et al. [48]). Language skills may help students to feel more confident to succeed in the study program, thereby increasing their task enjoyment and motivation to persist. Likewise, as language skills may reduce (cultural) barriers and facilitate participation in campus life, psychological costs such as the experience of loneliness and stress may decrease. Therefore, a measure of self-perceived study-related language skills was included in the analyses to analyze potential effects of these skills on international students' expectancies and value beliefs. Previous experiences in the host country were shown to have an impact on cultural transition experiences [49,50]. In particular, previous experiences in Germany might influence expectations of students to be able to handle (academic) challenges and thereby increase their academic self-efficacy. Furthermore, knowing what to expect in Germany may reduce fear and stress during the acculturation process, thus increasing enjoyment and decreasing psychological costs. Hence, we additionally considered previous residence in Germany.

Beyond the general importance of cultural norms and differences, the EVT also emphasizes that individual differences in approaching a culture may influence the formation of expectancy and value beliefs [8,42]. In addition, acculturation theory suggests that, for international students, differences in their approaches to both their host- and their home-culture are essential. According to Berry [51], such differences are captured in individuals' acculturation attitudes, i.e., their attitudes toward and interest in the involvement with the home- and the host-culture. Importantly, home- and host-culture orientation are considered to reflect two independent constructs [51,52]. Home- and host-culture orientations correlated positively with sociocultural adaption in the school context and beyond (e.g., Berry [51]). This suggests its relevance with regard to the formation of education-related expectancies and values.

However, a recent meta-analysis by Bierwiaczonek and Kunst [53] indicated that the relationship between host- and home-culture acculturation and adaptation might be weaker and less consistent than assumed. The only relationship that was consistently maintained even in longitudinal designs was the positive association between host-culture orientation and sociocultural adaptation. Yet, the authors concluded that the heterogeneity and instability of effects might be due to the heterogeneity in samples, e.g., with regard to the duration of their stay abroad. For example, it could be that acculturation orientations are more relevant regarding the sociocultural adaptation in the early stages of migration (e.g., at the beginning of a study program abroad) and become less relevant at later stages. The authors, thus, called for further studies on the relationship between acculturation orientations and sociocultural outcomes which—in the present case—are captured by students' study-related expectancies and values [53]. In line with previous findings, we speculated that a higher host-culture orientation could probably serve as a kind of resource during the adaption process and reduce the perceived costs of studying abroad during the adaption of international students in Germany.

1.4. Intersectionality

Intersectionality is defined as the interaction of multiple characteristics on diverse dimensions, which could build the base for discrimination or the experience of disadvantages [54,55]. Hence, the effects of these dimensions are developed in interaction/intersection, not just added up [10,11]. Results of studies in the US revealed that the confluence of multiple minority status-generating characteristics (e.g., gender, ethnicity, and parental academic background) conditioned lower levels of academic achievement in the school and higher education sector [56,57]. Most current studies on the interaction of gender, educational background, and ethnic/cultural background characteristics were conducted in the US. Most studies based on European data primarily examine the school sector [44,45]. With regard to higher education in Germany, intersectionality has mainly been considered at a descriptive level [38].

This unequal distribution of intersectional research between the US and Europe was also noted by Parker et al. [9]. Their meta-analysis provides the most comprehensive research work to date that considered an intersectional perspective on students' expectancy and value beliefs in different academic domains. Parker et al. [9] pointed out that gender influences the expectancy–value components not only alone, but also in interaction with multiple characteristics. In particular, their analysis revealed interactions of gender and social class (operationalized by socioeconomic status) in the math domain for expectancy, intrinsic value, and (by trend) for utility value. In all cases, the gender effect sizes were highest in high-socioeconomic status samples. Similar patterns were found for science expectancy. Yet, in view of the limited data on these interactions, the results need to be interpreted with caution and require further research.

Beyond these meta-analytic findings, a study that considered samples from 65 universities in the US revealed that female first-generation students reported significantly lower levels of self-efficacy than female continuous-generation students and male students, regardless of their academic family background [58]. In contrast, Else-Quest et al. [59] reported that, although gender gaps in self-beliefs and value beliefs in science and math slightly varied among Latin, Asian American, and Caucasian 10th grade high school students in the US, these differences tended not to be statistically significant. Seo et al. [60] used a similar but nationally representative sample of US 10th graders and showed that male adolescents' math self-concept was more positive than females among Whites and Latinxs but not among Blacks and Asians.

Given the heterogeneity and small number of previous findings, the current research considered intersectional effects between gender and students' academic and cultural background from an intersectional perspective to allow for a more specific consideration of potential disadvantage and dropout risks of specific student groups.

2. Purposes of the Present Study

The (S)EVT and previous findings emphasize that expectancy and value beliefs are important determinants of academic success [21–23]. However, to the best of our knowledge, there are no studies that specifically addressed expectations and value beliefs of international STEM students. In order to close this research gap, we examined how gender, parental academic background, and cultural characteristics (i.e., cultural background, study-related language skills, previous residence in Germany, and acculturation orientations) are related to the academic self-efficacy and value beliefs of a large sample of international STEM students. Moreover, despite our focus on students in STEM subjects, the design allows us to compare the pattern of results between STEM and non-STEM subjects; we can, thus, infer if any findings are indeed specific to the situation of international students in STEM or reflect a more general pattern.

On the basis of previous findings, we formulated the hypotheses below.

2.1. Relations with Gender

H1a. *Female international STEM students show lower levels of academic self-efficacy than male international STEM students.*

H1b. *Female international STEM students show higher levels of cost value than male international STEM students, but lower levels of intrinsic, attainment and utility values.*

2.2. Relations with Parental Academic Background

H2a. *International STEM students who are first-generation students report lower levels of academic self-efficacy than international STEM students who are continuous-generation students.*

H2b. *International STEM students who are first-generation students report higher levels of costs than international STEM students who are continuous-generation students. We did not assume differences in intrinsic, attainment and utility values, but assessed these for explorative reasons.*

2.3. Interactions of Gender and Parental Academic Background

H3a. *Female international STEM students who are first-generation students report lower levels of academic self-efficacy than female international STEM students who are continuous-generation students and than male STEM students.*

H3b. *Female international STEM students who are first-generation students report higher levels of cost values than female international STEM students are continuous-generation students and than male international STEM students.*

No interaction effects between gender and parental academic background were expected for intrinsic, attainment, and utility values, yet we assessed these for explorative reasons. We further pursued some exploratory questions on the associations of cultural background and its interaction with gender, as well as acculturation orientations. In particular, we assessed if academic self-efficacy and value beliefs of international STEM students differed by country groups and if these associations were moderated by students' gender.

Additionally, we examined how study-related language skills, previous stays in Germany, host-culture orientation, and home-culture orientation were related to academic self-efficacy and value beliefs of international STEM students.

In order to address these questions, we used data on bachelor students from a large and diverse panel sample of international students in Germany that provided information from students of different cultural and parental academic backgrounds, studying all over Germany in an extensive range of subject groups and at different higher education institutions. Notably, this dataset provided us with the great advantage of a large control group of international non-STEM students.

3. Method

3.1. Sample

The data come from the first wave of the 3 year German panel study on international students (*International Student Survey* [61]), which was conducted in the interdisciplinary research project *Academic Success and Withdrawal Among International Students in Germany in Bachelor's and Master's Programs (SeSaBa)*. The online survey (presented in German and English language) addressed international students, i.e., students who held a foreign citizenship and obtained their university entrance qualification outside of Germany or at a preparatory college (German: *Studienkolleg*) [1]. Further inclusion criteria were that international students had to study in the first semester of a bachelor's or master's degree program in winter term 2017/2018 and intended to obtain their degree in Germany [61]. The participants registered for the study by answering a short questionnaire (t0) to confirm they content the inclusion criteria. They also received information on data protection standards provided their informed consent here and at the beginning of each questionnaire. The first wave took place at the end of the first semester between January and April 2018.

The analysis sample included data from 1590 international bachelor students ($n_{\text{STEM}} = 882$, $n_{\text{non-STEM}} = 708$). They were studying at 123 universities (with the exception of music and art universities [61]) in all 16 federal states of Germany and in 34 subject groups. The participants came from 120 different countries. The largest group was from Syria (8.6%, $n = 136$), followed by China (6.5%, $n = 104$), Russia (5.6%, $n = 89$), Luxembourg (4.0%, $n = 63$), Bulgaria (3.8%, $n = 60$), and the United States and Ukraine (each 3.7%, $n = 59$). At t1, their mean age was 23.5 years ($SD = 4.25$); 48.3% ($n = 768$) were female and 457 (32.7%) were first-generation students. Further information on the demographic characteristics of the sample is shown in Table 1.

Table 1. Descriptive statistics of sample characteristics and control variables.

Variables	STEM		Non-STEM	
Mean age in years (*SD*)	22.95	(3.58)	24.18	(4.87)
Female (%)	283	(32.1%)	485	(68.5%)
Parental academic background (continuous-generation students) (%)	542	(67.2%)	445	(67.4%)
Country groups (%)				
Western Europe (Reference category)	95	(10.8%)	148	(20.9%)
Central and South Eastern Europe	117	(13.3%)	116	(16.4%)
Eastern Europe and Central Asia	74	(8.4%)	133	(18.8%)
North America	21	(2.4%)	43	(6.1%)
Latin America	67	(7.6%)	53	(7.5%)
North Africa and Middle East	269	(30.5%)	62	(8.8%)
Sub-Saharan Africa	45	(5.1%)	33	(4.7%)
Asia and Pacific	194	(22.0%)	120	(16.9%)
Previous residence in Germany (yes) (%)	253	(29.1%)	300	(43.0%)

Note. N = 1590, n_{STEM} = 882, $n_{\text{non-STEM}}$ = 708 (in the descriptive statistics, listwise deletion was used).

3.2. Instruments and Scales

Academic self-efficacy. Academic self-efficacy was measured using a three-item short scale on general self-efficacy expectations (ASKU) [62] that participants were instructed to answer with regard to their studies. An example item is "In difficult situations, I can rely on my abilities.". The participants indicated their level of agreement on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). The scale revealed a high internal consistency [63] with α = .78.

Study-related value beliefs. Intrinsic value was measured with two items from a scale by Westermann et al. [64] (German translation by Heise and Thies [65]). A sample item was "I find my studies really interesting." (ρ = .73). The study-related attainment value, utility value, and costs were measured with items based on Gaspard et al. [66]. There were two items each on attainment value (e.g., "My studies are very important to me personally."; the Spearman–Brown reliability coefficient was ρ = .69), utility value (e.g., "My studies will positively influence my future."; ρ = .72), and costs (e.g., "I have to make many personal sacrifices for my studies."; ρ = .57). Students indicated their level of agreement using a five-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*.

STEM versus non-STEM-subjects. The categorization of the study subject groups into STEM and non-STEM subjects was based on the specification of the Federal Statistical Office of Germany for the winter semester 2017/2018 [32] and the statistics of the Federal Employment Agency [33]. In this categorization, for example, human medicine and health science, agricultural science, forestry, nutritional science and veterinary medicine, business engineering specializing in economics, and social science were considered as non-STEM subjects, whereas, e.g., pharmacy, business engineering specializing in engineering sciences, architecture, interior design, and spatial planning were considered as STEM subjects, in addition to the more familiar STEM subjects [32,33].

Gender. Participants specified their gender using the categories 0 = *male*, 1 = *female*, and 2 = *diverse*. Because of the small number of students who chose the category "diverse" (< 0.1%) it was not possible to separately analyze this gender category. As a consequence, participants who indicated their gender as "diverse" (n = 6) were omitted from the sample.

Parental academic background. Students were categorized as first-generation students if they reported that *none* of their parents held a bachelor's, master's, or doctorate degree; otherwise, they were categorized as continuous-generation students (0 = *first-generation students*, 1 = *continuous-generation students*).

Cultural background. We operationalized the participants' cultural background by the country of students' university entrance qualification. These countries were categorized into eight regions according to the German Academic Exchange Service's (DAAD) regional coding system [4,67], i.e., Western Europe (which served as the reference category), Central and South Eastern Europe, Eastern Europe and Central Asia, North America, Latin America, North Africa and Middle East, Sub-Saharan Africa, and Asia-Pacific.

Study-related language skills. The participants indicated their study-related language skills by answering the self-formulated item "My language skills are sufficient to cope with my academic studies." on a five-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*).

Previous residence in Germany. Students responded to the question "Before your current stay in Germany, had you ever lived in Germany for at least 1 month?" with 0 = *no* or 1 = *yes* to provide information about a potential previous residence in Germany.

Acculturation orientations. Host-culture orientation (α = .81) and home-culture orientation (α = .81) were measured with four items each that were adapted from Demes and Geeraert [68]. The participants rated the statements on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Sample items were "It is important to me to have friends from my country of origin." for the home-culture orientation and "It is important to me to get to know and maintain German customs and traditions." for the host-culture orientation.

Age. Participants age at the first measurement was calculated from their year of birth.

3.3. Analytical Strategy

All analyses were calculated with IBM SPSS Statistics version 28 [69] and Mplus Version 7 [70]. Missing data were handled with the full information maximum likelihood (FIML) method as implemented in Mplus although missing rates of all constructs were very low (< 1%). Preliminary analyses confirmed that all conditions for linear regression analyses were met. To assess potential dependencies in the data due to the clustering of participants in country groups, we first inspected intraclass correlations (ICC's). Yet, all ICCs were low (< .01) which suggests that the independence of observations can be assumed. Thus, multi-group regression analyses were performed to evaluate significant differences between STEM and non-STEM subjects. All continuous predictors were grand-mean-centered as part of the analyses. In a first step, the predictor variables were inserted, and a Wald-test was run to test the appropriateness of equality constraints between the subject groups. According to non-significant Wald-tests (at $p > .05$), the respective model paths were restricted to be equal. In the next step, the interaction terms (interactions between gender and parental academic background and between gender and the country groups) were inserted, and it was tested if these paths could be equated across subject groups. If the interaction terms did not yield statistical significance in both subject groups, they were removed from the model. In the last step, the intercepts were tested and set equal if the Wald-test was not significant. In all analyses, effects of age were controlled.

3.4. Power Analyses

A post-hoc power analyses using G*Power Version 3.1.9.7 [71,72] revealed that, with a significance level of α = .05 and N = 1590, even small (interaction) effects of (f^2 = .02) could be detected with a probability of $1 - \beta$ = .96. Accordingly, a target power of at least .80 according to Cohen [73] was achieved.

4. Results

4.1. Descriptive Analyses

The international students showed high means of academic self-efficacy and all value components (see Table 2). Despite slightly different means, the overall pattern was the same in both subject groups. The highest values were observed for attainment value, intrinsic value, and utility value, which all clearly exceeded the scale mean of three. The means of academic self-efficacy and cost value were lower, but also exceeded the scale means. T-Tests revealed significant difference between STEM and non-STEM students in academic self-efficacy whilst all other differences were negligible (see Table 2). The bivariate correlations between the study variables in the STEM and non-STEM sample are presented in Table 3. Overall, the correlation pattern was similar in both subject groups and yielded small to moderate sizes associations.

Table 2. Descriptive statistics of main scale study variables.

Variables	STEM		Non-STEM		*t*-Tests		
	M	*SD*	*M*	*SD*	*t(df)*	*p*	*d*
Study-related language skills	3.98	0.92	4.03	0.96	0.96(1561)	.335	.05
Home-culture orientation	2.92	0.90	3.00	0.90	1.64(1518)	.102	.08
Host-culture orientation	3.51	0.84	3.54	0.78	0.68(1519)	.498	.04
Dependent variables							
Academic self-efficacy	3.71	0.71	3.81	0.69	2.72(1553)	.007	.14
Value beliefs							
Attainment value	4.49	0.68	4.48	0.60	−0.33(1559)	.745	−.02
Intrinsic value	3.99	0.82	4.06	0.80	1.78(1559)	.075	.09
Utility value	4.53	0.67	4.51	0.61	−0.53(1559)	.598	−.03
Cost value	3.45	0.89	3.39	0.91	−1.30(1559)	.193	−.07

Note. n_{STEM} = 882, $n_{\text{non-STEM}}$ = 708 (in the descriptive statistics, listwise deletion was used).

4.2. Multi-Group Analyses

All models explained substantial variance in academic self-efficacy (12% STEM; 15% non-STEM), attainment value (8% STEM; 11% non-STEM), intrinsic value (12% both subject groups), utility value (8% STEM; 12% non-STEM), and cost (7% STEM; 9% non-STEM). Overall, the multi-group analyses yielded very few differences between STEM- and non-STEM students (Table 4). That is, most of the reported results represent generalized patterns that apply to international students in STEM subjects and beyond. As a consequence, the below description of results is focused on the results for the STEM sample; findings from the comparison group of non-STEM students are only considered in cases of significant differences between the subject groups. In line with the results from the t-tests, differences in the adjusted intercepts were only identified for academic self-efficacy (adjusted M = 3.55, SE = 0.11, $p < .001$ for STEM students; M = 3.63, SE = 0.12, $p < .001$ for non-STEM; Wald χ^2 = 5.83, p = .016). All other Wald-tests on the adjusted intercepts—for attainment (M = 4.43, SE = 0.07), intrinsic (M = 4.13, SE = 0.06), utility (M = 4.48, SE = 0.07), and cost value (M = 3.12, SE = 0.07)—did not yield significance (all p-values > .136).

Contrary to our assumptions, we did not find associations between gender and any expectancy or value components (Table 4). That is, the hypotheses H1a and H1b had to be rejected.

Parental academic background showed a significant association with academic self-efficacy (b = .08, SE = .04, β = .12, p = .027) which implies that, in line with hypothesis 2a, continuous-generation STEM students showed higher levels of academic self-efficacy than first-generation STEM students.

Parental academic background showed no statistical significance in relation with cost value. Hence, hypothesis 2b was not supported.

Contrary to our assumptions, no significant interaction of parental academic background and gender in relation with academic self-efficacy (H3a) and cost values (H3b) could be identified which implies the rejection of both hypotheses.

Table 3. Correlations between the study variables.

Correlations	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.
1. Academic self-efficacy	-	.15 **	.30 **	.20 **	−.10 **	.39 **	.00	.01	.07	.03	.05	.15 **	−.02	−.06	−.04	−.17 **	.01	.02	.04
Value beliefs																			
2. Attainment value	.21 **	-	.36 **	.49 **	.10 **	.11 **	.07	−.01	−.09 *	.04	.00	−.17 **	.08 *	.05	.11 **	−.02	.04	−.02	.19 **
3. Intrinsic value	.34 **	.40 **	-	.30 **	−.22 **	.31 **	−.01	.06	−.05	−.01	.01	−.01	.03	.04	.04	−.17 **	.01	−.07	.10 *
4. Utility value	.21 **	.62 **	.38 **	-	.08 *	.16 **	.07	.01	−.00	.09 *	.10 **	−.13 **	.12 **	.07	.02	−.15 **	.10 *	−.04	.18 **
5. Cost value	−.07 *	.21 **	−.11 **	.17 **	-	−.22 **	−.01	.14 **	−.02	.03	−.00	.12 **	.03	.08 *	.02	.06	.03	.07	.13 **
6. Study-related language skills	.27 **	.11 **	.27 **	.13 **	−.15 **	-	−.05	−.11 **	.01	.08 *	.02	−.00	−.03	−.11 **	.06	−.41 **	.04	.01	−.02
7. Previous residence Germany	.00	−.06	−.05	−.03	.02	−.04	-	.12 **	−.01	.02	.21 **	−.07	.16 **	−.14 **	−.03	−.06	.12 **	−.13 **	.02
8. Age	−.04	.01	.01	−.05	.06	−.07	.08 *	-	−.12 **	−.15 **	.08 *	.15 **	.06	.14 **	.09 *	.02	−.03	−.07	.03
9. Parental academic background	.07 *	.02	.07	−.00	.03	.02	.08 *	−.24 **	-	−.05	.12 **	.14 **	.11 **	−.05	.01	−.06	−.08	.04	.02
Country groups																			
10. Central and South Eastern Europe	.05	−.06	.04	.04	.00	.09 **	.08 *	−.22 **	.01	-	−.21 **	−.11 **	−.13 **	−.14 **	−.01 **	−.20 **	.05	−.06	−.01
11. Eastern Europe and Central Asia	.04	.00	.04	.09 *	−.04	.07 *	.22 **	.03	.11 **	−.12 **	-	−.12 **	−.14 **	−.15 **	−.11 **	−.22 **	.16 **	−.08 *	.02
12. North America	.04	−.09 *	−.04	−.11 *	.04	.02	.02	.13 **	.06	−.06	−.05	-	−.07	−.08 *	−.06	−.12 **	−.02	.13 **	.08 *
13. Latin America	.07	.03	.05	.07 *	.10 **	.04	.13 **	.01	.11 **	−.11 **	−.09 **	−.05	-	−.09 *	−.06	−.13 **	−.03	−.02	.12 *
14. North Africa and Middle East	−.04	.12 **	−.10 **	.03	.07	−.10 **	−.15 **	.21 **	−.07 *	−.26 **	−.20 **	−.10 **	−.19 **	-	−.07	−.14 **	−.25 **	−.07	−.03
15. Sub-Sahara Africa	−.03	.13 **	.08 *	.12 **	.00	.04	−.12 **	.07	−.08 *	−.09 **	−.07 *	−.04	−.07 *	−.15 **	-	−.10 **	−.02	−.03	−.00
16. Asia and Pacific	−.07	−.09 *	−.09 **	−.18 **	−.05	−.26 **	−.02	−.01	−.01	−.21 **	−.16 **	−.08 *	−.15 **	−.35 **	−.12 **	-	−.00	.06	.03
17. Gender	−.04	−.03	−.00	.07 *	.00	.02	.07 *	−.03	.07 *	.08 *	.17 **	.00	.05	−.21 **	−.02	.04	-	−.01	−.01
18. Home-culture orientation	−.05	−.04	−.06	−.05	−.00	−.03	−.11 **	−.20 **	.07	−.04	−.11 **	.04	−.03	−.06	−.00	.12 **	−.02	-	.16 **
19. Host-culture orientation	.03	.12 **	.14 **	.09 *	.01	−.02	.01	.05	−.03	−.10 **	.03	.06	.05	.03	−.04	.04	.03	.08 *	-

Note. Correlations for the non-STEM group are above the diagonal. Correlations for the STEM group are below the diagonal. For categorial variables, Spearman's rho is presented. $n_{\text{STEM}} = 882$, $n_{\text{non-STEM}} = 708$. Listwise deletion was used. Country groups: Reference category, Western Europe. ** Correlation is significant at the level of $p < .01$. * Correlation is significant at the level of $p < .05$.

Table 4. Results of the multi-group models for STEM students.

Predictors	Academic Self-Efficacy			Attainment Value			Intrinsic Value			Utility Values			Cost Value		
Final Model	***b***	***SE***	***p***	***b***	***SE***	***p***	***b***	***SE***	***p***	***b***	***SE***	***p***	***b***	***SE***	***p***
Age	.00	.00	.468	−.00	.00	.830	.01	.01	.013	−.00	.00	.566	.01	.01	.017
Previous residence in Germany	−.02	.04	.634	.02	.04	.501	−.05	.05	.271	−.03	.04	.480	−.00	.05	.971
Study-related language skills	.27	.02	< .001	.09	.02	< .001	.23	.02	< .001	.08	.02	< .001	−.17	.03	< .001
Country groups[a]															
Central and South Eastern Europe	.14	.06	.027	−.13	.09	.160	−.11	.07	.147	.08	.09	.348	.35	.08	< .001
Eastern Europe and Central Asia	.10	.07	.122	.06	.11	.557	−.10	.08	.222	.17	.11	.110	.24	.09	.007
North America	.30	.10	.002	−.63	.13	< .001	−.25	.11	.027	−.61	.13	< .001	.56	.13	< .001
Latin America	.12	.08	.137	.05	.10	.606	−.05	.09	.561	.15	.10	.143	.50	.10	< .001
North Africa and Middle East	.06	.06	.335	.17	.08	.024	−.21	.07	.004	.11	.07	.135	.39	.08	< .001
Sub-Saharan Africa	−.03	.09	.759	.41	.12	< .001	.05	.11	.666	.21	.12	.063	.33	.12	.006
Asia and Pacific	.12	.06	.073	.11	.08	.194	−.15	.07	.047	−.09	.08	.292	.17	.09	.045
Gender	−.05	.04	.160	−.04	.09	.625	−.01	.04	.829	−.06	.09	.529	.05	.05	.303
Parental academic background	.08	.04	.027	−.04	.05	.459	.06	.04	.180	−.04	.05	.434	.00	.05	.998
Home-culture orientation	−.02	.02	.439	−.02	.02	.393	−.07	.02	.002	−.02	.02	.333	.05	.03	.055
Host-culture orientation	.03	.02	.151	.12	.02	< .001	.14	.03	< .001	.10	.02	< .001	.04	.03	.169
Gender × parental academic background				−.01	.07	.873				.03	.07	.618			
Central and South Eastern Europe × gender				.34	.12	.003				.19	.12	.096			
Eastern Europe and Central Asia × gender				.09	.13	.475				.11	.13	.400			
North America × gender				.47	.17	.006				.60	.17	< .001			
Latin America × gender				.26	.14	.066				.25	.14	.072			
North Africa and Middle East × gender				−.00	.12	.987				.11	.12	.347			
Sub-Saharan Africa × gender				−.14	.17	.397				.05	.17	.746			
Asia-Pacific × gender				.00	.11	.999				.12	.11	.263			

Note. n_{STEM} = 803. [a] Country groups used Western Europe as the reference category. The final model of each dependent variable is presented here.

With regard to cultural background, some country groups yielded statistically significant associations with academic self-efficacy: students from Central and South Eastern Europe (*b* = .14, *SE* = .06, β = .37, *p* = .027) and North America (*b* = .30, *SE* = .10, β = .41, *p* = .002) showed higher levels of study-related self-efficacy as compared to the reference group Western Europe. Beyond this, students' cultural background played a role with regard to all investigated value outcomes, with most associations being substantiated between cultural background and cost value, as presented in Table 4. In particular, students from all other country groups indicated higher cost value than their fellows from Western Europe (β's ranged from .19 to .63). Furthermore, students from North America (*b* = −.25, *SE* = .11, β = −.31, *p* = .027), from North Africa and Middle East (*b* = −.21, *SE* = .07, β = −.26, *p* = .004), and from Asia-Pacific (*b* = −.15, *SE* = .07, β = −.18, *p* = .047) showed lower levels of intrinsic value in comparison to the students from Western Europe.

No significant interactions of country groups and gender in relation with academic self-efficacy could be observed. Yet, some instances of interactions between country groups and gender were detected for attainment and utility value (Figures 1 and 2, Table 4), pointing toward differences between male and female students that only occurred amongst students from these country groups. In particular, female students from Central and South Eastern Europe (*b* = .34, *SE* = .12, β = .58, *p* = .003) and North America (*b* = .47, *SE* = .17, β = .80, *p* = .006) reported higher attainment values than their male fellows. The same applied to female students from North America with regard to their perceptions of utility values (*b* = .60, *SE* = .17, β = .92, *p* < .001). Regarding cost and intrinsic value, no intersections of country groups and gender showed statistical significance (all *p*-values > .060).

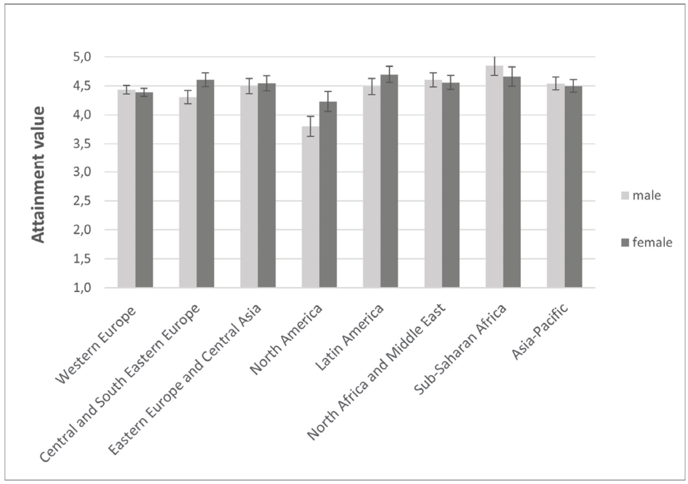


Figure 1. The interaction of country groups and gender in the prediction of attainment value. $n_{\mathrm{STEM}} = 803$. Error bars represent the standard errors. Reference category: Western Europe.

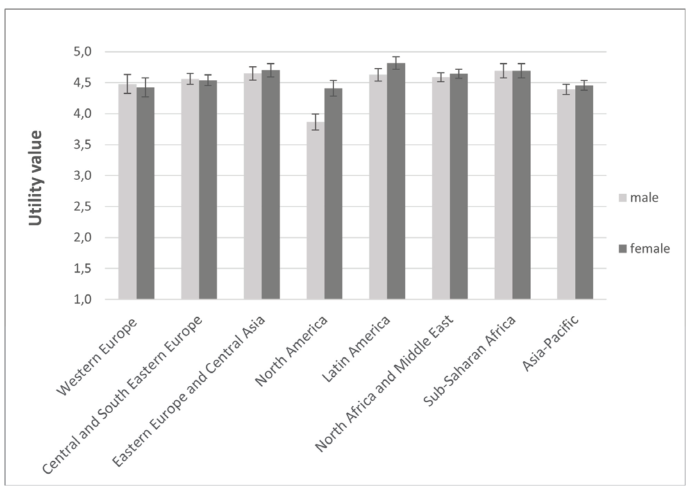


Figure 2. The interaction of country groups and gender in the prediction of utility value. $n_{\mathrm{STEM}} = 803$. Error bars represent the standard errors. Reference category: Western Europe.

Study-related language skills showed a positive significant relation with academic self-efficacy ($b = .27$, $SE = .02$, $\beta = .34$, $p < .001$), attainment value ($b = .09$, $SE = .02$, $\beta = .12$, $p < .001$), intrinsic value ($b = .23$, $SE = .02$, $\beta = .26$, $p < .001$), and utility value ($b = .08$, $SE = .02$, $\beta = .11$, $p < .001$), whereas a negative association with cost value ($b = -.17$, $SE = .03$, $\beta = -.17$, $p < .001$) could be identified. By contrast, previous residence in Germany showed no significant association with the dependent variables (all p-values > .270).

Moreover, acculturation orientations showed no significant relation with academic self-efficacy. However, acculturation orientations were positively related with attainment value ($b = .12$, $SE = .02$, $\beta = .15$, $p < .001$), intrinsic value ($b = .14$, $SE = .03$, $\beta = .15$, $p < .001$), and utility value ($b = .10$, $SE = .02$, $\beta = .13$, $p < .001$), i.e., a higher host-culture orientation was associated with higher levels in the three value components, whereas a higher home-culture orientation was only related to lower levels of intrinsic value ($b = -.07$, $SE = .02$, $\beta = -0.08$, $p = .002$).

With regard to age, a small positive relation with intrinsic value (b = .01, SE = .01, β = .06, p = .013), as well as cost (b = .01, SE = .01, β = .05, p = .017), showed statistical significance.

5. Discussion

Despite increasing numbers of international students in OECD countries such as Germany [3,67,74], little is known about their (psychological) prerequisites for study success. Given the relevance of expectations and value beliefs for a successful completion of the academic track [21–23], we explored their manifestation amongst international STEM students in Germany and investigated if former findings on differences in these variables by gender, parental academic background, and cultural background that were identified amongst school students or in US student samples were transferable to our sample. Beyond this, we also investigated associations with study-related language skills, previous residence in Germany, and the international students' acculturation orientations. To support our findings, we compared the associations between STEM and a non-STEM comparison group as a kind of robustness check and to explore potential differences between subject groups.

In general, the international students showed high levels of expectations and value beliefs. Yet, only the means of academic self-efficacy differed significantly between the subject groups, i.e., international STEM students showed significantly lower levels of academic self-efficacy than international non-STEM students. All other results (regression paths in the multi-group model, unadjusted and adjusted means) did not significantly differ between the subject groups. Against our expectations, there were no direct associations between gender and the expectancy–value components. Yet, the parental academic background was related to academic self-efficacy as continuous-generation students showed higher academic self-efficacy than first-generation students. By contrast, the relation between parental academic background and cost value was not significant, and there were no interactions between gender and parental academic background.

Furthermore, significant associations between cultural background and all expectancy–value components were identified, most of them referring to cost. In some cases, the associations differed by gender. Study-related language skills were positively related to academic self-efficacy and attainment, intrinsic, and utility value but negatively to costs. By contrast, no significant associations between previous residence in Germany and the outcome variables were substantiated. With regard to the acculturation orientations, host-culture orientation was positively related with attainment value, intrinsic value, and utility value, whereas home-culture orientation was negatively associated with intrinsic value. Below, we discuss the outlined results in more detail.

5.1. High Levels of Expectations and Value Beliefs Amongst International STEM Students

With regard to the manifestation of expectancies and values amongst international students, the analyses revealed high levels of all expectancy and value components. That is, at the beginning of their studies, international students in Germany seemed to be quite optimistic about their potential to meet the requirements of their degree courses and value their studies for several reasons, such as intrinsic joy, personal importance, and utility for their future. Yet, despite the rather high levels of academic self-efficacy being rather high, comparisons between the subject groups revealed that international STEM students scored lower than their fellows in non-STEM subjects. This applied to both the unadjusted means (that were compared via t-tests) and the intercepts in the final multi-group regression model that included all main predictors and, thus, controlled for potential differences between the subject groups in these variables. These results were in line with the findings by Lee et al. [75] who reported similar differences between the subject groups. This might be reasoned in the high reputation of STEM subjects as being particularly demanding and prestigious, which might decrease students' expectations to be able to master the upcoming academic challenges [76,77]. Furthermore, the international students also reported high levels of costs, i.e., they feel that their studies in Germany require great personal sacrifices and cause many worries. Interestingly, in opposition to the common finding that higher intrinsic, attainment, and utility values are accompanied by reduced costs [78], attainment and utility values were positively correlated with costs in the present sample. That is, international students who indicated that their studies are very important and useful to them also experienced substantial strain. This pattern might be specific to international students who have to master additional challenges beyond the academic context such as the cultural transition. It might also reflect an incident of introjected regulation, i.e., motivation from partially internalized values such as seeking approval and protecting the ego [79]. In the present case, the students' high levels of attainment and utility values might be influenced by the values and expectations of significant others. As a consequence, the students do not benefit from these elevating motivations but still suffer from increased costs. In the future, person-centered analyses that investigate if there are distinct motivational profiles amongst international (STEM) students and how these are linked to values and expectations of significant others might help to better understand the observed pattern. Furthermore, future research may also investigate the development of the expectancy and value components over the study course and how they interact in the prediction of study outcomes. Maybe the high costs decrease the positive effects of the other value components. This may provide further insights into the question why so many international students are dropping out of their degree courses [6] despite their rather high levels of self-efficacy, as well as intrinsic, attainment, and utility values at the beginning of their studies.

5.2. Differences in Expectations and Value Beliefs by Gender and Parental Academic Background

Against our expectations, no direct associations between gender and any expectancy or value components could be found. This points to an important difference between previous findings from school and general student samples where females revealed less favorable STEM expectations and value beliefs than males [9,29]. This pattern might be reasoned in the sample selectivity of international degree-seeking STEM students in Germany. International students have already accomplished several challenges before starting their studies abroad (e.g., getting visa and study permissions, dealing with expectations and potential stereotypes of family, teachers and peers [80], and obtaining all relevant (language) qualifications). They decided for a STEM degree abroad despite potential (gender) stereotypes and possible cultural barriers [22,25,28,80,81]. Hence, this sample can be assumed to be highly selective (which is corroborated by the high levels of expectancies and value beliefs). In this highly selective sample, gender might not play the same role as in general student samples. Yet, further longitudinal studies are needed to investigate if gender differences may unfold during later stages of the study programs.

Consistent with our hypothesis, continuous-generation students indicated higher levels of academic self-efficacy; that is, they were more optimistic to fulfil all necessary study requirements than their first-generation fellows. This is in line with the results of Wille et al. [25], who identified a significant positive correlation between socioeconomic status and math expectancy in a school sample. One explanation could be that parents with an academic background may be able to provide more (financial) support and helpful role models to their children. On the one hand, this may let them feel that their parents trust in their persistence and that they fit into their study environment. On the other hand, the parents' potential to provide financial support may reduce their fear of having to drop out of their studies due to a lack of financial means. Contrary to our hypotheses, no significant relations between parental academic background costs could be identified. Hence, our results did not support the findings of Goldman et al. [41], Meyer et al. [40], and Wille et al. [25] who showed that first-generation college students indicated higher levels of costs as compared to continuous-generation students, and that this effect even increased across the semester [41].

Lastly, and also contrary to our hypotheses, no interactions of parental academic background and gender in the prediction of academic self-efficacy and cost values yielded significance. That is, earlier findings which showed that female first-generation students are particularly disadvantaged in terms of lower self-efficacy [58] and increased cost value [9,82] could not be substantiated for international STEM students. Although the present findings may be interpreted as good news as (female) first-generation international STEM students do not seem to be more challenged by their studies than their fellows from academic households, further research may investigate the persistence of this pattern over the study course.

5.3. Differences in Expectations and Value Beliefs by Cultural Characteristics

The results of this study also indicate that students' cultural background played a role with regard to all investigated outcomes, with most effects being identified for costs. Beyond the rather high mean levels, substantial differences between the country groups were substantiated as students from all country groups indicated higher cost values than their fellows from the reference category Western Europe. One explanation could be that coping with cultural distance is a relevant factor. Students from more culturally distant countries may experience greater challenges [83,84] which makes them experience higher study costs. Yet, the relations between costs and further factors of adaptation abroad poses an important question that could be addressed in future explorational research. In particular, with regard to the (psychosocial) counselling of international students, it might be helpful to further explore if international students who perceive higher costs are more likely to suffer from burnout or reduced psychological wellbeing.

Some further associations between single country groups and expectancy and value components were observed, as well as some interactions between country groups and gender. The latter corroborate the findings of previous studies which suggested heterogeneous associations between gender and expectancy–value components in different ethnic groups in the US [59,85–87]. Overall, our results on cultural differences and their interplay with gender in the prediction of expectancy–value components enhance the knowledge on these contingencies amongst international students in Germany. Yet, in order to validate these findings and to explore the mechanisms that account for the observed effects, further research that more closely considers (cultural) differences between and within these country groups is deemed essential.

A previous residence in Germany showed no significant association with expectations and value beliefs. This might be reasoned in the rather unspecific measure that did not distinguish between different types of previous residence (e.g., holidays, school exchange, and family stay abroad) and the duration of these stays. Further research may thus benefit from considering such information. Contrary, study-related language skills were positively related with academic self-efficacy, attainment values, intrinsic value, and utility value, but negatively with cost value. This corroborates the importance of language skills with regard to academic success and its (psychological) prerequisites [48,88]. Further research might differentiate between skills in study program language (as assessed in our study) and host country language (which are not necessarily the same as, for example, students may be enrolled degree courses that are taught in English at German universities) and consider more extensive and objective measures of language competencies such as test scores.

Lastly, the analyses showed that acculturation orientations were not significantly related with academic self-efficacy, but were associated with attainment, intrinsic, and utility values; that is, a higher host-culture orientation was associated with higher levels in all three value components whereas a higher home-culture orientation was related to lower levels of the intrinsic component. These results partly support evidence from Berry [51] who pointed to the positive associations of host-culture orientations with adjustment and productive adaption in the school context and beyond. They are also in line with findings from a meta-analysis by Bierwiaczonek and Kunst [53], which suggested that the host-culture orientation was particularly relevant in the early stages of migration. Yet, whilst a higher host-culture orientation was approved as a valuable resource for international students, the results on the home-culture orientation showed a contrary pattern as a higher home-culture orientation was related to lower intrinsic value. That is, students who are very engaged with their home-culture seem to be less prone to enjoy their studies abroad.

5.4. Implications for Practice

Of course, it is important to mention that all reported results reflect small effect sizes. Yet, even small effect sizes matter, because these small effects may influence the whole academic career and, thus, gain practical relevance. Against this background, some practical implications that may be inferred from the present results are considered.

In line with former research, international first-generation students showed lower levels of self-efficacy than continuous-generation students in both subject groups. That implies that interventions promoting academic self-efficacy should particularly consider the encouragement of international first-generation students, e.g., by directly addressing them with teaser-texts such as "You are the first in the family who studies (abroad)? Meet others and find information regarding. . .".

One further important finding with regard to practical implications was the overall high levels of perceived costs amongst the international students that were increased for students from countries outside of Western Europe. These findings suggest that it will be most helpful to integrate interventions that address the reduction in psychological costs into the counselling programs for international students. Even though some promising examples of interventions that directly address psychological costs were described in previous studies [89,90], future research is needed to further elaborate and evaluate these interventions and to adapt them to the specific needs of international (STEM) students. In the meantime, interventions that address related topics and conditions of psychological costs may be usefully employed. On the one hand, interventions may address general study management skills such as time management (e.g., Middendorff [91]). On the other hand, our findings suggest that it might be beneficial to complement cost interventions for international students with elements of intercultural trainings as information on cultural differences and hands-on practices on how to handle these effectively as this may also contribute to the reduction of perceived costs. This suggestion is corroborated by findings from Poort et al. [81] who showed that the engagement in intercultural group work reduced perceived costs. With regard to higher education policy, establishing part-time study programs and blended learning models may also help to decrease perceived study burdens and, thus, psychological costs (e.g., Middendorff [91]). To conclude, addressing psychological costs is deemed particularly important as recent studies suggested that cost value moderate the effect of the other (S)EVT components on academic outcomes [78]. This also suggests that interventions which increase academic self-efficacy or attainment, intrinsic, and utility values may pay off in terms of reduced psychological costs. For example, previous studies suggested that the use of cooperative learning strategies, setting specific, short-term goals and encouraging goal reflection would be helpful to support students' academic self-efficacy [15,92], which in turn may also reduce the psychological costs.

Lastly, our findings on acculturation orientations suggest that it will be advantageous for the adaptation process to support international students' host-culture orientation. For example, promoting contact to the campus community may help in this regard [93]. Likewise, a deliberate approach toward the own home-culture orientation may help students to maintain their study motivation and enjoyment.

All implications need to consider the relations of study-related language skills and the expectancy–value components regardless of subject group. The more the students feel they can master study-related communication, the higher their study-related expectancy and the more adaptive their value beliefs will be. Wisniewski et al. [48,88] described language skills (particularly reading skills) to be highly predictive of academic success in the study entry phase of international students. Our findings support Wisniewski's [48,88] results and suggest that one mechanism by which language skills affect the academic success might be via international students' more adaptive study-related expectations and value constellations.

Importantly, as there were almost no differences between the subject groups, all of the suggested interventions may be equally beneficial for all international students in STEM and non-STEM subjects. Furthermore, in contrast to previous studies, almost no direct relations for gender and parental academic background with the value components were detected; that is, in this highly selected sample of international university students, it is probably not necessary to make targeted offers for female or for first-generation students.

5.5. General Limitations

First, it is important to consider that we investigated a specific sample, i.e., international students in Germany. The (international) student body in other countries might differ from this sample, e.g., with regard to their demographic and cultural characteristics. Likewise, the study conditions and environment will be different and, thus, present students with other chances and challenges that might influence their study-related expectations and value beliefs. Hence, caution is warranted with regard to the transferability of the present results to other samples. Likewise, our sample was limited to bachelor students. As master's students may have different and more heterogeneous (study) experiences in their subjects and beyond, their expectancies and value beliefs may be different and require further research.

Moreover, the distribution of country groups showed a small number of cases in some of the investigated country groups. This might be a reason why some effects may not have been identified for these groups or may have been inflated by few extreme cases. Furthermore, differences between the countries within the country groups were not investigated. As a consequence, it is questionable to what extent the results can be applied to each and every single country, particularly due to the small number of cases in some countries.

We examined the associations of gender, parental academic background, and cultural characteristics with the expectancy–value components in a cross-sectional design to explore contingencies between these variables at the beginning of the study time abroad. This was deemed important to get an understanding of which student groups might have increased risks of facing challenges in their studies or dropping out of university due to their maladaptive expectancy–value constellations. Further longitudinal analyses are needed to explore the differential trajectories of the expectancy–value components amongst STEM students over the study course and the longitudinal relations between study-related expectancy and value beliefs and academic success of international STEM students. Additionally, the interplay of acculturation orientations and study-related expectancy and value beliefs needs to be focused on future research, to more precisely understand their longitudinal interdependencies and the role of the duration of the stay abroad [53]. Furthermore, due to the non-experimental cross-sectional design of the present study, no causal interpretation of the results is possible.

Additionally, there were some limitations with regard to the used measurement instruments. First, only self-report measures were used which might imply some bias due to socially desirable responding [94,95]. Second, some of the subscales (particularly in value beliefs) were measured with only two items which may have restricted their content validity. Although this is a common approach in large panel studies, further research may benefit from enclosing more detailed measures of the investigated constructs.

6. Conclusions

Despite these limitations, the present study provided important insights with regard to the manifestation of and differences in study-related expectations and value beliefs as (psychological) prerequisites for study success amongst international STEM students (and non-STEM students) in Germany.

Importantly, without an intersectional perspective, some of the contingencies between demographic characteristics and the expectancy–value components would not have been detected. Hence, the intersectional perspective can be helpful in differentiating effects of belonging to multiple potentially disadvantaged groups [54,96]. We hope that the present results may encourage further research on the heterogeneity amongst internationals students in Germany and its implications with regard to their specific strengths and potential support needs.

Author Contributions: Conceptualization, methodology, data preparation, formal analysis, writing—original draft preparation, J.S.P.; conceptualization, writing—review and editing, J.Z. and K.J.; supervision, J.Z. and K.J.; project administration; J.Z. and J.S.P.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF), Grant number 16PX21018A, awarded to Dr. Julia Zimmermann.

Institutional Review Board Statement: Due to the noncontroversial nature of the content and design of the study, an ethics approval for this research was not required as per the ethical guidelines of the German Psychological Society (DGPs) and data protection standards at the time of data collection.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data used in this study are available upon request from the first author. This study's design and its analysis were not preregistered.

Acknowledgments: We thank Sandy Möckel for her valuable help with the preparation of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Federal Statistical Office—Destatis (Statistisches Bundesamt Destatis). *Bildung und Kultur: Studierende an Hochschulen. Sommersemester 2020*; Fachserie 11—Bildung und Kultur 4.1; Statistisches Bundesamt: Wiesbaden, Germany, 2020. Available online: https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00061732/2110410207314_fuer_Bibliothek.pdf (accessed on 26 July 2023).
2. OECD. *International Student Mobility (Indicator)*; OECD Publishing: Paris, France, 2023. [CrossRef]
3. OECD. *International Migration Outlook 2022*; OECD Publishing: Paris, France, 2022. [CrossRef]
4. DAAD; DZHW. *Wissenschaft Weltoffen 2021*; wbv Media GmbH & Co. KG: Bielefeld, Germany, 2021. [CrossRef]
5. Hoffmeyer-Zlotnik, P.; Grote, D. *Anwerbung und Bindung von internationalen Studierenden in Deutschland: Studie der Deutschen Nationalen Kontaktstelle für das Europäische Migrationsnetzwerk (EMN)*; Working Paper; Bundesamt für Migration und Flüchtlinge (BAMF) Forschungszentrum Migration, Integration und Asyl (FZ): Nürnberg, Germany, 2019. Available online: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-67593-8 (accessed on 26 July 2023).
6. Heublein, U.; Hutzsch, C.; Schmelzer, R. *Die Entwicklung der Studienabbruchquoten in Deutschland*; DSHW Brief; DZHW: Hannover, Germany, 2022. [CrossRef]
7. Eccles, J.S.; Adler, T.F.; Fattermann, R.; Goff, S.B.; Kaczala, C.M.; Meece, J.L.; Midgley, C. Expectancies, values and academics behaviors. In *Achievement and Achievement Motivation*; Spence, J.T., Ed.; Freeman: San Francisco, CA, USA, 1983; pp. 75–146.
8. Eccles, J.S.; Wigfield, A. From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemp. Educ. Psychol.* **2020**, *61*, 101859. [CrossRef]
9. Parker, P.D.; van Zanden, B.; Marsh, H.W.; Owen, K.; Duineveld, J.J.; Noetel, M. The intersection of gender, social class, and cultural context: A meta-analysis. *Educ. Psychol. Rev.* **2020**, *32*, 197–228. [CrossRef]
10. Cole, E.R. Intersectionality and research in psychology. *Am. Psychol.* **2009**, *64*, 170–180. [CrossRef]
11. Knapp, G.-A. Verhältnisbestimmungen: Geschlecht, Klasse, Ethnizität in gesellschaftstheoretischer Perspektive. In *Im Widerstreit: Feministische Theorie in Bewegung*; Knapp, G.-A., Ed.; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2012; pp. 429–460.
12. Eccles, J.S. Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychol. Women Quart.* **1994**, *18*, 585–609. [CrossRef]
13. Dickhäuser, O.; Seidler, A.; Kölzer, M. Kein Mensch kann alles? Effekte dimensionaler Vergleiche auf das Fähigkeitsselbstkonzept. *Z. Padagog. Psychol.* **2005**, *19*, 97–106. [CrossRef]
14. Marsh, H.W. Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years. *J. Educ. Psychol.* **2007**, *99*, 775–790. [CrossRef]
15. Bandura, A. *Self-Efficacy: The Exercise of Control*; W.H. Freeman & Co., Ltd.: New York, NY, USA, 1997.
16. Pajares, F. Self-efficacy beliefs in academic settings. *Rev. Educ. Res.* **1996**, *66*, 543–578. [CrossRef]
17. Eccles, J.S.; Wigfield, A. Motivational beliefs, values, and goals. *Annu. Rev. Psychol.* **2002**, *53*, 109–132. [CrossRef]
18. Bandura, A. *Social Foundations of Thought and Action: A Social Cognitive Theory*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1986.
19. Bandura, A.; Pastorelli, C.; Barbaranelli, C.; Caprara, G.V. Self-efficacy pathways to childhood depression. *J. Pers. Soc. Psychol.* **1999**, *76*, 258–269. [CrossRef]
20. Lent, R.W.; Brown, S.D.; Hackett, G. Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* **1994**, *45*, 79–122. [CrossRef]

21. Eccles, J.S.; Vida, M.N.; Barber, B. The relation of early adolescents' college plans and both academic ability and task-value beliefs to subsequent college enrollment. *J. Early Adolesc.* **2004**, *24*, 63–77. [CrossRef]
22. Wigfield, A.; Eccles, J.S. Expectancy-value theory of achievement motivation. *Contemp. Educ. Psychol.* **2000**, *25*, 68–81. [CrossRef] [PubMed]
23. Wigfield, A.; Eccles, J.S. 35 years of research on students' subjective task values and motivation: A look back and a look forward. *Adv. Motiv. Sci.* **2020**, *7*, 161–198. [CrossRef]
24. Trautwein, U.; Marsh, H.W.; Nagengast, B.; Lüdtke, O.; Nagy, G.; Jonkmann, K. Probing for the multiplicative term in modern expectancy–value theory: A latent interaction modeling study. *Jpn. J. Educ. Psychol.* **2012**, *104*, 763–777. [CrossRef]
25. Wille, E.; Stoll, G.; Gfrörer, T.; Cambria, J.; Nagengast, B.; Trautwein, U. It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemp. Educ. Psychol.* **2020**, *61*, 101858. [CrossRef]
26. Eccles, J.S.; Wang, M.-T. What motivates females and males to pursue careers in mathematics and science? *Int. J. Behav. Dev.* **2016**, *40*, 100–106. [CrossRef]
27. Oppermann, E.; Keller, L. *Geschlechtsunterschiede in der Frühen MINT-Bildung—Forschungsüberblick*; Stiftung-Kinder-Forschen: Berlin, Germany, 2018. Available online: https://www.haus-der-kleinen-forscher.de/ (accessed on 26 July 2023).
28. Jiang, S.; Simpkins, S.D.; Eccles, J.S. Individuals' math and science motivation and their subsequent STEM choices and achievement in high school and college: A longitudinal study of gender and college generation status differences. *Dev. Psychol.* **2020**, *56*, 2137–2151. [CrossRef]
29. Huang, C. Gender differences in academic self-efficacy: A meta-analysis. *Eur. J. Psychol. Educ.* **2013**, *28*, 1–35. [CrossRef]
30. Pajares, F.; Miller, M.D.; Johnson, M.J. Gender differences in writing self-beliefs of elementary school students. *J. Educ. Psychol.* **1999**, *91*, 50–61. [CrossRef]
31. Lee, Y.; Freer, E.; Robinson, K.A.; Perez, T.; Lira, A.K.; Briedis, D.; Walton, S.P.; Linnenbrink-Garcia, L. The multiplicative function of expectancy and value in predicting engineering students' choice, persistence, and performance. *J. Eng. Educ.* **2022**, *111*, 531–553. [CrossRef]
32. Federal Statistical Office—Destatis (Statistisches Bundesamt Destatis). *Bildung und Kultur: Studierende an Hochschulen. Wintersemester 2017/18*; Fachserie 11—Bildung und Kultur No. 4.1; Statistisches Bundesamt: Wiesbaden, Germany, 2018. Available online: https://www.statistischebibliothek.de/mir/receive/DEHeft_mods_00092410 (accessed on 26 July 2023).
33. Federal Employment Agency (Bundesagentur für Arbeit). *Berichte: Blickpunkt Arbeitsmarkt—MINT—Berufe*; Bundesagentur für Arbeit Statistik/Arbeitsmarktberichterstattung: Nürnberg, Germany, 2019.
34. Padgett, R.D.; Johnson, M.P.; Pascarella, E.T. First-generation undergraduate students and the impacts of the first year of college: Additional evidence. *J. Coll. Stud. Dev.* **2012**, *53*, 243–266. [CrossRef]
35. Pascarella, E.T.; Pierson, C.T.; Wolniak, G.C.; Terenzini, P.T. First-generation college students: Additional evidence on college experiences and outcomes. *J. High. Educ.* **2004**, *75*, 249–284. [CrossRef]
36. Ebert, J.; Heublein, U. *Studienabbruch bei Studierenden mit Migrationshintergrund: Eine Vergleichende Untersuchung der Ursachen und Motive des Studienabbruchs bei Studierenden mit und ohne Migrationshintergrund auf Basis der Befragung der Exmatrikulierten des Sommersemesters 2014*; DZHW: Hannover, Germany, 2017.
37. Janke, S.; Rudert, S.C.; Marksteiner, T.; Dickhäuser, O. Knowing one's place: Parental educational background influences social identification with academia, test anxiety, and satisfaction with studying at university. *Front. Psychol.* **2017**, *8*, 1326. [CrossRef] [PubMed]
38. Lörz, M. Intersektionalität im Hochschulbereich: In welchen Bildungsphasen bestehen soziale Ungleichheiten nach Migrationshintergrund, Geschlecht und sozialer Herkunft—und inwieweit zeigen sich Interaktionseffekte? *Z. Erzieh.* **2019**, *22*, 101–124. [CrossRef]
39. Gaspard, H.; Lauermann, F.; Rose, N.; Wigfield, A.; Eccles, J.S. Cross-domain trajectories of students' ability self-concepts and intrinsic values in math and language arts. *Child Dev.* **2020**, *91*, 1800–1818. [CrossRef] [PubMed]
40. Meyer, J.; Fleckenstein, J.; Köller, O. Expectancy value interactions and academic achievement: Differential relationships with achievement measures. *Contemp. Educ. Psychol.* **2019**, *58*, 58–74. [CrossRef]
41. Goldman, J.; Heddy, B.C.; Cavazos, J. First-generation college students' academic challenges understood through the lens of expectancy value theory in an introductory psychology course. *Teach. Psychol.* **2022**, *49*, 37–48. [CrossRef]
42. Tonks, S.M.; Wigfield, A.; Eccles, J.S. Expectancy-value theory in cross-cultural perspective: What have we learned in the last 15 years? In *Big Theories Revisited 2*; Liem, G.A.D., McInerney, D.M., Eds.; IAP—Information Age Publishing, Inc.: Charlotte, NC, USA, 2018; pp. 96–116.
43. Wigfield, A.; Eccles, J.S. The development of competence, belief, expectancies for success, and achievement values from childhood through adolescence. In *Development of Achievement Motivation: Educational Psychology*; Wigfield, A., Eccles, J.S., Eds.; Academic Press: San Diego, CA, USA, 2002; pp. 91–120. [CrossRef]
44. Else-Quest, N.M.; Hyde, J.S.; Linn, M.C. Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychol. Bull.* **2010**, *136*, 103–127. [CrossRef]
45. Goldman, A.D.; Penner, A.M. Exploring international gender differences in mathematics self-concept. *Int. J. Adolesc. Youth* **2016**, *21*, 403–418. [CrossRef]
46. Donohue, D. Culture, cognition, and college: How do cultural values and theories of intelligence predict students' intrinsic value for learning? *JCVE* **2020**, *4*, 1–14. [CrossRef]

47. Mok, S.Y.; Bakaç, C.; Froehlich, L. 'My family's goals are also my goals': The relationship between collectivism, distal utility value, and learning and career goals of international university students in Germany. *Int. J. Educ. Vocat. Guid.* **2021**, *21*, 355–378. [CrossRef]
48. Wisniewski, K.; Lenhard, W.; Möhring, J.; Spiegel, L. (Eds.) *Sprache und Studienerfolg bei Bildungsausländerinnen und Bildungsausländern*; Waxmann: Münster, Germany; New York, NY, USA, 2022.
49. Glorius, B. Gekommen, um zu bleiben? Der Verbleib internationaler Studierender in Deutschland aus einer Lebenslaufperspektive. *Raumforsch. Raumordn.* **2016**, *74*, 361–371. [CrossRef]
50. Zimmermann, J.; Greischel, H.; Jonkmann, K. The development of multicultural effectiveness in international student mobility. *High. Educ.* **2021**, *82*, 1071–1092. [CrossRef]
51. Berry, J.W. Acculturation: A conceptual overview. In *Monographs in Parenting Series: Acculturation and Parent-Child Relationships: Measurement and Development*, 1st ed.; Bornstein, M.H., Cote, L.R., Eds.; Lawrence Erlbaum: Boca Raton, FL, USA, 2006; pp. 13–32.
52. Berry, J.W. Lead Article—Immigration, acculturation, and adaptation. *Appl. Psychol.* **1997**, *46*, 5–34. [CrossRef]
53. Bierwiaczonek, K.; Kunst, J.R. Revisiting the integration hypothesis: Correlational and longitudinal meta-analyses demonstrate the limited role of acculturation for cross-cultural adaptation. *Psychol. Sci.* **2021**, *32*, 1476–1493. [CrossRef] [PubMed]
54. Crenshaw, K. *On Intersectionality: Essential Writings*; The New Press: New York, NY, USA, 2017. Available online: https://permalink.obvsg.at/AC15054604 (accessed on 26 July 2023).
55. Ferree, M.M. Filling the glass: Gender perspectives on families. *J. Marriage Fam.* **2010**, *72*, 420–439. [CrossRef]
56. MacPhee, D.; Farro, S.; Canetto, S.S. Academic self-efficacy and performance of underrepresented STEM majors: Gender, ethnic, and social class patterns. *Anal. Soc. Issues Public Policy* **2013**, *13*, 347–369. [CrossRef]
57. Riegle-Crumb, C.; Moore, C.; Ramos-Wada, A. Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity. *Sci. Ed.* **2010**, *95*, 458–476. [CrossRef]
58. Blaney, J.M.; Stout, J.G. Examining the relationship between introductory computing course experiences, self-efficacy, and belonging among first-generation college women. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*; Caspersen, M.E., Edwards, S.H., Barnes, T., Garcia, D.D., Eds.; ACM: New York, NY, USA, 2017; pp. 69–74.
59. Else-Quest, N.M.; Mineo, C.C.; Higgins, A. Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychol. Women Quart.* **2013**, *37*, 293–309. [CrossRef]
60. Seo, E.; Shen, Y.; Alfaro, E.C. Adolescents' beliefs about math ability and their relations to STEM career attainment: Joint consideration of race/ethnicity and gender. *J. Youth Adolesc.* **2019**, *48*, 306–325. [CrossRef] [PubMed]
61. Falk, S.; Thies, T.; Yildirim, H.H.; Zimmermann, J.; Kercher, J.; Pineda, J. *Methodenbericht zum "International Student Survey" aus dem Projekt "Studienerfolg und Studienabbruch bei Bildungsausländern in Deutschland im Bachelor- und Masterstudium" (SeSaBa): Release 2.0*; SSOAR, GESIS-Leibniz-Institut für Sozialwissenschaften e.V.; Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung (IHF): Mannheim, München, 2021.
62. Beierlein, C.; Kovaleva, A.; Kemper, C.J.; Rammstedt, B. *ASKU—Allgemeine Selbstwirksamkeit Kurzskala*; ZPID (Leibniz Institute for Psychology Information)-Testarchiv: Trier, Germany, 2012. [CrossRef]
63. Blanz, M. *Forschungsmethoden und Statistik für die Soziale Arbeit: Grundlagen und Anwendungen*, 2nd ed.; Kohlhammer: Stuttgart, Germany, 2015.
64. Westermann, R.; Heise, E.; Spies, K.; Trautwein, U. Identifikation und Erfassung von Komponenten der Studienzufriedenheit. *Psychol. Erzieh. Und Unterr.* **1996**, *43*, 1–22.
65. Heise, E.; Thies, B. Die Bedeutung von Diversität und Diversitätsmanagement für die Studienzufriedenheit. *Z. Padagog. Psychol.* **2015**, *29*, 31–39. [CrossRef]
66. Gaspard, H.; Dicke, A.-L.; Flunger, B.; Brisson, B.M.; Häfner, I.; Nagengast, B.; Trautwein, U. Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Dev. Psychol.* **2015**, *51*, 1226–1240. [CrossRef] [PubMed]
67. DAAD; DZHW. *Wissenschaft Weltoffen 2022: Daten und Fakten zur Internationalität von Studium und Forschung in Deutschland*; wbv Media: Bielefeld, Germany, 2022.
68. Demes, K.A.; Geeraert, N. Measures matter. *J. Cross Cult. Psychol.* **2014**, *45*, 91–109. [CrossRef]
69. IBM Corp. *IBM SPSS Statistics for Windows (Version 28.0)*; Released; IBM Corp.: Armonk, NY, USA, 2021.
70. Muthén, L.K.; Muthén, B.O. *Mplus Users Guide*, 7th ed.; Muthén & Muthén: Los Angeles, CA, USA, 1998–2015.
71. Faul, F.; Erdfelder, E.; Lang, A.-G.; Buchner, A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **2007**, *39*, 175–191. [CrossRef]
72. Faul, F.; Erdfelder, E.; Buchner, A.; Lang, A.-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* **2009**, *41*, 1149–1160. [CrossRef]
73. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Erlbaum: Hillsdale, NJ, USA, 1988. Available online: http://www.loc.gov/catdir/enhancements/fy0731/88012110-d.html (accessed on 26 July 2023).
74. DAAD; DZHW. *Wissenschaft Weltoffen 2019: Daten und Fakten zur Internationalität von Studium und Forschung in Deutschland. Fokus: Studienland Deutschland—Motive und Erfahrungen internationaler Studierender*; DAAD; DZHW: Bielefeld, Germany, 2019. Available online: https://www.wissenschaft-weltoffen.de/content/uploads/2021/09/wiwe_2019_verlinkt.pdf (accessed on 26 July 2023).
75. Lee, H.R.; von Keyserlingk, L.; Arum, R.; Eccles, J.S. Why do they enroll in this course? Undergraduates' course choice from a motivational perspective. *Front. Educ.* **2021**, *6*, 641254. [CrossRef]

76. Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF). *Deutschland ist international Spitzenreiter in MINT*; BMBF: Berlin, Germany, 2022.
77. Stifterverband; McKinsey & Company. *Vom Arbeiterkind zum Doktor: Der Hürdenlauf auf dem Bildungsweg der Erststudierenden*; Stifterverband für die Deutsche Wissenschaft e.V.: Essen, Germany, 2021.
78. Schnettler, T.; Scheunemann, A.; Bäulke, L.; Thies, D.O.; Kegel, L.S.; Bobe, J.; Dresel, M.; Fries, S.; Leutner, D.; Wirth, J.; et al. Studienmotivation und Studienerfolg: Zusammenhänge zwischen Erwartung, Wert und motivationalen Kosten mit Studienleistungen und Studienabbruchintentionen [Paper presentation]. In Proceedings of the 10th Annual Conference Gesellschaft für Empirische Bildungsforschung GEBF, Duisburg-Essen, Germany, 2 March 2023.
79. Ryan, R.M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68–78. [CrossRef]
80. Froehlich, L.; Tsukamoto, S.; Morinaga, Y.; Sakata, K.; Uchida, Y.; Keller, M.M.; Stürmer, S.; Martiny, S.E.; Trommsdorff, G. Gender stereotypes and expected backlash for female STEM students in Germany and Japan. *Front. Educ.* **2022**, *6*, 793486. [CrossRef]
81. Poort, I.; Jansen, E.; Hofman, A. Promoting university students' engagement in intercultural group work: The importance of expectancy, value, and cost. *Res. High. Educ.* **2023**, *64*, 331–348. [CrossRef]
82. Gaspard, H.; Dicke, A.-L.; Flunger, B.; Schreier, B.; Häfner, I.; Trautwein, U.; Nagengast, B. More value through greater differentiation: Gender differences in value beliefs about math. *J. Educ. Psychol.* **2015**, *107*, 663–677. [CrossRef]
83. Bethel, A.; Ward, C.; Fetvadjiev, V.H. Cross-cultural transition and psychological adaptation of international students: The mediating role of host national connectedness. *Front. Educ.* **2020**, *5*, 539950. [CrossRef]
84. Suanet, I.; van de Vijver, F.J.R. Perceived cultural distance and acculturation among exchange students in Russia. *J. Community Appl. Soc. Psychol.* **2009**, *19*, 182–197. [CrossRef]
85. Hsieh, T.; Simpkins, S.D.; Eccles, J.S. Gender by racial/ethnic intersectionality in the patterns of adolescents' math motivation and their math achievement and engagement. *Contemp. Educ. Psychol.* **2021**, *66*, 101974. [CrossRef]
86. Else-Quest, N.M.; Grabe, S. The political is personal. *Psychol. Women Quart.* **2012**, *36*, 131–144. [CrossRef]
87. Zarrett, N.; Malanchuk, O.; Davis-Kean, P.E.; Eccles, J.S. Examining the gender gap in IT by race: Young adults' decisions to pursue an IT career. In *Women and Information Technology*; Cohoon, J., Aspray, W., Eds.; The MIT Press: Cambridge, MA, USA, 2006; pp. 55–88.
88. Wisniewski, K.; Möhring, J.; Lenhard, W.; Seeger, J. Zum Zusammenhang sprachlicher Kompetenzen mit dem Studienerfolg von Bildungsausländer/-innen im ersten Studiensemester. In *Testen Bildungssprachlicher Kompetenzen und Akademischer Sprachkompetenzen: Zugaenge Für Schule und Hochschule*; Drackert, A., Mainzer-Murrenhoff, M., Soltyska, A., Timukova, A., Eds.; Peter Lang GmbH, Internationaler Verlag der Wissenschaften: Berlin, Germany, 2020; pp. 279–319.
89. Rosenzweig, E.Q.; Wigfield, A.; Hulleman, C.S. More useful or not so bad? Examining the effects of utility value and cost reduction interventions in college physics. *J. Educ. Psychol.* **2020**, *112*, 166–182. [CrossRef]
90. Rosenzweig, E.Q.; Song, Y.; Clark, S. Mixed effects of a randomized trial replication study testing a cost-focused motivational intervention. *Learn. Instr.* **2022**, *82*, 101660. [CrossRef]
91. Middendorff, E. Studienbelastung im Bachelor-Studium—alles nur gefühlt? Befunde der 19. Sozialerhebung des DSW durchgeführt vom HIS-Institut für Hochschulforschung. In *Rückenwind—Was Studis gegen Stress tun können: Ein Ratgeber mit informativen Texten und hilfreichen Tipps zum Umgang mit Stress für Studierende und Hochschulen*; [Beiträge, die in den vergangenen drei Jahren bei den Karlsruher Stresstagen entstanden sind]; Duriska, M., Ed.; KIT Karlsruher Institut für Technologie: Karlsruhe, Germany, 2011; pp. 42–45.
92. Schunk, D.H.; Pajares, F. The development of academic self-efficacy. In *Development of Achievement Motivation: Educational Psychology*; Wigfield, A., Eccles, J.S., Eds.; Academic Press: San Diego, CA, USA, 2002; pp. 15–31.
93. Serrano-Sánchez, J.; Zimmermann, J.; Jonkmann, K. When in Rome... A longitudinal investigation of the predictors and the development of student sojourners' host cultural behavioral engagement. *Int. J. Intercult. Rel.* **2021**, *83*, 15–29. [CrossRef]
94. Holtrop, D.; Hughes, A.W.; Dunlop, P.D.; Chan, J.; Steedman, G. Do social desirability scales measure dishonesty? *Eur. J. Psychol. Assess.* **2021**, *37*, 274–282. [CrossRef]
95. Mummendey, H.D. Methoden und Probleme der Kontrolle sozialer Erwünschtheit (Social Desirability). *Z. Differ. Diagn. Psychol.* **1981**, *2*, 199–218.
96. Harackiewicz, J.M.; Canning, E.A.; Tibbetts, Y.; Priniski, S.J.; Hyde, J.S. Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *J. Pers. Soc. Psychol.* **2016**, *111*, 745–765. [CrossRef] [PubMed]

education sciences

Article

Increasing Self-Concept and Decreasing Gender Stereotypes in STEM through Professional Development for Early Childhood Educators

Sarah Feierabend [1,*], Sabine Hasenhütl [2], Manuela Paechter [2], Silke Luttenberger [3], Marina W. T. Eglmaier [4] and Lars Eichen [1]

1 Institute of Education Research and Teacher Education, University of Graz, 8010 Graz, Austria; lars.eichen@uni-graz.at
2 Institute of Psychology, University of Graz, 8010 Graz, Austria; sabine.hasenhuetl@uni-graz.at (S.H.); manuela.paechter@uni-graz.at (M.P.)
3 Practical Education and Action Research, University College of Teacher Education Styria, 8010 Graz, Austria; silke.luttenberger@phst.at
4 Division of Physiology, Otto Loewi Research Center, Medical University of Graz, 8010 Graz, Austria; m.eglmaier@medunigraz.at
* Correspondence: sarah.feierabend@uni-graz.at

Abstract: Starting early in life, children, especially girls, experience obstacles when it comes to developing an interest in STEM. Although early childhood (EC) educators face an important task in promoting girls (and boys) in STEM, they often face challenges in doing so. Therefore, it is crucial for EC educators to cultivate positive attitudes, self-concepts, and STEM skills. To address these identified issues, a three-month professional development program was created for EC educators. This professional development program was evaluated using a pre–post design with a focus on the self-concept and gender stereotypes of EC educators. The program involved 30 female EC educators in evaluating these aspects. The statistical analyses show positive results in enhancing educators' self-concepts and reducing gender stereotypes over the course of this professional development program. The results suggest the potential of the blended learning design in this professional development program and indicate that this program could serve as a promising model for future interventions.

Keywords: STEM education; self-concept; gender stereotypes; gender and STEM; early childhood education; professional development; early childhood educators

Citation: Feierabend, S.; Hasenhütl, S.; Paechter, M.; Luttenberger, S.; Eglmaier, M.W.T.; Eichen, L. Increasing Self-Concept and Decreasing Gender Stereotypes in STEM through Professional Development for Early Childhood Educators. *Educ. Sci.* **2024**, *14*, 81. https://doi.org/10.3390/educsci14010081

Academic Editor: James Albright

Received: 24 October 2023
Revised: 23 December 2023
Accepted: 3 January 2024
Published: 11 January 2024

1. Introduction

Proficiency and interest in STEM empower the succeeding generation to actively participate in shaping a sustainable, technology-oriented future [1]. Early childhood lays the foundation for the development of STEM proficiency and interest [2]. Positive learning experiences enhance children's interest, boost their confidence in their STEM abilities [3,4], contribute to their later academic success [5,6], and may even impact career choices [7,8]. However, already in early childhood, disparities in STEM to the disadvantage of girls can be observed. They manifest in negative self-assessments and attitudes, such as a lower self-concept, lower interest, and gender stereotypes [9–12], as well as lower STEM achievements and stereotyped career wishes [13,14].

Children's development in STEM is shaped by educators who wield significant influence, molding children's thoughts, behaviors, and attitudes through their own actions, communication, and expectations [15–18]. However, EC educators, too, can be influenced by gender stereotypes and harbor negative self-assessments. In light of these challenges, a professional development program focusing on gender-sensitive STEM pedagogy was developed and evaluated to determine its potential to positively impact EC educators' self-assessments and stereotypes.

1.1. Self-Assessments and Stereotypes as a Barrier to Gender-Sensitive STEM Education

From early childhood, individuals develop self-concepts, i.e., temporally relatively stable conceptions and evaluations of themselves, in different domains [19]. The self-concept in domains such as STEM (often referred to as academic STEM self-concept) develops through earliest experiences in the family, kindergarten, school, etc., and interpretations of one's own environment in relation to feelings of self-confidence, competence, and capability [20,21]. Evaluations from significant others, reinforcements, and attributions of one's own behavior have an influence on the formation of the self-concept [22]. With the start of vocational training, a professional self-concept is added. It refers to how individuals perceive themselves within their professional or work-related roles and encompasses whether an individual sees themselves as proactive and self-responsible in their work [23,24]. Such self-assessments within the self-concept influence EC educators' instructional behavior and, consequently, the achievements of children [25].

Not only the self-concept but also stereotypes develop already in early childhood. Gender stereotypes can be defined as shared beliefs that link women and men with specific characteristics, abilities, and interests while mostly judging women as less competent and less committed [26]. For Austria (the country in which the study was conducted), this can be attributed to the fact that in Austria and generally in the German-speaking countries, traditional family structures and gender-specific role divisions are still strongly adhered to. This further contributes to the reinforcement of gender stereotypes and roles [27,28]. Gender stereotypes have extensive effects on individuals, influencing the development of academic self-concept in STEM subjects, as well as self-concept within one's profession.

Given that early childhood education as a profession is predominantly female and closely associated with interest and expertise in the social domain [29], it is not surprising that EC educators are susceptible to self-critical evaluations and gender stereotypes, both regarding themselves and women/girls in general [30–32]. For example, a study in Sweden revealed that educators are influenced by preconceived notions about children's STEM development and believe that gender differences are innate [33].

Research in the field of early childhood education indicates that self-concepts and stereotypes held by EC educators wield a significant influence on their interactions with children. For instance, educators' self-concepts are closely tied to their actual professional knowledge and skills. A low self-concept has been associated with limited foundational knowledge and pedagogical skills for effectively introducing STEM topics to children [34]. Conversely, a positive assessment of STEM skills and genuine enjoyment of STEM correlate with an increased frequency of providing STEM activities [35]. Prior success and enjoyment in delivering STEM activities for children can further bolster an educator's self-concept. In a study conducted by Erden and Sönmez [36], positive prior experiences were positively correlated with positive attitudes among EC educators towards science education. In contrast, uncertainties about one's knowledge and skills, negative emotions, or even anxiety stemming from EC educators' past school experiences with science subjects can have a detrimental impact on their later professional work [37].

Educators may manifest stereotypes about the interests of girls and boys in STEM in various ways. This can involve restricting or directing children's play choices based on activities traditionally deemed suitable for boys or girls, focusing more on developing boys' abilities in what is perceived as masculine domains, or generally holding different expectations for girls and boys [33,38–40]. As Wang [41] points out, merely verbally stating an intention to have the same expectations for boys and girls in STEM does not necessarily equate to educators acting without bias in their professional practice. Educators may unintentionally convey to girls that STEM is a boy-dominated field, imposing negative impacts on girls' STEM development. These findings illuminate the conflicting beliefs and practices of EC educators and underscore the critical role that their stereotypes and self-concepts play when working with children, potentially transferring their expectations, attitudes, or uncertainties. They also emphasize the necessity of continuously addressing and expanding EC educators' beliefs, attitudes, and knowledge in vocational training and professional development [32].

1.2. Professional Development Concept for EC Educators

The previous explanations have illustrated the challenges that EC educators face in STEM and have highlighted the need for appropriate professional development. This holds also (or even especially) true for EC educators in German-speaking countries who often lack expertise in STEM pedagogy. This knowledge gap is a direct result of inadequate initial training and the absence of continuous professional development [42–44]. To address this issue, a comprehensive professional development concept tailored to the needs and the education system in Austria was developed. Through professional development, newly acquired skills and knowledge can be immediately applied in practice. It is believed that these acquired competencies may have a positive impact on self-concept [45]. Its primary objective is to enhance the STEM pedagogical skills of EC educators, with a specific emphasis on gender-sensitive education as well as on educators' attitudes and stereotypes.

Theoretical model for a professional development program. The concept aligns with the SciMath-DLL Professional Development Three-Component Model by Brenneman and colleagues [46]. This theoretical model encompasses the acquisition of knowledge in STEM fields as well as pedagogical knowledge and emphasizes the significance of opportunities to apply the newfound knowledge in practical pedagogical settings. It also aims to explore EC educators' beliefs and attitudes towards STEM while supporting opportunities for discussion, reflection, and the creation of a community of learners. By incorporating these components, the intention is to bolster EC educators' pedagogical abilities, encourage critical reflection on their professional activities, enhance motivation, and cultivate positive attitudes toward STEM subjects. The Three-Component Model considers cognitive learning objectives such as understanding, applying knowledge and procedures, evaluating or creating (e.g., developing learning units). In addition, it addresses affective learning objectives, such as values, esteem, and attitudes [46–49].

Advancement of a professional development program. For the newly involved professional development program, a blended learning design was chosen, interweaving online and face-to-face learning in different learning phases (see Table 1) and presenting a flexible and learner-friendly alternative to traditional face-to-face instruction [50]. Especially in early childhood education, online professional development, such as blended learning, has gained importance in improving educators' knowledge and skills, thus advancing the professionalization of the field [51].

Table 1. Temporal design, measures.

	Kick-Off Event, Introduction (Phase 1)	Online Modules (Phase 2)	In-Person Workshop (Phase 3)	Implementation (Phase 4)	Closing Event (Phase 5)
Point in Time	**t1**	**t2**	**t3**	**t4**	**t5**
Aims, setup	Opening session; getting to know each other, reflection on own practice; online, synchronous.	Self-regulated learning online, asynchronous. Reflection and discussions.	In-person workshop with didactic examples.	Implementation of the didactic examples. Reflection and discussions.	Exchange of experiences; online synchronous.
Assessments	Academic and professional self-concept, stereotypes.				Overall evaluation. Self-concept, stereotypes.

After an (online) introduction to the course (phase 1), the focus is on fundamentals and background knowledge for teaching and learning in STEM (phase 2). In several learning units, overarching themes such as “gender-sensitive teaching”, “motivation and interest”, or “impact of stereotypes” are addressed. Each learning unit commences with a more theoretically oriented introduction and progresses forward to practical implications and pedagogical practice. Phase 2 is implemented as asynchronous online learning, providing participants with the flexibility to engage with the learning materials independently, at their convenience. In accordance with the Three-Component Model [46] this self-regulated approach is complemented with tasks, opportunities for online discussions, and reflective dialogues with both fellow participants and the course instructors.

The main segment of the professional development program on didactic skills and expertise in STEM is conducted as a face-to-face workshop (phase 3). The focus is on enabling participants to translate knowledge of gender-sensitive STEM didactics into pedagogical activities. The fundamentals for the topic are illustrated with didactic good practice examples from different STEM areas, with the age range of four- to five-year-old children in focus. For science, didactic examples on the topic “human body” had been developed, the “principle of chain reactions” for technology, “programming unplugged” ideas for engineering and “measures and sizes” using the example of the human body for mathematics. The intent behind designing the didactic examples was to make sure participants had ample opportunity to put their own ideas into practice. For children, the didactic examples emphasize hands-on experiences, achieving immediate learning success, and fostering a positive self-concept and interest. The examples are enriched by female role models, which offer girls in particular opportunities for identification. In the face-to-face workshop, the EC educators are introduced to these didactic examples and the materials, they try them out for themselves, and discuss their applicability to their everyday professional life. Care was taken to ensure that the examples provided the EC educators with ample room to implement their own creative ideas. All documents and materials are also available online, enabling participants to delve deeper into the knowledge acquired during the workshop.

Following the face-to-face workshop, the participants put their acquired skills into practice at the workplace (phase 4). For this purpose, they can borrow all the learning and teaching materials from the course provider. As a result, the workshop not only provides EC educators with practical experience in teaching the content during the workshops but also allows them to continue applying their skills in their professional settings afterward. The professional development program concludes with an online meeting and an exchange of experiences (phase 5).

All learning units are accompanied by discussions and reflections on what has been learned and its significance for one’s own profession and attitudes. Feedback is provided by the course instructors as well as by fellow course participants. Not only cognitive learning objectives but also objectives that focus on the development of positive attitudes and mindsets are significant for the professional development program, not least because the educators are important role models, especially female educators for girls [52].

2. Research Questions

The present study aims to evaluate this professional development program regarding its effects on the participants’ self-concept and gender stereotypes. The following research questions and hypotheses are being investigated:

1. How do participants evaluate the professional development program? As the evaluation is an exploratory question, no hypothesis was formed in this regard.
2. Is participation in the professional development program related to changes in educators’ self-concept?

H1. *In comparison to a control group that does not participate in the professional development program, the participants’ self-concept will change over the course of the training.*

3. Is participation in the professional development program related to changes in educators’ gender stereotypes concerning children?

H2. *In comparison to a control group that does not take part in the professional development program, the participants' gender stereotypes concerning children will change over the course of the training.*

3. Method

3.1. Samples and Study Design

Samples. The professional development program described in 1.2. was evaluated by using a study design with a training group and a control group.

The training group included EC educators who had voluntarily registered for a professional development program on promoting girls in STEM. Participation was acknowledged as part of the annual continuing education hours that EC educators are required to complete on a regular basis. Altogether, 31 EC educators, all employed in an early childhood education facility, participated in the professional development program (30 women, one man; data analyses were carried out only for the female participants). The age range varied from 20 to 54 years (M = 33.97, SD = 10.105). The participants' work experience varied from one to 35 years (M = 11.38, SD = 9.777); 58.1% had never taken any STEM professional development program. Regarding occupational positions, the training group had 9 (30%) participants in management roles with childcare responsibilities, 14 (46.7%) worked as EC educators as group leaders, 7 (23.3%) had other functions.

The control group included 25 female EC educators, all employed in an early childhood education facility. They did not participate in the professional development program and were recruited from different early childhood education facilities. The age range for the participants in the control group varied between 20 and 45 years (M = 30.72, SD = 6.262). Work experience varied from one to 24 years (M = 9.29, SD = 6.682); 48% had never taken any STEM training. Regarding occupational positions, the control group had one (4%) participant in management roles with childcare responsibilities, 13 (52%) worked as EC educators as group leaders, 11 (44%) had other functions.

Study design. The professional development program was implemented as a blended-learning design with online and in-person learning phases. It was conducted at the authors' institution and lasted 14 weeks with a workload of approximately 40 h. The program comprised five different learning phases with respective modules (see Table 1):

- Phase 1—kick-off meeting. A kick-off meeting marked the start of the professional development program. It was conducted online synchronously, aiming to familiarize participants with each other, giving opportunities for reflection on individual professional experiences and discussing organizational matters;
- Phase 2—self-regulated learning with online modules. After the kick-off meeting, the participants learned with online modules that covered topics such as gender-sensitive didactics, motivation in STEM subjects, stereotypes, and reflections on pedagogical practices and roles. Participants could decide on their own learning time and place;
- Phase 3—in-person workshop. In a two-day long in-person workshop at the institution of the course providers, participants got acquainted with didactic examples pertaining to four different STEM fields;
- Phase 4—implementation of the didactic examples in one's own professional practice. The participants tried out didactic examples at their workplaces;
- Phase 5—closing event. The EC educators gathered online to exchange experiences, reflect on their work, and engage in synchronous discussions with fellow participants about the training.

Table 1 gives an overview of the time schedule and the measures that are relevant for the present study.

The training group members participated at all points in time. They completed different survey instruments at the five time points; only the data collections at time points 1 and 5 are relevant to this study. Members of the control group did not take part in the professional development program and, therefore, did not receive any of its content. They completed the same surveys as the training group at the first point in time (approximately at t1, the start of the professional development program for the training group) and (t5).

3.2. Variables

Gender, age, professional profile, and work experience were recorded at the kick-off event (t1) for the training group and at the first survey for the control group (t1). The training and the control group completed assessments on different facets of the self-concept and on gender stereotypes at the first point in time (t1).

Additionally, the training group evaluated the training at the closing event (t5) by three items. Overall satisfaction, assessment of learning achievements, and usefulness of the training for professional practice were measured by three items ("overall, I am with the training . . . 1 = not satisfied to 6 = very satisfied"; "altogether I have learned . . . 1 = little to 6 = a lot"; "the training is . . . 1 = not useful to 6 = very useful for my professional practice").

Three facets of self-concept were measured.

Social academic self-concept in STEM was measured by the respective scale of the SASK (Skalen zum akademischen Selbstkonzept (Academic Self Concept Scales) [20]. It comprises four items by which participants rate their abilities in comparison to others on a 7-point-Likert scale (see example item: "I believe I am. . . 1 = less talented to 7 = more talented. . . in STEM than other people). Reliability calculated by Cronbach's α calculated for the whole sample at the two points in time was good, with $\alpha = 0.82$ at t1 and $\alpha = 0.92$ at t5.

Absolute self-concept was also measured by the SASK. It comprises four items by which participants rate their overall abilities without a reference to others (example item: I assess my talent in STEM as being . . . 1 = low to 7 = high). Reliability was very good, with $\alpha = 0.91$ at t1 and $\alpha = 0.92$ at t5.

Professional self-concept was measured by eight items, six from the Monitoring Report [53] of the "Haus der kleinen Forscher" initiative (House of Little Researchers) and two specifically developed for the present study (example items: "I feel confident in exploring everyday natural phenomena with children"; "I can identify children's interest in STEM topics). Participants rate their self-concept on a 4-point Likert scale, with values ranging from 1 = strongly disagree to 4 = strongly agree. Higher values indicate a more positive self-concept. Reliability was good, with $\alpha = 0.80$ at t1 and $\alpha = 0.86$ at t5.

Two facets of gender stereotypes were measured.

Stereotypes about girls' and boys' interest in STEM were measured by seven items using a 5-point Likert scale, ranging from 1 = strongly disagree to 5 = strongly agree [21] (example item: "girls are not as interested as boys in STEM subjects"). Higher values indicate more pronounced stereotypes. Reliability was good to satisfactory, with $\alpha = 0.80$ at t1 and $\alpha = 0.74$ at t5.

Stereotypes about girls' and boys' abilities in STEM were measured by seven items; four items from the stereotype scale of Ertl et al. [21] (example item: "girls perform not as well as boys at STEM subjects"), one item from the scale of Mösko [54] (example item: "boys are more talented in STEM fields than girls") another one from the scale by Grosch [55] (example item: "among the highly skilled in STEM fields, there are fewer girls than boys"), and one more item from the KomMa questionnaire ("girls need more support in STEM fields") [56]. All items were rated on a 5-point Likert scale, ranging from 1 = strongly disagree to 5 = strongly agree [21]. Higher values indicate more pronounced stereotypes. Reliability was good, with $\alpha = 0.87$ at t1 and $\alpha = 0.85$ at t5.

All analyses were carried out using IBM SPSS Statistics 29. Hypotheses (as described in Section 2) were tested by MANOVA and *t*-tests. For MANOVA results, effect sizes are described with $\eta^2 < 0.06$ indicating a small effect, $\eta^2 \geq 0.06$ a medium and $\eta^2 > 0.14$ a large effect. The study was performed in accordance with the American Psychological Association's Ethics Code and the Declaration of Helsinki.

4. Results

4.1. Evaluation of the Professional Training Program

The first research question referred to the evaluation of the professional development program. Overall, the training was evaluated very positively, with mean values above 5.0 on the 6-point Likert scale (see Table 2). On all items, at least 50% of the participants evaluated the training with the highest value.

Table 2. Descriptive statistics for evaluation variables, training sample (n = 30).

Overall Satisfaction				**Assessment of Learning Achievements**				**Usefulness of Training**			
M	***SD***	**min**	**max**	***M***	***SD***	**min**	**max**	***M***	***SD***	**min**	**max**
5.20	1.031	3	6	5.10	1.269	2	6	5.43	0.898	3	6

Note. n = 30; all assessments measured on a 6-point scale. Higher values indicate a more positive assessment.

4.2. MANOVA Results for the Social, Absolute, and Professional Self-Concept

Self-concept was measured at two points in time. Table 3 describes the descriptive statistics for the variables.

Table 3. Descriptive statistics for the self-concept variables, whole sample, training and control group at t1 and t5.

	Whole Sample (n = 55)				**Training Group (n = 30)**				**Control Group (n = 25)**			
Variable, Point in Time	***M***	***SD***	**min**	**max**	***M***	***SD***	**min**	**max**	***M***	***SD***	**min**	**max**
Social self-concept (t1)	4.31	0.829	2.25	6.00	4.64	0.635	3.00	6.00	3.91	0.869	2.25	5.75
Social self-concept (t5)	4.57	0.993	2.00	7.00	4.85	0.703	4.00	6.25	4.19	1.171	2.00	7.00
Absolute self-concept (t1)	4.73	0.975	2.75	7.00	5.02	0.777	3.50	6.50	4.39	1.09	2.75	7.00
Absolute self-concept (t5)	4.87	1.043	2.25	7.00	5.28	0.726	3.75	6.75	4.39	1.17	2.25	7.00
Professional self-concept (t1)	3.25	0.408	2.13	4.00	3.29	0.370	2.50	3.88	3.19	0.449	2.13	4.00
Professional self-concept (t5)	3.41	0.455	2.00	4.00	3.58	0.315	2.75	4.00	3.20	0.512	2.00	4.00

Note. Social and absolute self-concept were measured on a 7-point scale; professional self-concept was measured on a 4-point scale. Higher values indicate a higher (more positive) self-concept.

Initial differences between the control group and the training group at test time 1 were evaluated using t-tests. These tests were applied to the social self-concept $t(43{,}087) = 3.502$, $p < 0.001$ *, the absolute self-concept $t(42{,}334) = 2.410$, $p = 0.020$ and the professional self-concept $t(46{,}525) = 0.904$, $p = 0.371$. Considering Bonferroni Alpha correction for three variables, p-values below $p = 0.016$ can be considered significant. Therefore, only the t-test for the social self-concept indicates a significant difference.

The second research question addresses whether a change in the direction of a more positive self-concept can be observed due to the training, i.e., whether an interaction between group × time can be observed. A multivariate analysis of variance (MANOVA) with the factors group and time was carried out with the factors group and time. Overall, MANOVA yielded significant results for the factors group (Pillai's $F(3{,}51) = 4.232$, $p = 0.010$, $\eta^2 = 0.199$), time (Pillai's $F(3{,}51) = 4.290$, $p = 0.009$, $\eta^2 = 0.202$), and for the interaction group × time (Pillai's $F(3{,}51) = 3.772$, $p = 0.016$, $\eta^2 = 0.182$).

Univariate analyses for the single variables show a significant interaction group × time for the professional self-concept (see Table 4). While the values for the control group remained nearly the same, they increased for the training group.

Figures 1–3 depict the results for the interaction. The training group starts on all three variables with a higher self-concept. A significant interaction in the sense that the training group improves while values stay the same for the control group can only be observed for the professional self-concept.

Table 4. MANOVA results for single variables (*F*-values, *df*, *p*-values, η^2) for factors time, group, group × time ($n = 55$).

	Variable	*F*	*df*	*p*-Value	η^2
Group	Social self-concept	12.287	1	0.001	0.188
	Absolute self-concept	9.791	1	0.003	0.156
	Professional self-concept	5.937	1	0.018	0.101
Time	Social self-concept	6.386	1	0.015	0.108
	Absolute self-concept	2.482	1	0.121	0.045
	Professional self-concept	8.946	1	0.004	0.144
Group × time	Social self-concept	0.020	1	0.887	0.001
	Absolute self-concept	2.482	1	0.121	0.045
	Professional self-concept	7.799	1	0.007	0.128

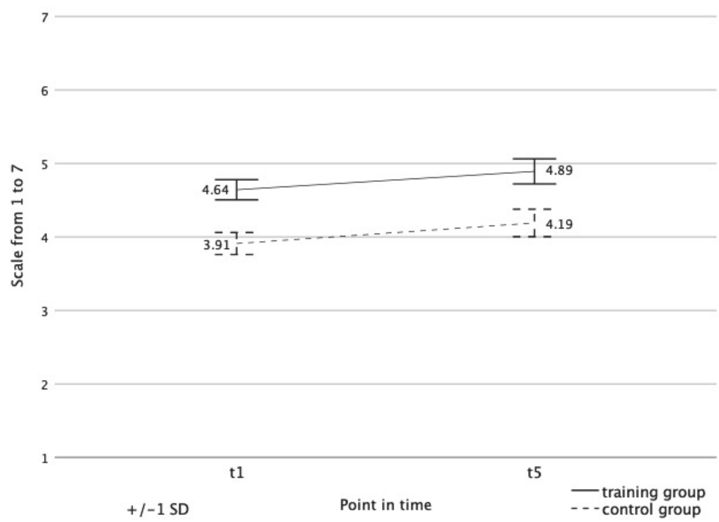


Figure 1. Social self-concept.

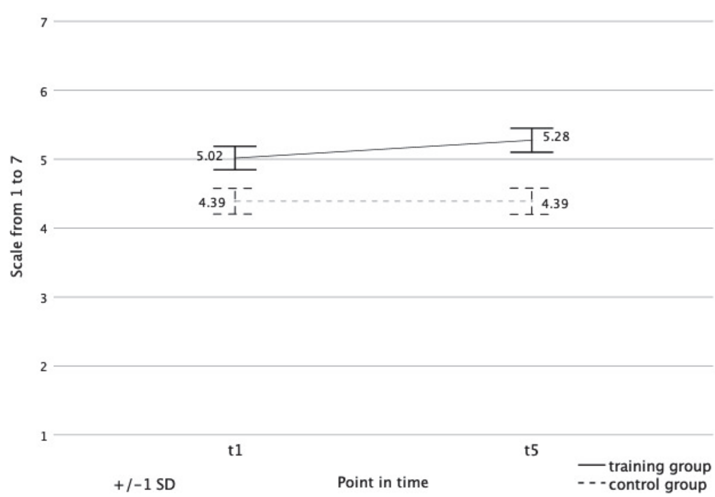


Figure 2. Absolute self-concept.

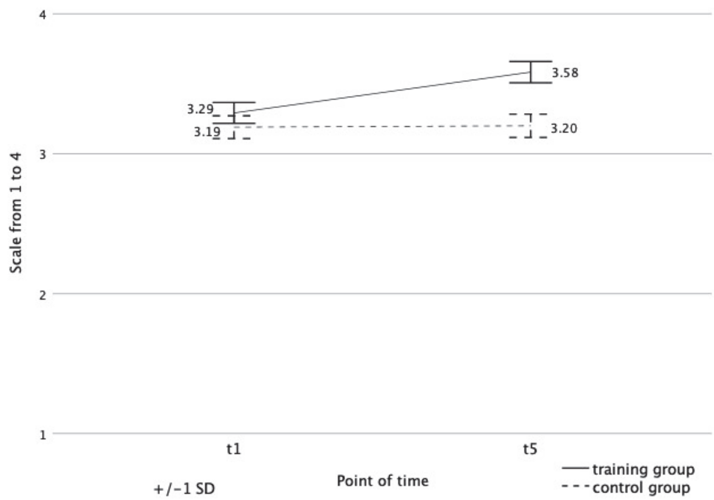


Figure 3. Professional self-concept.

4.3. MANOVA Results for Stereotypes

Two stereotype variables were measured at two points in time in the training and the control group. Table 5 describes the descriptive statistics for the variables.

Table 5. Descriptive statistics for the stereotype variables, whole sample, training and control group at t1 and t5.

	Whole Sample (*n* = 55)				Training Group (*n* = 30)				Control Group (*n* = 25)			
Variable, Point in Time	***M***	***SD***	**min**	**max**	***M***	***SD***	**min**	**max**	***M***	***SD***	**min**	**max**
Stereotypes on interests (t1)	2.58	0.779	1.00	4.25	2.85	0.709	1.00	4.25	2.24	0.741	1.00	3.75
Stereotypes on interests (t5)	2.29	0.691	1.00	3.75	2.35	0.618	1.50	3.75	2.23	0.780	1.00	3.75
Stereotypes on abilities (t1)	1.83	0.736	1.00	3.43	1.92	0.680	1.00	3.43	1.69	0.785	1.00	3.43
Stereotypes on abilities (t5)	1.64	0.640	1.00	3.57	1.68	0.740	1.00	2.57	1.60	0.544	1.00	3.57

Note. Stereotypes were measured on a 5-point scale. Higher values indicate stronger gender stereotypes.

Initial differences between the control group and the training group at test time 1 were evaluated using *t*-tests. These tests were applied to the stereotypes on interests $t(50{,}689) = 3.077$, $p < 0.002$ * and the stereotypes on abilities $t(48{,}3095) = 1.228$, $p = 0.225$. Considering Bonferroni Alpha correction for two variables, *p*-values below $p = 0.025$ can be considered significant. Therefore, only the *t*-test for the stereotypes on interests indicates a significant difference.

The third research question addresses whether a change towards lower stereotype levels can be observed due to the professional development program, i.e., whether an interaction between group × time can be observed. MANOVA with the factors group and time yielded significant results for the factors time (Pillai's $F(2{,}52) = 5.650$, $p = 0.006$, $\eta^2 = 0.179$), group (Pillai's $F(2{,}52) = 3.139$, $p = 0.052$, $\eta^2 = 0.108$), and group × time (Pillai's $F(2{,}52) = 5.109$, $p = 0.009$, $\eta^2 = 0.164$). The univariate results for the two stereotype variables for differences between groups, time and for the interaction group × time are displayed in Table 6.

Table 6. MANOVA results for single variables (*F*-values, *df*, *p*-values, η^2) for factors time, group × time.

	Variable	*F*	*df*	*p*-Value	η^2
Group	Stereotypes on interests	4.223	1	0.045	0.074
	Stereotypes on abilities	0.251	1	0.618	0.005
Time	Stereotypes on interests	9.101	1	0.04	0.147
	Stereotypes on abilities	7.283	1	0.009	0.121
Group × time	Stereotypes on interests	8.434	1	0.005	0.137
	Stereotypes on abilities	6.336	1	0.015	0.107

Figures 4 and 5 depict the changes for the gender stereotype variables. For the stereotypes on interests, the training group starts with more pronounced stereotypes than the control group. A significant interaction in the sense that the training group's stereotypes decrease while values stay the same for the control group can be observed for both variables.

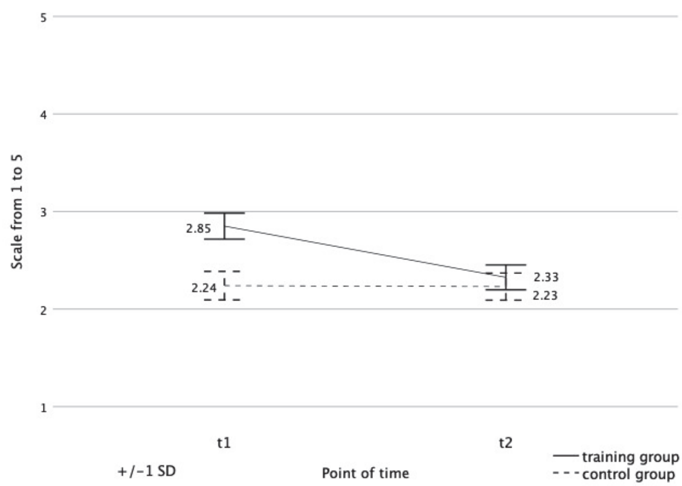


Figure 4. Stereotypes on interest.

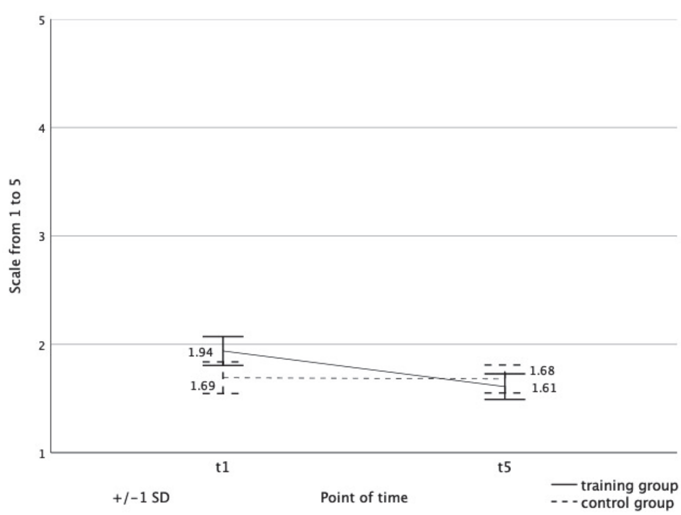


Figure 5. Stereotypes on abilities.

5. Discussion

Research on STEM education underscores the importance of educators' stereotypes and self-assessments in fostering gender-sensitive teaching in STEM. The lack of stereotypes or at least low stereotype levels and positive self-concepts are vital to ensure that both girls and boys have equal opportunities to nurture their interests, develop a positive attitude toward STEM subjects and fulfill their potential in the field [57,58]. In light of this perspective, the professional development concept also aimed to improve educators' STEM self-concepts and decreasing their gender stereotypes.

5.1. Training Participants' Overall Evaluation of the Professional Development Program

An important prerequisite for attitude change is that training is seen as important and useful [59]. Indeed, the participants in the training group evaluated the training very positively: 50% of them rated their overall course satisfaction at the highest level (a score of 6 on the scale), 56.7% reported the highest level of satisfaction with their learning outcomes, and 66.7% found the training highly beneficial for their professional practice. It should be considered that participants voluntarily chose to enroll in the training and registered for it, potentially introducing a bias in the assessments. Nonetheless, it is crucial to highlight that, despite this potential bias, the participants expressed high levels of satisfaction, which is pertinent to interpreting the findings, particularly regarding changes in self-concept and gender stereotypes.

5.2. Changes in Self-Concept

Three facets of the self-concept were assessed: social and absolute self-concept in STEM overall, plus professional self-concept in STEM.

The results indicate that the training had no noticeable impact on the participants' absolute self-concept. Although the social self-concept decreased, this change was not attributed to the group that received the training. It appears that the participants did not translate the training experiences into their self-assessments, even though the didactic examples specifically addressed STEM content. In contrast, a training effect on the participants' professional self-concept could be observed. This outcome can be attributed to the content of the training, which encompassed various elements such as gender-sensitive teaching in STEM, strategies to enhance children's motivation in STEM, the role of stereotypes in education, and providing didactic good practice examples of how to foster both girls' (and boys') engagement in STEM in early childhood education. Although the different self-concepts influence each other, due to the focus of the training program, the professional self-concept might have been more strongly influenced. This outcome is also important as the self-concept inherently includes a motivational dimension, as underscored by Marsh et al. [60] in their Reciprocal Effects Model. A positive self-concept not only relates to perceived competence and performance but also encompasses motivational factors and effort. Therefore, in a professional development setting such as the one investigated, it is important not only to impart professional skills and knowledge but also to address and nurture the participants' self-concept.

In addition, educators have an important role model function for the children entrusted to their care. They play a crucial role as influential socializers, wielding substantial influence over gender disparities in academic motivation, educational choices, and overall achievement [61,62]. Therefore, a positive self-concept in STEM is important not only for educators' pedagogical practice, but also for how the children perceive their educators and develop STEM attitudes and assessments in interaction with these role models [33]. Female educators especially serve as role models for girls; younger girls, in particular, orient themselves strongly to their educators [52]. It is, therefore, all the more important that the educators themselves dispose of an interest, positive self-assessments, and high self-efficacy beliefs in STEM. In this context, alterations in educators' professional self-concept can be regarded as an important outcome of the training.

Based on the results, the second hypothesis can only be partially confirmed. Examining three aspects of the self-concept—the social, absolute, and professional self-concept—significant changes were only observed in the professional self-concept.

5.3. Changes in Stereotypes

Research on STEM shows that (similar to other educators) also EC educators possess stereotypes about girls in STEM (e.g., [41,63]). In the present study, two categories of stereotypes had been measured: stereotypes regarding girls' and boys' interests (assuming lower interest of girls in STEM) and abilities in STEM (assuming inferiority of girls).

When examining the mean values of the study groups on both scales, it is first noticeable that agreement with the items is relatively low, with mean values falling below the scale mean of three and indicating a lower level of stereotypes. Nonetheless, interesting and important distinctions emerge between the means on the two stereotype scales. Statements regarding girls' alleged inferiority in STEM fields are strongly repudiated compared to statements concerning their purported lower interest. These findings have been corroborated in other studies with varying populations (e.g., in Ertl et al. [21] in a sample of female college students). From an educator's perspective, this composition of stereotypes would imply tailoring strategies for girls' interests above all else when teaching STEM and focusing less on alleged ability differences.

The professional development program explicitly included content concerning gender stereotypes in STEM, encompassing their origins and the potential consequences they may have on children's academic journeys, accompanied by discussions and reflections of the educators on their pedagogical role and personal attitudes. Moreover, the didactic examples were tailored to girls, including role models from diverse STEM fields. On the whole, the training appeared to effectively achieve its learning objectives aimed at mitigating stereotypes. The study results revealed training effects for both categories of stereotypes. In the case of interest-related stereotypes, it is noteworthy that the training group initially exhibited higher levels of stereotypes than the control group but managed to reduce these stereotypes through the professional training program while values in the control group stayed the same. As for stereotypes regarding abilities, both groups started with similar baseline values, but only the training group exhibited a reduction in their stereotypes. In light of research that even very young children can harbor gender stereotypes pertaining to play activities, interests, and occupations deemed suitable for boys/men or girls/women [13,62–64], it becomes important that EC educators themselves do not carry such biases. All the more important findings are suitable training concepts.

5.4. Limitations

This study is not without limitations. Limitations concern the sample size, with 30 participants in the training group. It is advisable to expand the sample size in future investigations to enhance the statistical power. Additionally, conducting more comprehensive analyses, such as exploring correlations and regression coefficients between educators' attitudes and assessments, could yield valuable insights (but were not possible due to limited sample size). Further limitations pertain to the participants' evaluations, particularly given that the training program introduced a novel and innovative concept. It is possible that participants' assessments were influenced by a novelty effect. Furthermore, the study may be susceptible to limitations related to social desirability effects, although these limitations should not significantly impact the results pertaining to the training's effectiveness. The critical factors here are the differences observed in judgments between the training and control groups and any changes over time. The self-selection of participants can lead to a bias in results, as individuals are not randomly chosen but rather opt into the study voluntarily. This might result in differences in characteristics, behaviors, or attitudes between participants and nonparticipants [65].

6. Summary and Conclusions

Despite the generally understood importance of STEM skills and attitudes, studies point to obstacles hindering the development of STEM competencies and interests of children that begin in their early educational development. These obstacles affect both boys and girls, with girls being particularly impacted. While EC educators face an important, difficult task in supporting children in STEM, they are also affected by barriers to STEM. It is, therefore, all the more important that EC educators are able to develop positive attitudes and self-assessments as well as high skills in STEM didactics. Against this background, training was developed, and its impact on EC educators' self-concepts and stereotypes was evaluated. Overall, the results speak for the training as participants could increase their professional self-concept and decrease gender stereotypes.

The need for such a professional development program arises from a neglect of gender-sensitive pedagogical concepts for STEM in the Austrian vocational education curriculum for EC educators within the STEM domain. They do not receive sufficient pedagogical training during their education to effectively integrate STEM content into their teaching practices. Improvements are necessary in the vocational education of educators and for teachers responsible for the training of EC educators, including an assessment of their knowledge and pedagogical skills in STEM [66]. Altogether, there is a need to integrate this subject matter early in vocational education as well as to provide ongoing professional development opportunities for educators already established in their careers. Ideally, such training would be integrated into the curriculum and be continuous rather than an isolated, one-shot training to ensure a successful practical transfer in early childhood education. (see recommendations by Ari et al., [63]; Wei et al., [67]). The professional development program, characterized by features such as practical application of acquired knowledge in pedagogical contexts, facilitation of discussion, reflection, and the creation of a supportive learning community, can serve as a template for further refinement.

Conclusions also concern the implementation of the professional development program as blended learning with a mix of asynchronous and synchronous online phases and face-to-face workshops. This blended approach extends the reach of participation and temporal flexibility, accommodating educators in active employment residing at a distance. However, phases of instruction focusing on the manipulation of pedagogical materials, the application of didactic strategies, and immediate discussion and reflection necessitated face-to-face workshops. The effectiveness of this blended learning design in instigating positive changes in participants' attitudes and self-assessments suggests its potential as a model for future training.

Author Contributions: Conceptualization, S.F., M.P. and L.E.; methodology, S.F., S.H. and M.P.; formal analysis, S.F., S.H., M.P. and M.W.T.E.; investigation, S.F. and S.H.; resources, M.P., S.L. and L.E.; data curation, M.P.; writing—original draft preparation, S.F. and M.P.; writing—review and editing, S.F., S.H., M.P., S.L., M.W.T.E. and L.E.; supervision, M.P., S.L. and L.E.; project administration, S.F., S.H., M.P. and L.E.; funding acquisition, M.P., S.L. and L.E. All authors have read and agreed to the published version of the manuscript.

Funding: The project was funded by the Austrian Chancellery; Section III—Women's Affairs and Equality.

Institutional Review Board Statement: The study was performed in accordance with the American Psychological Association's Ethics Code and the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: The authors acknowledge the financial support of the Federal Chancellery for their support. The authors also acknowledge the Open Access Funding by the University of Graz.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. OECD. PISA 2024 Strategic Vision and Direction for Science. Available online: https://www.oecd.org/pisa/publications/PISA-2024-Science-Strategic-Vision-Proposal.pdf (accessed on 6 September 2023).
2. Campbell, C.; Speldewinde, C.; Howitt, C.J.; MacDonald, A. STEM practice in the early years. *Creat. Educ.* **2018**, *9*, 11–25. [CrossRef]
3. Patrick, H.; Mantzicopoulos, P.; Samarapungavan, A.; French, B.F. Patterns of young children's motivation for science and teacher-child relationship. *J. Exp. Educ.* **2008**, *76*, 121–144. [CrossRef]
4. Eshach, H.; Fried, M.N. Should science be taught in early childhood? *J. Sci. Educ. Technol.* **2005**, *14*, 315–336. [CrossRef]
5. Morgan, P.L.; Farkas, G.; Hillemeier, M.M.; Maczuga, S. Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educ. Res.* **2016**, *45*, 18–35. [CrossRef]
6. Watts, T.W.; Duncan, G.J.; Siegler, R.S.; Davis-Kean, P.E. What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educ. Res.* **2014**, *43*, 352–360. [CrossRef]
7. Saçkes, M.; Trundle, K.C.; Bell, R.L.; O'Connell, A.A. The influence of early science experience in kindergarten on children's immediate and later science achievement: Evidence from the early childhood longitudinal study. *J. Res. Sci. Teach.* **2011**, *48*, 217–235. [CrossRef]
8. Osborne, J. Attitude towards science: A review of the literature and its implications. *Int. J. Sci. Educ.* **2003**, *25*, 1049–1079. [CrossRef]
9. Dresel, M.; Schober, B.; Ziegler, A. Golem und "Pygmalion: Scheitert die Chancengleichheit von Mädchen im mathematisch-naturwissenschaftlich-technischen Bereich am geschlechtsstereotypen Denken der Eltern? In *Erwartungen in Himmelblau und Rosarot: Effekte, Determinanten und Konsequenzen von Geschlechterdifferenzen in der Schule*; Ludwig, H.P., Ludwig, H., Eds.; Juventa: München, Germany, 2007; pp. 61–81.
10. Master, A.; Meltzoff, A.N.; Cheryan, S. Gender stereotypes about interests start early and cause gender disparities in computer science and engineering. *Psychol. Cogn. Sci.* **2021**, *118*, e2100030118. [CrossRef]
11. Mustafa, N.A.; Shah, N.M.; Hashim, N.W.; Desa, M.M. An overview of stem education and industry 4.0 for early childhood education in Malaysia. *J. Posit. Sch. Psychol.* **2022**, *6*, 53–62.
12. Watt, H.M.; Eccles, J.S. *Gender and Occupational Outcomes: Longitudinal Assessments of Individual, Social, and Cultural Influences*; American Psychological Association: Washington, DC, USA, 2008; pp. 22–384.
13. Häfele, E. Was Ich Einmal Werden Möchte Was Ich Einmal Werden Möchte... Zukunftsvorstellungen bei Kindern [What I Want to Be When I Grow Up... Children's Ideas about the Future]. Available online: https://presse.vorarlberg.at/land/servlet/AttachmentServlet?action=show&id=24763 (accessed on 25 October 2023).
14. Muntoni, F.; Retelsdorf, J. At their children's expense: How parents' gender stereotypes affect their children's reading outcomes. *Learn. Instr.* **2019**, *60*, 95–103. [CrossRef]
15. Acker, S. *Gendered Education: Sociological Reflections on Women, Teaching and Feminism*; McGraw-Hill, Open University Press: London, UK, 1994.
16. Barrie, T. *Gender Play: Girls and Boys in School*; Rutgers University Press: New Brunswick, NJ, USA, 1993; ISBN 0813519233.
17. Clark, S.; Paechter, C. 'Why can't girls play football?' Gender dynamics and the playground. *Sport Educ. Soc.* **2007**, *12*, 261–276. [CrossRef]
18. Tiedemann, J. Teachers' Gender Stereotypes as Determinants of Teacher Perceptions in Elementary School Mathematics. *Educ. Stud. Math.* **2002**, *50*, 49–62. [CrossRef]
19. Marsh, H.W.; Martin, A.J. Academic self-concept and academic achievement: Relations and causal ordering. *Br. J. Educ. Psychol.* **2011**, *81*, 59–77. [CrossRef]
20. Dickhäuser, O.; Schöne, C.; Spinath, B.; Stiensmeier-Pelster, J. Die Skalen zum akademischen Selbstkonzept: Konstruktion und Überprüfung eines neuen Instrumentes. *Z. Differ. Diagn. Psychol.* **2002**, *23*, 393–405. [CrossRef]
21. Ertl, B.; Luttenberger, S.; Paechter, M. The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Front. Psychol.* **2017**, *8*, 703. [CrossRef] [PubMed]
22. Marsh, H.W.; Scalas, L.F. Self-concept in learning: Reciprocal effects model between academic self-concept and academic achievement. In *Social and Emotional Aspects of Learning*; Järvela, S., Ed.; Elsevier: Amsterdam, The Netherlands, 2011; pp. 191–197. [CrossRef]
23. Abele, A.E. Prädiktoren des Berufserfolgs von Lehrkräften. Befunde der Langzeitstudie MATHE. *Z. Pädagogik* **2011**, *57*, 674–694. [CrossRef]
24. Koch, S. Berufliches Selbstkonzept und eigenverantwortliche Leistung. *Gruppendynamik* **2005**, *36*, 157–174. [CrossRef]
25. Tschannen-Moran, M.; Hoy, A.W. Teacher efficacy: Capturing an elusive construct. *Teach. Teach. Educ.* **2001**, *17*, 783–805. [CrossRef]
26. Master, A.; Tang, D.; Forsythe, D.; Alexander, T.M.; Cheryan, S.; Meltzoff, A.N. Gender equity and motivational readiness for computational thinking in early childhood. *Early Child. Res. Q.* **2023**, *64*, 242–254. [CrossRef]
27. Bacher, J.; Beham-Rabanser, M.; Forstner, M. Can work value orientations explain the gender wage gap in Austria? *Int. J. Sociol.* **2022**, *52*, 208–228. [CrossRef]

28. Leitner, A. Gender Als Mainstream: Doing Gender in Theorie und Politischer Praxis. (Reihe Soziologie/Institut für Höhere Studien, Abt. Soziologie, 70). Institut für Höhere Studien (IHS). Available online: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-196889 (accessed on 1 October 2023).
29. Holland, J.L. *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments*, 3rd ed.; Psychological Assessment Resources: Odessa, Ukraine, 1997; ISBN 0-911907-27-0.
30. Cohrssen, C.; Page, J. Articulating a rights-based argument for mathematics teaching and learning in early childhood education. *Australas. J. Early Child.* **2016**, *41*, 104–108. [CrossRef]
31. Hedlin, M.; Gunnarsson, G. Preschool student teachers, technology, and gender: Positive expectations despite mixed experiences from their own school days. *Early Child Dev. Care* **2014**, *184*, 1948–1959. [CrossRef]
32. McClure, E.R.; Guernsey, L.; Clements, D.H.; Bales, S.N.; Nichols, J.; Kendall-Taylor, N.; Levine, M.H. STEM Starts Early: Grounding Science, Technology, Engineering, and Math Education in Early Childhood. Available online: https://joanganzcooneycenter.org/wp-content/uploads/2017/01/jgcc_stemstartsearly_final.pdf (accessed on 1 October 2023).
33. Gullberg, A.; Andersson, K.; Danielsson, A. Pre-Service Teachers' Views of the Child—Reproducing or Challenging Gender Stereotypes in Science in Preschool. *Res. Sci. Educ.* **2018**, *48*, 691–715. [CrossRef]
34. Spektor-Levy, O.; Baruch, Y.K.; Mevarech, Z. Science and Scientific Curiosity in Pre-school—The teacher's point of view. *Int. J. Sci. Educ.* **2013**, *35*, 2226–2253. [CrossRef]
35. Gerde, H.K.; Pierce, S.J.; Lee, K.; Van Egeren, L.A. Early childhood educators' self-efficacy in science, math, and literacy instruction and science practice in the classroom. *Early Educ. Dev.* **2018**, *29*, 70–90. [CrossRef]
36. Erden, F.; Sönmez, S. The effect of science teaching on Turkish preschool teachers' attitudes towards science teaching. *Early Child Dev. Care* **2011**, *181*, 285–300.
37. Edwards, K.; Loveridge, J. The inside story: Looking into early childhood teachers' support of children's scientific learning. *Australas. J. Early Child.* **2011**, *36*, 28–35. [CrossRef]
38. Çetin, M.; Demircan H, Ö.; Şenyurt, E.; Aktürk, A.A. An Analysis of young children's preferences on STEM Activities in terms of gender. *J. Educ. Future* **2020**, *18*, 1–15. [CrossRef]
39. Fleer, M. When preschool girls engineer: Future imaginings of being and becoming an engineer. *Learn. Cult. Soc. Interact.* **2019**, *30*, 100372. [CrossRef]
40. Stephenson, T.; Fleer, M.; Fragkiadaki, G. Increasing girls' STEM engagement in early childhood: Conditions created by the conceptual playworld model. *Res. Sci. Educ.* **2021**, *52*, 1243–1260. [CrossRef]
41. Wang, S. Exploring Early Childhood Educators' Perceptions and Practices Towards Gender Differences in STEM Play: A Multiple-Case Study in China. *Early Child. Educ. J.* **2023**, 1–14. [CrossRef] [PubMed]
42. Plöger-Werner, M. *Epistemologische Überzeugungen von Erzieherinnen und Erziehern: Die Bedeutung für das Pädagogische Handeln in Kindertageseinrichtungen*; Springer: Berlin/Heidelberg, Germany, 2015. [CrossRef]
43. Rank, A. Subjektive Theorien von Erzieherinnen zu vorschulischem Lernen und Schriftspracherwerb. In *Chancengleichheit in der Grundschule: Ursachen und Wege aus der Krise*; Ramseger, J., Wagener, M., Eds.; Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2008. [CrossRef]
44. Steffensky, M.; Anders, Y.; Barenthien, J.; Hardy, I.; Hartinger, A.; Kästner, R.; Leuchter, M.; Oppermann, E.; Pauen, S.; Rank, A.; et al. Wirkungen Naturwissenschaftlicher Bildungsangebote auf Pädagogische Fachkräfte und Kinder. Available online: https://www.pedocs.de/volltexte/2019/17871/pdf/Anders_et_al2019_Wirkungen_naturwissenschaftlicher_Bildungsangebote.pdf (accessed on 29 September 2023).
45. Roche, L.A.; Marsh, H.W. Multiple dimensions of university teacher self-concept. *Instr. Sci.* **2000**, *28*, 439–468. [CrossRef]
46. Brenneman, K.; Lange, A.; Nayfeld, I. Integrating STEM into preschool education; designing a professional development model in diverse settings. *Early Child. Educ. J.* **2019**, *47*, 15–28. [CrossRef]
47. Adams, N.E. Bloom's taxonomy of cognitive learning objectives. *J. Med. Libr. Assoc.* **2015**, *103*, 152–153. [CrossRef] [PubMed]
48. Anderson, L.W.; Krathwohl, D.R.; Airsasian, P.W.; Cruikshank, K.A.; Mayer, R.E.; Pintrich, P.R.; Raths, J.; Wittrock, M.C. A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Available online: https://www.uky.edu/~rsand1/china2018/texts/Anderson-Krathwohl%20-%20A%20taxonomy%20for%20learning%20teaching%20and%20assessing.pdf (accessed on 29 August 2023).
49. Koballa, T. Framework for the Affective Domain in Science Education. Available online: https://serc.carleton.edu/NAGTWorkshops/affective/framework.html (accessed on 1 October 2023).
50. Rzejak, D.; Lipowsky, F.; Bleck, V. Synopse bedeutsamer Merkmale von Lehrkräftfortbildungen. Wirkungsebenen und konzeptionelle Aspekte. *J. LehrerInnenbildung* **2020**, *20*, 18–30. [CrossRef]
51. Pölzl-Stefanec, E. Challenges and barriers to Austrian early childhood educators' participation in online professional development programmes. *Br. J. Educ. Technol.* **2021**, *52*, 2192–2208. [CrossRef]
52. Skipper, Y.; Leman, P.J. The role of feedback in young people's academic choices. *Int. J. Sci. Educ.* **2017**, *39*, 453–467. [CrossRef]
53. Stiftung Haus der Kleinen Forscher. Monitoring Bericht 2016/2017. Available online: https://www.stiftung-kinder-forschen.de/fileadmin/Redaktion/4_Ueber_Uns/Evaluation/Monitoring-Bericht_2016_2017.pdf (accessed on 29 August 2023).
54. Mösko, M. Elterliche Geschlechtsstereotype und Deren Einfluss auf das Mathematische Selbstkonzept von Grundschulkindern. Available online: https://www.fachportal-paedagogik.de/literatur/vollanzeige.html?FId=3302644 (accessed on 29 August 2023).

55. Grosch, K.; Häckl, S.; Kocher, M.G.; Bauer, C. MINT-Interesse Steigern bei Kindern. Ein Feldexperiment an Volksschulen in Österreich. Institut für Höhere Studien (IHS) Institut for Advanced Studies. 2020. Available online: https://irihs.ihs.ac.at/id/eprint/5558/1/ihs-report-2020-grosch-haeckl-kocher-bauer-mint-interesse-bei-kindern-steigern.pdf (accessed on 26 September 2023).
56. Blömeke, S.; Gustafsson, J.-E.; Shavelson, R.J. Beyond dichotomies—Competence viewed as a continuum. *Z. Psychol.* **2015**, *223*, 3–13. [CrossRef]
57. Chapman, R. A case study of gendered play in preschools: How early childhood educators' perceptions of gender influence children's play. *Early Child Dev. Care* **2016**, *186*, 1271–1284. [CrossRef]
58. Luttenberger, S.; Wimmer, S.; Paechter, M. Spotlight on math anxiety. *Psychol. Res. Behav. Manag.* **2018**, *11*, 311–322. [CrossRef]
59. Truitt, D.L. The Effect of Training and Development on Employee Attitude as it Relates to Training and Work Proficiency. *SAGE Open* **2011**, *1*. [CrossRef]
60. Marsh, H.W.; Trautwein, U.; Ludtke, O.; Koller, O.; Baumert, J. Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Dev.* **2005**, *76*, 397–416. [CrossRef]
61. Heyder, A.; Steinmayr, R.; Kessels, U. (Do teachers' beliefs about math aptitude and brilliance explain gender differences in children's math self-concept? *Front. Educ.* **2019**, *4*, 34. [CrossRef]
62. Kollmayer, M.; Schober, B.; Spiel, C. Gender stereotypes in education: Development, consequences, and interventions. *Eur. J. Dev. Psychol.* **2018**, *15*, 361–377. [CrossRef]
63. Ari, F.; Arslan-Ari, I.; Vasconcelos, L. Early Childhood Preservice Teachers' Perceptions of Computer Science, Gender Stereotypes, and Coding in Early Childhood Education. *Technol. Trends* **2022**, *66*, 539–546. [CrossRef] [PubMed]
64. Chick, K.A.; Heilman-Houser, R.A.; Hunter, M.W. The Impact of Child Care on Gender Role Development and Gender Stereotypes. *Early Child. Educ. J.* **2002**, *29*, 149–154. [CrossRef]
65. Heckman, J.J. Selection bias and self-selection. In *Econometrics*; Durlauf, S.N., Blume, L.E., Eds.; Palgrave Macmillan: London, UK, 1990; pp. 201–224.
66. Wrumnig, D. Masterarbeitsvorstellung. MINT in der Elementarpädagogik. Eine triangulative Studie zur Bedeutung von MINT in der elementarpädagogischen Praxis und den aktuell fachdidaktischen MINT-Kompetenzen frühpädagogischer Fachkräfte in Österreich. *ElFo Elem. Forschungsbeiträge* **2021**, *3*, 98–102.
67. Wei, F.-Y.F.; Hendrix, K.G. Gender differences in preschool children's recall of competitive and noncompetitive computer mathematics games. *Learn. Media Technol.* **2009**, *34*, 27–43. [CrossRef]

education sciences

Article

Measuring Implicit STEM and Math Attitudes in Adolescents Online with the Brief Implicit Association Test

Regina Reichardt [1,*], Celina Rottmann [1], Laura Russo [1], Kathrin J. Emmerdinger [2] and Sigrun Schirner [2]

1 Department of Psychology, University of Regensburg, 93040 Regensburg, Germany
2 Department of Educational Sciences, University of Regensburg, 93040 Regensburg, Germany; kathrin.emmerdinger@ur.de (K.J.E.); sigrun.schirner@ur.de (S.S.)
* Correspondence: regina.reichardt@ur.de

Abstract: Despite societal efforts toward enhancing gender equality, females are still underrepresented in STEM (science, technology, engineering, mathematics). Prominent explanations draw on gender differences in attitudes about STEM (with females holding more negative attitudes than males), which result from the gender stereotype that STEM is a male domain. While a lot of research has focused on explicit attitudes, little is known about implicit attitudes toward STEM. The present research sought to examine implicit attitudes among adolescents, and how they relate to other STEM cognitions. We measured implicit attitudes about the STEM concept as a whole, and about math in particular. For this purpose, we developed two Brief Implicit Associations Tests (BIATs) and administered them online in a sample of adolescents (N = 517). We additionally measured a variety of self-reported motivational and social-psychological variables (interest, aspiration, self-concept of ability, and sense of belonging to the math and STEM community, respectively), which previous research has identified as factors contributing to the gender gap in STEM participation. Our findings confirm the reliability and validity of both the STEM BIAT and the Math BIAT. Moreover, implicit STEM attitudes predicted interest in and aspiration for STEM, self-concept of STEM ability, and sense of belonging to the STEM community. Similarly, implicit math attitudes predicted interest in and aspiration for math, and sense of belonging to the math community (but not self-concept of math ability). Our findings confirm that our novel online BIATs are efficient measurement tools of implicit attitudes in adolescents. Moreover, our findings underscore the significance of implicit attitudes in the STEM domain.

Keywords: implicit attitudes; STEM; mathematics; gender stereotypes

Citation: Reichardt, R.; Rottmann, C.; Russo, L.; Emmerdinger, K.J.; Schirner, S. Measuring Implicit STEM and Math Attitudes in Adolescents Online with the Brief Implicit Association Test. *Educ. Sci.* **2023**, *13*, 899. https://doi.org/10.3390/educsci13090899

Academic Editors: Erin Mackenzie, Bernhard Ertl, Christine R. Starr, Silke Luttenberger and Manuela Paechter

Received: 6 June 2023
Revised: 2 September 2023
Accepted: 4 September 2023
Published: 6 September 2023

1. Introduction

Despite societal efforts toward enhancing gender equality, females are still underrepresented in the STEM domain, i.e., science, technology, engineering, and mathematics [1,2]. A tremendous amount of research has been devoted to understanding the factors determining the gender gap in STEM participation, identifying factors at the individual as well as the environmental level [3–6]. Importantly, gender gaps in STEM participation persist despite no or small gender gaps in STEM-related achievement [7–10]. Instead, cognitive, emotional, and motivational factors appear to play a major role. In particular, the stereotype that STEM is a male domain negatively affects females' attitudes about STEM, their self-concept of abilities in STEM, their STEM identity, and their sense of belonging to the STEM community, eventually influencing decisions to enter or leave STEM fields [6,11–16].

The present research focused on attitudes about STEM as an important factor contributing to the gender gap in STEM participation. We adopt the widespread definition of *attitude* as a global evaluation of an entity with some degree of favor or disfavor [17,18]. Attitudes are key determinants of motivation, decision-making, and behavior [19,20]. By and large, numerous studies demonstrate that females compared to males report more negative attitudes toward math and science [7,9,21–25]. Furthermore, math and science attitudes have been shown to be related to STEM choices and achievement [26–30].

Most research so far has focused on explicit or self-report measures of STEM attitudes and other cognitions. However, self-reports do not capture the entire spectrum of cognitions relevant to motivation and behavior. A tremendous amount of research in social psychology and beyond has shown that implicit attitudes and other implicit cognitions play an important role in motivation, behavior and decision-making (for a review, see [31]). At the same time, little is known about implicit STEM attitudes in adolescents who are at a stage in life where they are setting the course for their future professional careers. To fill this gap, the present research investigated implicit attitudes about STEM among adolescents and how they relate to other STEM cognitions.

1.1. Implicit Cognitions

In contrast to self-reported or explicit cognitions, implicit cognitions reflect automatic associations that are activated quickly and independently of goals [32,33]. Implicit cognitions are not easily accessible to introspection, and even if accessible, participants may not report them as such but adjust them based on their ideals, norms, and values. Both implicit and explicit cognitions contribute to motivation, behavior, and decision-making, yet under different conditions [20,34,35]. For instance, implicit attitudes about a political issue have been shown to better predict future decisions among undecided voters than explicit attitudes [36]. With respect to STEM, implicit gender stereotypes have been shown to predict women's commitment and fit in STEM [37]. In sum, a vast amount of research has shown the relevance of implicit cognitions in predicting a variety of outcomes [31,38]. Against this background, it seems reasonable that implicit cognitions contribute to the gender gap in STEM participation. Previous research on implicit STEM cognitions can be distinguished based on (1) the domain (math, science, STEM, etc.), (2) the type of association (stereotypes, attitudes, etc.), and (3) the age of the population (children, adolescents, adults).

To begin with, implicit stereotypes about math or science as male domains are prevalent around the world, yet with considerable variability between countries [12,39–41]. Implicit math gender stereotypes have been observed not only among adults but also among adolescents [42,43] (but see [44]), and even among elementary school children [45]. Similarly, implicit physics gender stereotypes have been observed among adults [46] and adolescents [47]. Among females, implicit math gender stereotypes predicted more negative implicit attitudes about math [41,48], a lower self-reported self-concept of abilities [41,43], and worse math achievement [41,43,49]. Furthermore, among females, implicit math gender stereotypes as well as implicit science gender stereotypes predicted lower interest or participation in the respective domain [14,41,42,50,51].

Most relevant to the present research are findings on implicit attitudes about STEM. However, we are not aware of a study on implicit attitudes about the concept of STEM as a whole. So far, research has investigated implicit attitudes about specific STEM-subjects, such as math or physics. In particular, implicit attitudes about math or physics were related to participant gender, with females exhibiting more negative implicit attitudes than males [41,47,48,52]. Gender differences in implicit attitudes about math have been observed already among elementary school children [52], and gender differences in implicit attitudes about physics have been observed among adolescents [47]. Implicit attitudes about math were related to interest and participation in math [41], self-concept of math abilities [41], and math achievement [41,48,52].

To summarize, implicit stereotypes and attitudes have been shown to play a pivotal role in STEM-related cognitions, motivation, and behavior. While there has been extensive research on implicit gender stereotypes about science or math as well as implicit attitudes about math or physics, we are not aware of a study on implicit attitudes about the concept of STEM as a whole. Moreover, most studies investigated implicit cognitions in adults, and the few studies that did examine implicit cognitions in children or adolescents focused on specific school subjects such as math or physics.

1.2. The Present Research

The main objective of the present research was to investigate implicit attitudes about STEM among secondary school students and their relation to other STEM cognitions. A better understanding of implicit attitudes about STEM in this population is essential for several reasons. First, previous research shows that interest in STEM rapidly declines throughout secondary school [53], making it especially important to learn more about implicit STEM attitudes and their contribution to STEM motivational factors in this age group. Second, the concept of STEM is omnipresent in education, and students are confronted with several STEM-related choices before and during secondary school in the German education system. Transition from elementary to secondary school already requires setting the course for future STEM-related educational options (i.e., selecting a school that offers a STEM profile). Depending on the school type, students at a higher grade level (e.g., seventh grade) can choose whether they want to intensify STEM subjects or other subjects (e.g., languages). Furthermore, students in secondary school can often choose to participate in extracurricular programs regarding STEM. Third, many intervention programs that aim at increasing female participation in STEM address the overarching concept of STEM, that is, they do not focus on single (school) subjects, but on various STEM domains and interdisciplinary aspects [54,55]. An instrument to capture implicit attitudes about STEM in school students would thus be a valuable evaluation tool for such programs. As STEM is an important concept in our society and education system, it is essential to better understand attitudes about this concept in school students and how they relate to other STEM cognitions.

Related to our main objective of investigating implicit STEM attitudes, our research had a further methodological goal. We sought to develop a novel measure of implicit STEM attitudes and evaluate whether this measure yields reliable and valid results when adolescents complete it online on a voluntary basis, at a location and time of their own choice. Previous research on implicit attitudes has mostly used the Implicit Association Test (IAT) [56]. In the IAT, participants rapidly classify stimuli belonging to four categories using two response keys. For instance, in a math-language-attitude IAT, participants classify stimuli belonging to the categories *math*, *language*, *positive*, and *negative*. They see one stimulus at a time on the computer screen (e.g., a math-related word such as “number”, a language-related word such as “word”, a positive word such as “love”, or a negative word such as “hate”). In one block, they use one response key (e.g., right key) for math and positive words, and the other response key (e.g., left key) for language and negative words. In the other block, they use one response key for language and positive words, and the other response key for math and negative words. The difference in the average response latencies of the two blocks represents the extent to which math or language is associated with positivity or negativity. For instance, faster responses in the math/positive–language/negative block than in the language/positive–math/negative block indicate positive associations with math relative to language.

Completion of the IAT is cognitively demanding and lengthy. As such, it is a suboptimal instrument when administered online on a voluntary basis in the age group of adolescents because it may result in large dropout rates. Previous research on implicit cognitions in underaged participants has been conducted in the lab and has used a simplified variant of the IAT with fewer exemplar stimuli to classify [47,52]. While reducing the number of stimuli is appropriate when the IAT assesses attitudes about a single subject such as math or physics, it is not suitable for measuring attitudes about a multifaceted construct such as STEM.

To develop a measurement instrument that adequately captures the multifaceted construct of STEM and is, at the same time, short and appealing for adolescents, we adapted the Brief Implicit Association Test (BIAT) [57]. The BIAT is a variant of the IAT that is easier and takes less time to complete than the IAT because the classification task is simplified, while, at the same time, the number of different exemplar stimuli is retained. Importantly, the BIAT has comparable psychometric properties in terms of reliability and validity as the IAT, as demonstrated in several adult samples [58]. However, the BIAT has not yet been used with underaged participants. Thus, it remains to be shown that the BIAT is a reliable and valid instrument in this age group.

To evaluate the BIAT as a measurement instrument in our target group, we implemented two BIATs that we aimed to compare, one BIAT assessing implicit attitudes about STEM and one BIAT assessing implicit attitudes about math. We know from previous research that implicit attitudes about math can be measured in underaged participants with a child-friendly IAT [52]. Therefore, we expect similar results with our novel Math BIAT. We will further investigate whether our novel STEM BIAT yields comparable results.

First, we aim to evaluate the psychometric properties of our Math and STEM BIATs. To evaluate the reliability of the BIATs we analyzed internal consistency. To evaluate the construct validity of the BIATs, we analyzed correlations with self-report measures of attitudes about math or STEM, respectively. Implicit and self-report measures of attitudes are thought to assess distinct, but related constructs [59]. According to previous research on the relation between implicit and explicit attitude measures, we should find small-to-medium-sized correlations [31].

Second, we examine the predictive validity of the BIATs by investigating their relations to other math- or STEM-related measures, respectively. Based on what has been observed with respect to implicit math attitudes [41,52], we expect implicit math attitudes to be related to interest in math, to aspiration in math, and to self-concept of math abilities. In a similar vein, we expect implicit STEM attitudes to be related to interest in STEM, to aspiration in STEM, and to self-concept of STEM abilities. We additionally examine whether implicit attitudes are related to sense of belonging to the respective community. Recent research suggests that gender stereotypes lead to a lower sense of belonging to the STEM community among females, which contributes to lower participation rates in females [6,60–65]. To our knowledge, the relation between sense of belonging and implicit attitudes has not yet been studied. Therefore, we seek to provide new evidence on a potential relation between these constructs. We expect that implicit math attitudes are related to sense of belonging to the math community. In a similar vein, implicit STEM attitudes should be related to sense of belonging to the STEM community.

2. Methods

The study was implemented and presented online using the PsyToolkit platform [66,67]. We report all manipulations, measures, and exclusions. Materials are available at https://osf.io/9sz32 (accessed on 26 May 2023).

2.1. Design

Participants completed either the Math BIAT or the STEM BIAT. They were randomly assigned to the BIAT conditions. Except for the BIAT, the conditions were identical, i.e., participants completed the same questionnaires in both conditions.

2.2. Participants

Participants were recruited on various online platforms, social media channels, and through personal contacts. The study was advertised as a study on attitudes about STEM. As compensation, participants could take part in a lottery of 50 vouchers for 10 EUR for an online shop of their choice. Informed consent was obtained from participants as well as their parents at the beginning of the study. The study was automatically terminated if consent was not given by either the participants or the parents.

N = 862 participants started the study after consent was provided. The participation criteria (1) aged between 10 and 17 years and (2) being a school student were assessed at the beginning of the study, and the study was terminated if participants did not meet the criteria. N = 844 participants met the participation criteria. N = 557 (66%) participants fully completed the study. Participants were excluded from the data analysis if they did not correctly answer the two attention check questions that were interspersed in the questionnaires (21 participants), if they indicated at the end of the study that they did not seriously answer the questions (five participants), if they were not yet in fifth grade (five participants), if there were technical problems (five participants), or if more than 10% of the responses in the BIAT were faster than 300 ms or slower than 10,000 ms (15 participants), which is the standard performance-based exclusion criterion in the BIAT [57,68]. We aimed for a total sample size of $N = 500$ ($N = 250$ per BIAT condition) to achieve sufficient power for correlational analyses [69]. The final sample size was $N = 517$ (262 participants completed the Math BIAT, and 255 participants completed the STEM BIAT). The demographics are shown in Tables 1 and 2.

Table 1. Numbers of participants.

Demographic Variables	Math BIAT Condition		STEM BIAT Condition	
	Male	Female	Male	Female
	104	**158**	**85**	**170**
School Type				
Comprehensive School	7	9	10	3
Lower-Track Sec. School [1]	0	2	1	0
Medium-Track Sec. School [1]	4	8	5	8
Higher-Track Sec. School [1]	89	136	66	149
Other	4	3	3	10
School Profile				
STEM profile	64	68	49	68
Language profile	19	33	17	41
Other or no profile	21	57	19	61
Subjects Taken				
Mathematics	104	158	85	170
Biology	100	146	81	159
Physics	93	126	66	135
Chemistry	85	115	60	117
Computer Science	73	108	54	112
German	104	158	85	170
English	104	158	82	168
French	48	86	34	87
Spanish	21	28	15	28
Latin	32	42	25	54

[1] Sec. = Secondary.

Table 2. Mean age and school grade level [1].

Demographic Variables	Math BIAT Condition		STEM BIAT Condition	
	Male	Female	Male	Female
Age	14.54 (2.02)	13.88 (2.17)	14.20 (2.20)	13.80 (2.11)
Grade level [2]	9.18 (2.16)	8.60 (2.11)	8.86 (2.21)	8.49 (2.13)

[1] Standard deviations are in parenthesis. [2] School grade level varied from 5th to 13th grade.

We examined whether demographic characteristics were distributed equally across the Math and STEM BIAT conditions. There was no significant association between condition and gender, $\chi^2(1) = 2.26$, $p = 0.133$. Participants did not differ between conditions in age, $t(515) = 1.10$, $p = 0.270$, $d = 0.10$ or grade level, $t(515) = 1.17$, $p = 0.245$, $d = 0.10$. There was no significant association between condition and school type (recoded as Higher-Track Secondary School vs. other), $\chi^2(1) = 0.25$, $p = 0.617$ or between condition and school profile (recoded as STEM profile vs. other/none), $\chi^2(1) = 1.05$, $p = 0.306$. In sum, the conditions did not differ in demographic variables.

Furthermore, we examined whether gender was associated with relevant demographic variables. Male participants were, on average, older than female participants, $t(515) = 2.82$, $p = 0.005$, $d = 0.26$. Correspondingly, male participants were, on average, in a higher grade level than female participants, $t(515) = 2.53$, $p = 0.012$, $d = 0.23$. There was no significant association between gender and school type (recoded as Higher-Track Secondary School vs. other), $\chi^2(1) = 2.25$, $p = 0.133$. However, there was a significant association between gender and school profile, $\chi^2(1) = 16.13$, $p < 0.001$. There were more male participants than expected and fewer female participants than expected in the STEM profile.

2.3. Procedure

Participants first answered demographic questions about their age, whether they were school students, the school type (response options were four different types of German secondary schools and other, see Table 1), the school profile if any (response options were language-oriented, STEM-oriented, and other), their grade level, and their gender. Then, participants were informed about the meaning of the acronym STEM (in German MINT) and that the STEM subjects in school were mathematics, computer science, biology, physics, and chemistry. They were told that some of the following questions would refer to STEM subjects, while other questions would refer to language subjects (e.g., German, Latin, English, French, Spanish). We selected the languages that are typically taught at secondary school in Germany. Participants were asked to indicate which of the following subjects they were currently taking or had taken before in school (mathematics, computer science, biology, physics, chemistry, German, Latin, English, French, Spanish).

Afterward, participants completed self-report measures on interest, aspiration, self-concept of abilities, attitudes, feeling thermometers, and feelings of belonging (for details see Section 2.4). Then, participants were randomly assigned to the BIAT condition (Math BIAT vs. STEM BIAT), with the constraint that gender was balanced across BIAT conditions.

After completion of the BIAT, participants could enter comments on the study, and they were asked to indicate whether they had seriously answered all question. Finally, participants were informed about the procedure of the lottery. To take part in the lottery, they were asked to enter a self-generated code that would be stored separately from their data to ensure anonymity of the study data and to send an E-Mail with the code and their name to the researcher.

2.4. Materials

2.4.1. Self-Report Measures

Interest in STEM and interest in languages were assessed with respect to the school subjects of mathematics, computer science, biology, physics, chemistry, German language, Latin, English, French, and Spanish [55,70]. A sample item is "Please indicate the extent to which you are interested in mathematics." Participants indicated their interest on a 6-point Likert scale from (*no interest at all*) to (*very strong interest*). For STEM-related analyses, responses to the STEM subjects were averaged to a STEM-interest scale ($\alpha = 0.71$), and responses to the language subjects were averaged to a language-interest scale ($\alpha = 0.62$).

Aspiration for STEM and aspiration for languages were assessed with respect to the subselection of school subjects that participants were currently taking (out of the school subjects mathematics, computer science, biology, physics, chemistry, German, Latin, English, French, and Spanish). For each of the subjects that participants were currently taking, they were asked "to specify the grade with which they would be satisfied in their next school report" [71]. The German grading system ranges from 1 (highest grade) to 6 (lowest grade). For STEM-related analyses, responses to the STEM subjects were averaged to a STEM-aspiration scale, and responses to the language subject were averaged to a language-aspiration scale. Because the subselection of school subjects, which comprised the STEM-aspiration and language-aspiration scales, respectively, differed between participants, internal consistency scores cannot be calculated across all participants.

Self-concept of abilities was assessed with respect to STEM abilities in general, as well as with respect to mathematics in particular. To this end, we used an adapted four-item scale version of the *belief in one's own abilities* scale [71,72]. The items were presented as 6-point bipolar Likert scale items with the poles labeled in item-specific ways ("I doubt that I am talented for the STEM subjects." vs. "I believe that I am talented for the STEM subjects."; "I am not sure whether I am good enough to succeed in the STEM subjects." vs. "I am sure that I am good enough to succeed in the STEM subjects."; "I don't have a lot of confidence in my STEM abilities." vs. "I have full confidence in my STEM abilities.", "When I get new learning material in the STEM subjects, I often think that I may not be able to understand it." vs. "When I get new learning material in the STEM subjects, I am usually able to understand it."). To assess the self-concept of mathematic abilities, we replaced the term *STEM subjects* with *mathematics.* The internal consistency was excellent for both the self-concept of STEM-ability scale (α = 0.92) and the self-concept of mathematics-ability scale (α = 0.94).

Attitudes were assessed with respect to STEM and languages in general, as well as with respect to mathematics and German in particular. To this end, we used an adapted three-item scale [48,73]. The items were presented as 6-point bipolar Likert scale items with the poles labeled in item-specific ways ("I don't favor the STEM subjects." vs. "I favor the STEM subjects."; "I don't like the STEM subjects at all." vs. "I like the STEM subjects a lot."; "The STEM subjects are absolutely boring." vs. "The STEM subjects are a lot of fun."). To assess attitudes toward the other domains, the term *STEM subjects* was replaced with *mathematics, language subjects*, and *German language*, respectively. Internal consistency was excellent for all attitude scales (STEM attitudes: α = 0.94; math attitudes: α = 0.96; language attitudes: α = 0.94; German attitudes: α = 0.95).

In addition to attitude scales, we administered feeling thermometers to assess the affective component of attitudes [48]. Feelings of unpleasantness/pleasantness were assessed with respect to the subselection of the school subjects, that participants were currently taking (out of the subjects mathematics, computer science, biology, physics, chemistry, German, Latin, English, French, and Spanish). Participants were asked to imagine that they were working on a task from each of these subjects. They should imagine the feelings they were experiencing while working on the task as vividly as possible. They gave their response on a slider scale ranging from *unpleasant* to *pleasant*. Responses were coded from 1 to 100. For STEM-related analyses, responses to the STEM subjects were averaged to a STEM-feeling thermometer scale, and responses to the language subjects were averaged to a language-feeling thermometer scale. Because the subselection of school subjects, which comprised the STEM-feeling thermometer and the language-feeling thermometer scales, respectively, differed between participants, internal consistency scores cannot be calculated across all participants.

Sense of belonging was assessed with respect to the STEM community and the math community. To this end, we adapted the Sense of Belonging Scale from Good et al. [62], following the German translation from Ladewig et al. [60]. We used a short four-item version ("I feel that I belong to the STEM people."; "I perceive myself as a member of the STEM community."; "I feel connected to the STEM people."; "I have the feeling that I am part of the STEM world."). Participants gave their responses on a 6-point Likert scale ranging from (*not at all*) to (*completely).* To assess sense of belonging to the math community, the term *STEM* was replaced with *mathematics.* Internal consistency was excellent for both scales (sense of belonging to the STEM community: α = 0.95; sense of belonging to the math community: α = 0.96).

2.4.2. Implicit Measures

Like the IAT, the BIAT is a speeded classification task, in which participants are presented with stimuli—one at a time on the screen—belonging to one of four categories. Different from the IAT, participants focus on just two of the four categories when classifying the stimuli, which makes the task easier. For instance, when *STEM* and *good* are the focal categories, participants use the right response key for all stimuli belonging to the categories *STEM* or *good*, and the left response key for all other stimuli. Conversely, when *languages* and *good* are the focal categories, participants use the right response key for all stimuli belonging to the categories *languages* or *good*, and the left response key for all other stimuli. Like in the IAT, however, stimuli of all four categories are presented during a block, and the mapping of stimuli to response keys is the same as in the IAT. That is, in one block participants respond with the right key to *STEM* and *good* stimuli and with the left key to *languages* and *bad* stimuli. In the other block, they respond with the right key to *languages* and *good* stimuli and with the left key to *STEM* and *bad* stimuli. The BIAT is easier to complete than the IAT because participants must keep in working memory only the two focal categories (not all four categories as in the IAT). Furthermore, the BIAT needs fewer practice trials than the IAT, shortening total completion time.

The STEM BIAT was modeled after science IATs [12]. The stimuli of the STEM BIAT were the STEM subjects that are typically taught at German secondary schools (see Table 3). We used *languages* as the comparison category because this category is easy for adolescents to understand and because secondary schools in Germany offer a language profile as an alternative to the STEM profile. The stimuli of the category *languages* were the language subjects that are typically taught at German secondary schools. The attribute dimensions were *good* and *bad* as in similar attitude IATs [47,52,73]. We used the standard procedure of the BIAT [57,68] and adapted the instructions and the practice block to make the task easier for our age group of adolescents. Participants were told that their task was to decide as quickly as possible whether a word presented in the center of the screen matched a category presented at the top of the screen. They were shown several example screens (e.g., Does "dog" match the category "animal"?). They were told to press the right-hand key if the word matched the category (focal category), and the left-hand key if the word did not match the category (nonfocal category). The response keys were L and A on a QWERTZ keyboard.

Table 3. Category labels and stimuli presented in the STEM and Math BIATs.

Category Labels (English Translation)	Stimuli (English Translation)	Category Labels (Original German)	Stimuli (Original German)
STEM	Mathematics Biology Chemistry Physics Computer Science	MINT	Mathematik Biologie Chemie Physik Informatik
Languages	German Latin English French Spanish	Sprachen	Deutsch Latein Englisch Französisch Spanisch
Mathematics	Numbers Compute Summate Multiply Geometry	Mathematik	Zahlen Rechnen Addieren Multiplizieren Geometrie
German	Words Verbs Read Orthography Poem	Deutsch	Wörter Verben Lesen Rechtschreibung Gedicht
Good	Happy Love Laughing Pleasure Wonderful	Gut	Glücklich Liebe Lachen Freude Wundervoll
Bad	Agony Nasty Awful Terrible Horrible	Schlecht	Qual Übel Schrecklich Grausam Scheußlich

Participants first completed a practice block with the concept category *animals* (e.g., dog, cat) and the attribute category *good words* as the focal categories and the concept category *trees* (e.g., oak, beech) and the attribute category *bad words* as nonfocal categories (see Table 4 for an overview of the procedure). The first four trials presented only stimuli of the concept categories *animals* and *trees*, and the following 16 trials alternated stimuli of the concept categories and stimuli of the attribute categories. After the practice block, the concept categories *STEM* and *languages* were introduced and the lists of *STEM*- and *languages*-stimuli were shown on the screen. Furthermore, the lists of *good* and *bad* words were shown. Participants completed four test blocks. Following the recommendation of Nosek et al. [68], the attribute category *good* was always focal, and the attribute category *bad* was always nonfocal. Which of the two concept categories of *STEM* and *languages* was focal alternated between test blocks. It was counterbalanced between participants whether they started with the *STEM*-*good*-focal block or with the *languages*-*good*-focal block. The first four trials of each block presented only stimuli of the concept categories and were not analyzed. The following 20 trials of each block alternated stimuli of the concept categories and stimuli of the attribute categories. The order of the stimuli was determined randomly.

Table 4. BIAT procedure [1].

Block Type	N° Test Block	Trials	Nonfocal Categories (Left Key)	Focal Categories (Right Key)
Practice Block		4 trials with concepts only 16 trials alternating concepts and attributes	Tree Bad	Animal Good
Test Block STEM-Good	1	4 trials with concepts only 20 trials alternating concepts and attributes	Languages Bad	STEM Good
Test Block Languages-Good	2	4 trials with concepts only 20 trials alternating concepts and attributes	STEM Bad	Languages Good
Test Block STEM-Good	3	4 trials with concepts only 20 trials alternating concepts and attributes	Languages Bad	STEM Good
Test Block Languages-Good	4	4 trials with concepts only 20 trials alternating concepts and attributes	STEM Bad	Languages Good

[1] The table presents the procedure of the STEM BIAT. The procedure of the Math BIAT was identical. In each BIAT, blocks 1 and 3 are identical, and blocks 2 and 4 are identical. The order of blocks 1 and 3 with blocks 2 and 4 was counterbalanced. From the trials of blocks 1 through 4, only the 20 trials alternating concepts and attributes are analyzed.

During all trials of a block, the focal category labels were shown on the top of the screen (e.g., "STEM or Good"), and the response keys were shown on the bottom of the screen, with the response key for nonfocal stimuli in brackets ("L-key [A-key]"). The response keys were presented as a reminder to reduce potential error variance stemming from careless reading of the instructions. The concept category labels and stimuli were presented in green. The attribute labels and stimuli were presented in yellow. The background color was black. On each trial, a stimulus was shown in the center of the screen until participants responded. In case of an incorrect response, a red "X" appeared below the stimulus until participants gave the correct response. The intertrial interval was 400 ms.

The stimuli of the Math BIAT were taken from previous research using Math IATs [41,42,48]. As our sample was German speaking, we used *German language* as the comparison category (see Table 3). Otherwise, the Math BIAT was identical to the STEM BIAT.

3. Results

3.1. Analyses of the BIATs

We used the scoring algorithm recommended by Nosek et al. [68]. Responses from practice trials and the first four trials of each block were deleted. The dependent variable was the latency from stimulus onset to the correct response. Recall that when participants made an error, they had to correct their response. On these trials, the total latency from stimulus onset to the final correct response was used as dependent variable. Responses with latencies slower than 10,000 ms were deleted. Latencies faster than 400 ms were recoded to 400 ms, and latencies slower than 2000 ms were recoded to 2000 ms. Separate *D* scores were computed for the first two blocks and the second two blocks, and then averaged. To compute the *D* score for the first two blocks of the STEM BIAT, we subtracted the mean response latencies in the STEM-good block from the mean response latencies in the languages-good block, and divided the resulting difference score by the standard deviation of response latencies across both blocks. The *D* score is an individual effect size estimate that is similar to Cohen's *d*. A positive-*D* score indicates a preference for STEM relative to languages, and a negative-*D* score indicates a preference for languages relative to STEM. The Math-BIAT score was calculated in the same vein. A positive-*D* score, thus, indicates a preference for math relative to German, and a negative-*D* score indicates a preference for German relative to math.

3.1.1. BIAT Completion Time

Participants took on average $M = 5.76$ ($SD = 1.68$) minutes to complete the Math BIAT, and $M = 5.89$ ($SD = 1.73$) minutes to complete the STEM BIAT, with no significant difference between BIAT conditions, $t(515) = -0.87$, $p = 0.384$. Older participants took less time than younger participants, as indicated by a significant negative correlation between age and completion time, $r = -0.32$, $p < 0.001$.

3.1.2. BIAT Internal Consistency

As an index of internal consistency, we calculated the Guttman Split-Half coefficient from the *D* scores of the first two and the second two blocks. The internal consistency was excellent for both, the Math BIAT ($\alpha = 0.97$) and the STEM BIAT ($\alpha = 0.95$).

3.1.3. BIAT Construct Validity

To examine the construct validity of the BIATs, we calculated correlations of the *D* scores with explicit attitudes and feeling thermometer scales about math and STEM, respectively. Correlations with self-report measures about German and languages are presented in the Supplementary Materials (Table S1). As expected, we observed a significant positive correlation between implicit math attitudes and explicit math attitudes, $r = 0.14$, $p = 0.020$, as well as between implicit math attitudes and math feelings, $r = 0.16$, $p = 0.011$. Similarly, we observed a significant positive correlation between implicit STEM attitudes and explicit STEM attitudes, $r = 0.17$, $p = 0.005$, as well as between implicit STEM attitudes and STEM feelings, $r = 0.16$, $p = 0.009$.

3.2. Gender Differences in Math- and STEM-Related Measures

We examined gender differences in all math- and STEM-related measures with one-tailed *t*-Tests. Gender differences in German- and languages-related measures are presented in the Supplementary Materials (Table S2). As can be seen in Table 5, we observed the well-known gender differences in almost all math and STEM-related measures. Regarding implicit and explicit attitudes, an interesting pattern emerged. The genders differed in their implicit attitudes about math, with girls showing more negative implicit attitudes about math than boys, $t(260) = 2.68$, $p = 0.004$. However, the genders did not differ in their explicit attitudes about math, $t(515) = 1.52$, $p = 0.064$. Conversely, the genders differed in their explicit attitudes about STEM, with girls showing less positive explicit attitudes about STEM than boys, $t(515) = 2.74$, $p = 0.003$, but not in their implicit attitudes about STEM, $t(253) = 0.98$, $p = 0.165$. When asked about their feelings, participants showed significant gender differences. Girls reported less positive feelings than boys about math, $t(418.7) = 2.87$, $p = 0.002$ (unequal variances assumed) and about STEM, $t(515) = 3.34$, $p < 0.001$.

Table 5. Means, standard deviations, and Cohen's *d* of gender differences in Math- and STEM-related measures.

Measure	Math			STEM		
	Male	Female	Gender Difference	Male	Female	Gender Difference
	M (*SD*)	*M* (*SD*)	*d*	*M* (*SD*)	*M* (*SD*)	*d*
Implicit attitudes	0.07 (0.56)	−0.11 (0.53)	0.34 **	0.02 (0.51)	−0.05 (0.59)	0.13
Explicit attitudes	4.59 (1.37)	4.40 (1.42)	0.14	4.87 (1.14)	4.58 (1.16)	0.25 **
Feeling thermometer	73.81 (25.70)	66.87 (27.88)	0.26 **	69.79 (17.46)	64.14 (19.05)	0.31 ***
Interest	4.66 (1.32)	4.43 (1.40)	0.17 *	4.44 (0.98)	4.20 (1.01)	0.25 **
Aspiration	2.25 (0.97)	2.27 (0.89)	−0.03	2.26 (0.74)	2.24 (0.75)	0.04
Self-concept of ability	4.80 (1.20)	4.27 (1.46)	0.38 ***	4.78 (1.02)	4.20 (1.33)	0.48 ***
Sense of belonging	3.88 (1.44)	3.50 (1.46)	0.26 **	4.17 (1.30)	3.57 (1.35)	0.45 ***

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$ (one-tailed *t*-tests).

Furthermore, the genders differed by interest. Girls reported a lower interest in math, $t(515) = 1.83$, $p = 0.034$ and a lower interest in STEM, $t(515) = 2.70$, $p = 0.004$ compared to boys. Interestingly, males and females did not differ in math aspiration, $t(515) = -0.31$, $p = 0.379$ nor in STEM aspiration, $t(515) = 0.38$, $p = 0.351$. The genders did, however, differ in self-concept of abilities. Girls compared to boys reported lower self-concepts of math ability, $t(455.1) = 4.40$, $p < 0.001$ (unequal variances assumed) and lower self-concepts of STEM ability, $t(475.1) = 5.58$, $p < 0.001$ (unequal variances assumed). Finally, the genders differed in sense of belonging. Girls compared to boys reported a lower sense of belonging to the math community, $t(515) = 2.83$, $p = 0.002$ and a lower sense of belonging to the STEM community, $t(515) = 4.95$, $p < 0.001$.

3.3. Predictive Validity of Implicit Attitudes

We examined whether implicit attitudes predicted interest, aspiration, self-concept of ability, and sense of belonging separately for the math and STEM domains. To this end, we first analyzed zero-order correlations (Tables 6 and 7). All math-related self-report measures were highly correlated with one another. Implicit math attitudes were related to all math-related self-report measures except for self-concept of math ability. In a similar vein, all STEM-related self-report measures were highly correlated with one another. Implicit STEM attitudes were related to all STEM-related self-report measures.

To examine the predictive validity of implicit attitudes, we conducted separate linear regression analyses on the various STEM- and math-related outcome variables. In the first step of each regression analyses, we entered the demographic variables gender, grade level, and school profile (dummy coded as STEM profile vs. not) as control variables. In the second step, we entered implicit attitudes as predictors. Dependent variables were STEM- and math-related interest, aspiration, self-concept of ability, and sense of belonging. The results are summarized in Table 8 (math-related regression analyses) and Table 9 (STEM-related regression analyses).

As can be seen in Table 8, a STEM school profile (but not gender and grade level) significantly predicted math interest. Students attending a school with a STEM profile reported higher math interest than other students. Most importantly, implicit math attitudes significantly predicted math interest over and beyond the control variables. Students with more positive implicit math attitudes reported a higher math interest than students with more negative implicit math attitudes.

Table 6. Correlations among Math-related measures [1].

Measure	1	2	3	4	5	6	7
1. Implicit math attitude	-	0.14 *	0.16 *	0.16 *	−0.14 *	0.12	0.16 *
2. Explicit math attitude		-	0.86 ***	0.87 ***	−0.61 ***	0.79 ***	0.80 ***
3. Math feeling thermometer			-	0.80 ***	−0.61 ***	0.79 ***	0.76 ***
4. Math interest				-	−0.63 ***	0.71 ***	0.75 ***
5. Math aspiration [2]					-	−0.59 ***	−0.55 ***
6. Self-concept of math ability						-	0.77 ***
7. Sense of belonging to math community							-

[1] Correlations with implicit attitudes are based on the sample in the Math-BIAT condition (N = 262), correlations among self-report measures are based on the entire sample (N = 517). [2] Math aspiration was measured in terms of aspired grades, ranging from 1 (highest grade) to 6 (lowest grade). * $p < 0.05$. *** $p < 0.001$.

Table 7. Correlations among STEM-related measures [1].

Measure	1	2	3	4	5	6	7
1. Implicit STEM attitude	-	0.17 **	0.16 **	0.21 ***	−0.22 ***	0.22 ***	0.21 ***
2. Explicit STEM attitude		-	0.71 ***	0.79 ***	−0.49 ***	0.73 ***	0.76 ***
3. STEM feeling thermometer			-	0.70 ***	−0.51 ***	0.64 ***	0.63 ***
4. STEM interest				-	−0.48 ***	0.65 ***	0.70 ***
5. STEM aspiration [2]					-	−0.49 ***	−0.44 ***
6. Self-concept of STEM ability						-	0.72 ***
7. Sense of belonging to STEM community							-

[1] Correlations with implicit attitudes are based on the sample in the STEM BIAT condition (N = 255), correlations among self-report measures are based on the entire sample (N = 517). [2] STEM aspiration was measured in terms of aspired grades, ranging from 1 (highest grade) to 6 (lowest grade). ** $p < 0.01$. *** $p < 0.001$.

Furthermore, a STEM school profile (but not gender and grade level) significantly predicted math aspiration. Students attending a school with a STEM profile reported higher math aspirations than other students. Most importantly, implicit math attitudes significantly predicted math aspiration over and beyond the control variables. Students with more positive implicit math attitudes reported higher math aspirations than students with more negative implicit math attitudes.

Furthermore, gender, grade level, and a STEM school profile significantly predicted self-concept of math ability. Boys reported a higher self-concept of math ability than girls. Students at a higher grade level reported a lower self-concept of math ability than students at a lower grade level. Students attending a school with a STEM profile reported a higher self-concept of math ability than other students. However, implicit math attitudes did not predict self-concept of math ability over and beyond the control variables.

Finally, grade level and a STEM school profile (but not gender) significantly predicted sense of belonging to the math community. Students at a higher grade level reported a lower sense of belonging to the math community than students at a lower grade level. Students attending a school with a STEM profile reported a higher sense of belonging to the math community than other students. Most importantly, implicit math attitudes significantly predicted sense of belonging over and beyond the control variables. Students with more positive implicit math attitudes reported a higher sense of belonging to the math community than students with more negative implicit math attitudes.

As can be seen in Table 9, grade level and a STEM school profile (but not gender) significantly predicted STEM interest. Students at a higher grade level reported lower STEM interest than students at a lower grade level. Students attending a school with a STEM profile reported higher STEM interest than other students. Most importantly, implicit STEM attitudes significantly predicted STEM interest over and beyond the control variables. Students with more positive implicit STEM attitudes reported a higher STEM interest than students with more negative implicit STEM attitudes.

Table 8. Regression analyses on math variables [1].

Variable	Model 1			Model 2		
	B **[95% CI]**	***SE B***	**β**	***B*** **[95% CI]**	***SE B***	**β**
			Math interest			
Gender [a]	0.16 [−0.17; 0.49]	0.17	0.06	0.10 [−0.23; 0.43]	0.17	0.04
Grade level	<0.01 [−0.08; 0.08]	0.04	<0.01	<0.01 [−0.08; 0.08]	0.04	<0.01
STEM school profile [b]	0.55 ** [0.22; 0.88]	0.17	0.21	0.55 ** [0.22; 0.88]	0.17	0.21
Implicit math attitudes				0.35 * [0.06; 0.64]	0.15	0.14
R^2	0.052 **			0.072 ***		
ΔR^2				0.020 *		
			Math aspiration			
Gender [a]	−0.03 [−0.25; 0.20]	0.11	−0.02	0.01 [−0.21; 0.24]	0.11	0.01
Grade level	−0.02 [−0.08; 0.03]	0.03	−0.06	−0.02 [−0.08; 0.03]	0.03	−0.06
STEM school profile [b]	−0.30 ** [−0.52; −0.08]	0.11	−0.17	−0.30 ** [−0.52; −0.08]	0.11	−0.17
Implicit math attitudes				−0.22 * [−0.42; −0.03]	0.10	−0.14
R^2	0.039 *			0.057 **		
ΔR^2				0.018 *		
			Self-concept of math ability			
Gender [a]	0.47 ** [0.14; 0.79]	0.16	0.17	0.43 * [0.10; 0.75]	0.17	0.16
Grade level	−0.08 * [−0.15; −0.004]	0.04	−0.13	−0.08 * [−0.16; −0.01]	0.04	−0.13
STEM school profile [b]	0.48 ** [0.16 0.81]	0.16	0.18	0.48 ** [0.16 0.80]	0.16	0.18
Implicit math attitudes				0.22 [−0.07; 0.50]	0.15	0.09
R^2	0.074 ***			0.082 ***		
ΔR^2				0.008		
			Sense of belonging to math community			
Gender [a]	0.30 [−0.04; 0.65]	0.18	0.11	0.24 [−0.11; 0.59]	0.18	0.08
Grade level	−0.09 * [−0.17; −0.01]	0.04	−0.14	−0.09 * [−0.17; −0.01]	0.04	−0.14
STEM school profile [b]	0.70 *** [0.35; 1.05]	0.18	0.25	0.70 *** [0.35; 1.04]	0.18	0.25
Implicit math attitudes				0.36 * [0.05; 0.66]	0.16	0.14
R^2	0.079 ***			0.097 ***		
ΔR^2				0.019 *		

[1] In Model 1, we entered the control variables gender, grade level, and STEM school profile to predict the outcome variables. In Model 2, we entered implicit math attitudes as predictor. [a] 0 = female, 1 = male. [b] 0 = no STEM profile, 1 = STEM profile. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Table 9. Regression analyses on STEM variables [1].

Variable	Model 1			Model 2		
	B [95% CI]	*SE B*	β	*B* [95% CI]	*SE B*	β
STEM interest						
Gender [a]	0.14 [−0.13; 0.40]	0.14	0.06	0.12 [−0.15; 0.38]	0.13	0.05
Grade level	−0.12 *** [−0.18; −0.06]	0.03	−0.25	−0.12 *** [−0.18; −0.06]	0.03	−0.25
STEM school profile [b]	0.41 ** [0.16; 0.67]	0.13	0.20	0.37 ** [0.12; 0.62]	0.13	0.18
Implicit STEM attitudes				0.36 ** [0.15; 0.58]	0.11	0.20
R^2	0.088 ***			0.126 ***		
ΔR^2				0.038 **		
STEM aspiration						
Gender [a]	0.11 [−0.09; 0.32]	0.10	0.07	0.13 [−0.07; 0.33]	0.10	0.08
Grade level	−0.02 [−0.06; 0.03]	0.02	−0.05	−0.02 [−0.06; 0.03]	0.02	−0.05
STEM school profile [b]	−0.2 ** [−0.44; −0.05]	0.10	−0.16	−0.21 * [−0.41; −0.02]	0.10	−0.14
Implicit STEM attitudes				−0.29 *** [−0.46; −0.13]	0.09	−0.21
R^2	0.030			0.075 ***		
ΔR^2				0.045 ***		
Self-concept of STEM ability						
Gender [a]	0.48 ** [0.14; 0.82]	0.17	0.17	0.46 ** [0.12; 0.79]	0.17	0.16
Grade level	−0.05 [−0.12; 0.03]	0.04	−0.07	−0.05 [−0.12; 0.03]	0.04	−0.07
STEM school profile [b]	0.53 ** [0.20; 0.85]	0.17	0.20	0.47 ** [0.15; 0.80]	0.16	0.18
Implicit STEM attitudes				0.46 ** [0.18; 0.74]	0.14	0.19
R^2	0.078 ***			0.115 ***		
ΔR^2				0.037 **		
Sense of belonging to STEM community						
Gender [a]	0.52 ** [0.17; 0.87]	0.18	0.18	0.49 ** [0.15; 0.84]	0.18	0.17
Grade level	−0.09 * [−0.16; −0.01]	0.04	−0.13	−0.09 * [−0.16; −0.01]	0.04	−0.13
STEM school profile [b]	0.72 *** [0.38; 1.05]	0.17	.26	0.67 *** [0.33; 1.00]	0.17	0.24
Implicit STEM attitudes				0.44 ** [0.15; 0.72]	0.15	0.18
R^2	.114 ***			0.145 ***		
ΔR^2				0.031 **		

[1] In Model 1, we entered the control variables gender, grade level, and STEM school profile to predict the outcome variables. In Model 2, we entered implicit STEM attitudes as predictor. [a] 0 = female, 1 = male. [b] 0 = no STEM profile, 1 = STEM profile. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Furthermore, a STEM school profile (but not gender and grade level) significantly predicted STEM aspiration. Students attending a school with a STEM profile reported higher STEM aspirations than other students. Most importantly, implicit STEM attitudes significantly predicted STEM aspiration over and beyond the control variables. Students with more positive implicit STEM attitudes reported higher STEM aspirations than students with more negative implicit STEM attitudes.

Furthermore, gender and a STEM school profile (but not grade level) significantly predicted self-concept of STEM ability. Boys reported a higher self-concept of STEM ability than girls. Students attending a school with a STEM profile reported a higher self-concept of STEM ability than other students. Most importantly, implicit STEM attitudes significantly predicted self-concept of STEM ability over and beyond the control variables. Students with more positive implicit STEM attitudes reported a higher self-concept of STEM ability than students with more negative implicit STEM attitudes.

Finally, gender, grade level, and a STEM school profile significantly predicted sense of belonging to the STEM community. Boys reported a higher sense of belonging to the STEM community than girls. Students at a higher grade level reported a lower sense of belonging to the STEM community than students at a lower grade level. Students attending a school with a STEM profile reported a higher sense of belonging to the STEM community than other students. Most importantly, implicit STEM attitudes significantly predicted sense of belonging over and beyond the control variables. Students with more positive implicit STEM attitudes reported a higher sense of belonging to the STEM community than students with more negative implicit STEM attitudes.

4. Discussion

The goals of the present study were twofold. First, we sought to evaluate the reliability and validity of our novel implicit attitude measures in a sample of adolescents participating online on a voluntary basis, at a location and time of their own choice. Second, we sought to provide first evidence on implicit STEM attitudes in adolescents and extend our knowledge on implicit math attitudes and how they relate to other STEM and math cognitions.

Regarding our first goal, the present study confirms the reliability and validity of both the Math BIAT and the STEM BIAT in our adolescent sample under the conditions of self-determined online participation. The internal consistency of both BIATs was excellent and slightly better than the internal consistencies of the BIAT and the IAT reported in previous research [58,68]. Criterion validity was evaluated based on correlations with self-report measures of attitudes. To this end, we administered explicit attitude scales consisting of three items as well as feeling thermometer items regarding the school subjects [48]. As implicit and self-report measures are thought to assess distinct, but related constructs [59], we expected small-to-medium-sized correlations. Similar to previous research on implicit–explicit math attitude correlations in underaged participants [52], we observed small-sized correlations of implicit and explicit math attitudes as well as of implicit and explicit STEM attitudes. This pattern of results was confirmed across both self-report measures, the explicit attitude scale and the feeling thermometer items, with the latter tapping into the affective component of attitudes. Taken together, the present study provides evidence that implicit attitudes about math and STEM can be measured reliably and validly in adolescents using our adaptation of the BIAT. As such, the BIAT constitutes a useful alternative to the standard IAT, possessing several advantages (shorter completion time and less cognitively taxing than the standard IAT), without compromising reliability and validity.

Regarding our second goal, the present study reveals several interesting results. We replicate previous findings that females have more negative implicit attitudes about math than males [41,52]. At the same time, the genders did not differ in their explicit attitudes about math. This pattern underscores the added value of implicit measures when investigating attitudes about math. Interestingly, we did not observe gender differences in implicit attitudes about STEM. This is particularly noteworthy because we observed gender differences in explicit attitudes about STEM, with boys reporting more positive attitudes and more positive feelings about STEM. At present, it is difficult to conclusively interpret the observed pattern of gender differences. It is important to note that we did not draw a representative sample and, therefore, our results on gender differences must be interpreted with caution. Our self-selected sample may likely be biased toward STEM-interested adolescents because we advertised the study as a study on STEM. Consequently, the observed gender difference may underestimate the actual gender differences in the population.

Most importantly, our results show that both implicit attitudes about math and implicit attitudes about STEM predicted a variety of other math- and STEM-related cognitions, respectively. In particular, implicit attitudes about math as well as implicit attitudes about STEM predicted interest in and aspiration for math and STEM, respectively. Furthermore, implicit STEM (but not math) attitudes predicted self-concept of ability. Finally, implicit attitudes about math as well as implicit attitudes about STEM predicted sense of belonging to the math and STEM communities, respectively. These findings are noteworthy, because they demonstrate that implicit attitudes contribute to several motivational factors, such as interest, aspiration, and self-concept of ability, which eventually play a role in career decisions. Moreover, the nearly parallel pattern of relations of implicit math attitudes and self-reported math cognitions on one side and implicit STEM attitudes and self-reported STEM cognitions on the other side confirms the generality of the findings. Finally, this is, to our knowledge, the first evidence showing that implicit attitudes predict sense of belonging to the respective community. Sense of belonging has recently been identified as one major factor contributing to gender gaps in STEM participation [6]. Our findings add to this literature by showing that implicit attitudes are related to a sense of belonging with respect to both math and STEM.

Furthermore, the present research points to the idea that intervention programs aiming at increasing female participation rates in STEM should focus not only on changing explicit cognitions but also on changing implicit cognitions [73–75]. Decades of social-psychological research on the change of implicit attitudes demonstrate that implicit attitudes can be changed in the short term by a variety of interventions [76]. However, long-term change is difficult to maintain [77]. At the same time, implicit attitudes play an important role in behavior, decision-making, and motivation [31]. Thus, it is essential to investigate how long-term change of implicit attitudes about STEM can be achieved. Instilling and maintaining positive association with STEM in girls may constitute one route to eventually increasing the rate of females entering and staying in STEM.

In interpreting our findings, it must be kept in mind that one characteristic of IAT measures is that they are inherently relative in nature. That is, they measure associations with one category relative to a contrast category. When drawing conclusions from IAT findings, one must therefore keep in mind that the choice of the contrast category contributes to the final IAT score. In our case, we selected German as the contrast category for the Math BIAT and languages as the contrast category for the STEM BIAT. Participants' national language has been the standard contrast category in Math IATs [41,52]. This choice is based on the fact that the national language and math are school subjects students typically take from first grade on. Moreover, gender stereotypes with respect to math and reading develop early in childhood [45]. Thus, the national language is an obvious contrast category to math.

Our novel STEM BIAT was modeled after science IATs that contrast science with liberal arts or humanities [12]. While the concepts of liberal arts or humanities are known to university students, school students are not yet familiar with these concepts. Therefore, we chose languages as the contrast category. The category of languages is easy to understand, and it is obvious which exemplar stimuli belong to this category. Furthermore, the category of languages appears to be an appropriate contrast category to STEM because the higher-track secondary schools in Germany typically offer a STEM or a language profile. Therefore, school students need to reflect on their preferences, interests, and abilities regarding STEM and languages to make an informed decision about which profile they want to choose. While we and many other researchers investigated STEM as a unitary concept, it is also worthwhile to differentiate between STEM subjects that are more or less gender balanced [78]. Whereas physics and computer science still have a low proportion of women, biology and chemistry have achieved almost equal representation of genders, at least among university students [2]. Research has discussed several factors contributing to the different gender distributions in the various STEM subjects, including stereotypes about the particular subjects [47,61,64,79,80]. Thus, future research may disentangle implicit attitudes about the STEM subfields that are more or less gender-balanced [81]. Nevertheless, measuring implicit attitudes about STEM as a unitary concept is a useful and important endeavor because STEM is presented as a unitary concept in education systems, and many secondary schools offer a STEM or a language profile (among other profiles), which requires students to reflect on their STEM interests and abilities.

5. Conclusions

In sum, the present research demonstrates for the first time that implicit attitudes about math and STEM can be measured reliably and validly in an online sample of adolescents. As such, our online BIATs provide efficient measurement tools for future research. Our findings demonstrate that implicit attitudes are related to motivational and social-psychological factors (interest, aspiration, self-concept of abilities, and sense of belonging), which are known factors contributing to gender gaps in STEM participation. Thus, our findings underscore the significance of implicit attitudes in the STEM domain.

Future research may further investigate the role of implicit attitudes in behavior, decision making, and emotional experiences. Dual-process models [34,35] and models of implicit cognition [31–33] provide a solid theoretical basis on which several predictions can be derived. For instance, it would be interesting to examine the predictive validity of implicit attitudes with respect to classroom behavior such as asking questions or approaching the teacher after class [73]. Furthermore, it would be interesting to investigate the predictive validity of implicit attitudes with respect to future career decisions such as entering or dropping out of the STEM domain. Also, it would be interesting to investigate the relation of implicit attitudes to well-being and stress experiences when working on STEM tasks. Taken together, implicit attitudes may play a significant role in a variety of STEM-related variables.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci13090899/s1, Table S1: Correlations among German- and languages-related measures; Table S2: Means, standard deviations, and Cohen's *d* of gender differences in German- and languages-related measures.

Author Contributions: Conceptualization, R.R., K.J.E. and S.S.; methodology, R.R., C.R. and L.R.; software, R.R., C.R. and L.R.; formal analysis, R.R.; investigation, C.R. and L.R.; data curation, R.R.; writing—original draft preparation, R.R.; writing—review and editing, R.R., C.R., L.R., K.J.E. and S.S.; supervision, R.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of the University of Regensburg (protocol code 21-5659-101, date of approval 17 January 2021).

Informed Consent Statement: Informed consent was obtained from all subjects and their parents involved in the study.

Data Availability Statement: The data presented in this study are openly available at https://osf.io/9sz32 (accessed on 26 May 2023).

Acknowledgments: We are thankful to all who distributed the link to the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. National Center for Education Statistics. Science, Technology, Engineering, and Mathematics (STEM) Education by Gender. Available online: https://nces.ed.gov/fastfacts/display.asp?id=899 (accessed on 4 May 2023).
2. Statistisches Bundesamt [German Federal Statistical Office]. Students Enrolled in STEM Courses. Available online: https://www.destatis.de/EN/Themes/Society-Environment/Education-Research-Culture/Institutions-Higher-Education/Tables/students-in-stem-courses.html (accessed on 4 May 2023).
3. Charlesworth, T.E.S.; Banaji, M.R. Gender in Science, Technology, Engineering, and Mathematics: Issues, Causes, Solutions. *J. Neurosci.* **2019**, *39*, 7228–7243. [CrossRef] [PubMed]

4. Dasgupta, N.; Stout, J.G. Girls and Women in Science, Technology, Engineering, and Mathematics: STEMing the Tide and Broadening Participation in STEM Careers. *Policy Insights Behav. Brain Sci.* **2014**, *1*, 21–29. [CrossRef]
5. Wang, M.-T.; Degol, J.L. Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. *Educ. Psychol. Rev.* **2017**, *29*, 119–140. [CrossRef]
6. Master, A.; Meltzoff, A.N. Cultural Stereotypes and Sense of Belonging Contribute to Gender Gaps in STEM. *Int. J. Gend. Sci. Technol.* **2020**, *12*, 152–198. Available online: https://genderandset.open.ac.uk/index.php/genderandset/article/view/674 (accessed on 18 January 2021).
7. Else-Quest, N.M.; Hyde, J.S.; Linn, M.C. Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis. *Psychol. Bull.* **2010**, *136*, 103–127. [CrossRef]
8. Hyde, J.S.; Mertz, J.E. Gender, Culture, and Mathematics Performance. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8801–8807. [CrossRef]
9. Reiss, K.; Weis, M.; Klieme, E.; Köller, O. *PISA 2018*; Waxmann Verlag GmbH: Münster, Germany, 2019; ISBN 978-3-8309-4100-2.
10. Schwippert, K.; Kasper, D.; Köller, O.; McElvany, N.; Selter, C.; Steffensky, M.; Wendt, H. *TIMSS 2019. Mathematische und Naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im Internationalen Vergleich*; Waxmann Verlag GmbH: Münster, Germany, 2020; ISBN 978-3-8309-4319-8.
11. Master, A. Gender Stereotypes Influence Children's STEM Motivation. *Child Dev. Perspect.* **2021**, *15*, 203–210. [CrossRef]
12. Nosek, B.A.; Smyth, F.L.; Sriram, N.; Lindner, N.M.; Devos, T.; Ayala, A.; Bar-Anan, Y.; Bergh, R.; Cai, H.; Gonsalkorale, K.; et al. National Differences in Gender–Science Stereotypes Predict National Sex Differences in Science and Math Achievement. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10593–10597. [CrossRef]
13. Schmader, T. Gender Inclusion and Fit in STEM. *Annu. Rev. Psychol.* **2023**, *74*, 219–243. [CrossRef]
14. Starr, C.R. "I'm Not a Science Nerd!": STEM Stereotypes, Identity, and Motivation Among Undergraduate Women. *Psychol. Women Q.* **2018**, *42*, 489–503. [CrossRef]
15. Master, A.; Meltzoff, A.N.; Cheryan, S. Gender Stereotypes about Interests Start Early and Cause Gender Disparities in Computer Science and Engineering. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2100030118. [CrossRef] [PubMed]
16. Kessels, U.; Heyder, A.; Latsch, M.; Hannover, B. How Gender Differences in Academic Engagement Relate to Students' Gender Identity. *Educ. Res.* **2014**, *56*, 220–229. [CrossRef]
17. Eagly, A.H.; Chaiken, S. *The Psychology of Attitudes*; Harcourt Brace Jovanovich: Fort Worth, TX, USA, 1993.
18. Zanna, M.P.; Rempel, J.K. Attitudes: A New Look at an Old Concept. In *The Social Psychology of Knowledge*; Bar-Tal, D., Kruglanski, A.W., Eds.; Cambridge University Press: Cambridge, UK, 1988; pp. 315–334.
19. Eccles, J.S.; Wigfield, A. Motivational Beliefs, Values, and Goals. *Annu. Rev. Psychol.* **2002**, *53*, 109–132. [CrossRef] [PubMed]
20. Fazio, R.H.; Zanna, M.P. Multiple Processes by Which Attitudes Guide Behavior: The Mode Model as an Integrative Framework. In *Advances in Experimental Social Psychology*; Zanna, M.P., Ed.; Academic Press: San Diego, CA, USA, 1990; Volume 23, pp. 75–109.
21. Frenzel, A.C.; Pekrun, R.; Goetz, T. Girls and Mathematics—A "Hopeless" Issue? A Control-Value Approach to Gender Differences in Emotions towards Mathematics. *Eur. J. Psychol. Educ.* **2007**, *22*, 497–514. [CrossRef]
22. Gaspard, H.; Dicke, A.-L.; Flunger, B.; Schreier, B.; Häfner, I.; Trautwein, U.; Nagengast, B. More Value through Greater Differentiation: Gender Differences in Value Beliefs about Math. *J. Educ. Psychol.* **2015**, *107*, 663–677. [CrossRef]
23. Reilly, D.; Neumann, D.L.; Andrews, G. Investigating Gender Differences in Mathematics and Science: Results from the 2011 Trends in Mathematics and Science Survey. *Res. Sci. Educ.* **2019**, *49*, 25–50. [CrossRef]
24. Watt, H.M.G. Development of Adolescents' Self-Perceptions, Values, and Task Perceptions According to Gender and Domain in 7th- through 11th-Grade Australian Students. *Child Dev.* **2004**, *75*, 1556–1574. [CrossRef]
25. Osborne, J.; Simon, S.; Collins, S. Attitudes towards Science: A Review of the Literature and Its Implications. *Int. J. Sci. Educ.* **2003**, *25*, 1049–1079. [CrossRef]
26. Guo, J.; Parker, P.D.; Marsh, H.W.; Morin, A.J.S. Achievement, Motivation, and Educational Choices: A Longitudinal Study of Expectancy and Value Using a Multiplicative Perspective. *Dev. Psychol.* **2015**, *51*, 1163–1176. [CrossRef]
27. Jiang, S.; Simpkins, S.D.; Eccles, J.S. Individuals' Math and Science Motivation and Their Subsequent STEM Choices and Achievement in High School and College: A Longitudinal Study of Gender and College Generation Status Differences. *Dev. Psychol.* **2020**, *56*, 2137–2151. [CrossRef]
28. Levine, S.C.; Pantoja, N. Development of Children's Math Attitudes: Gender Differences, Key Socializers, and Intervention Approaches. *Dev. Rev.* **2021**, *62*, 100997. [CrossRef]
29. Mao, P.; Cai, Z.; He, J.; Chen, X.; Fan, X. The Relationship Between Attitude Toward Science and Academic Achievement in Science: A Three-Level Meta-Analysis. *Front. Psychol.* **2021**, *12*, 784068. [CrossRef]
30. Watt, H.M.G.; Shapka, J.D.; Morris, Z.A.; Durik, A.M.; Keating, D.P.; Eccles, J.S. Gendered Motivational Processes Affecting High School Mathematics Participation, Educational Aspirations, and Career Plans: A Comparison of Samples from Australia, Canada, and the United States. *Dev. Psychol.* **2012**, *48*, 1594–1611. [CrossRef]
31. Greenwald, A.G.; Lai, C.K. Implicit Social Cognition. *Annu. Rev. Psychol.* **2020**, *71*, 419–445. [CrossRef] [PubMed]
32. Bargh, J.A. The Four Horsemen of Automaticity: Awareness, Intention, Efficiency, and Control in Social Cognition. In *Handbook of Social Cognition*; Wyer, R.S., Srull, T.K., Eds.; Erlbaum: Hillsdale, NJ, USA, 1994; Volume 1, pp. 1–40.
33. Greenwald, A.G.; Banaji, M.R. Implicit Social Cognition: Attitudes, Self-Esteem, and Stereotypes. *Psychol. Rev.* **1995**, *102*, 4–27. [CrossRef] [PubMed]

34. Gawronski, B.; Bodenhausen, G.V. Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit and Explicit Attitude Change. *Psychol. Bull.* **2006**, *132*, 692–731. [CrossRef]
35. Strack, F.; Deutsch, R. Reflective and Impulsive Determinants of Social Behavior. *Personal. Soc. Psychol. Rev.* **2004**, *8*, 220–247. [CrossRef] [PubMed]
36. Galdi, S.; Arcuri, L.; Gawronski, B. Automatic Mental Associations Predict Future Choices of Undecided Decision-Makers. *Science* **2008**, *321*, 1100–1102. [CrossRef]
37. Block, K.; Hall, W.; Schmader, T.; Inness, M.; Croft, E. Should I Stay or Should I Go? Women's Implicit Stereotypic Associations Predict Their Commitment and Fit in STEM. *Soc. Psychol.* **2018**, *49*, 243–251. [CrossRef]
38. Greenwald, A.G.; Poehlman, T.A.; Uhlmann, E.; Banaji, M. Understanding and Using the Implicit Association Test: III. Meta-Analysis of Predictive Validity. *J. Pers. Soc. Psychol.* **2009**, *97*, 17–41. [CrossRef] [PubMed]
39. Miller, D.I.; Eagly, A.H.; Linn, M.C. Women's Representation in Science Predicts National Gender-Science Stereotypes: Evidence from 66 Nations. *J. Educ. Psychol.* **2015**, *107*, 631–644. [CrossRef]
40. Nosek, B.A.; Smyth, F.L.; Hansen, J.J.; Devos, T.; Lindner, N.M.; Ranganath, K.A.; Tucker Smith, C.; Olson, K.R.; Chugh, D.; Greenwald, A.G.; et al. Pervasiveness and Correlates of Implicit Attitudes and Stereotypes. *Eur. Rev. Soc. Psychol.* **2007**, *18*, 36–88. [CrossRef]
41. Nosek, B.A.; Smyth, F.L. Implicit Social Cognitions Predict Sex Differences in Math Engagement and Achievement. *Am. Educ. Res. J.* **2011**, *48*, 1125–1156. [CrossRef]
42. Steffens, M.C.; Jelenec, P.; Noack, P. On the Leaky Math Pipeline: Comparing Implicit Math-Gender Stereotypes and Math Withdrawal in Female and Male Children and Adolescents. *J. Educ. Psychol.* **2010**, *102*, 947–963. [CrossRef]
43. Steffens, M.C.; Jelenec, P. Separating Implicit Gender Stereotypes Regarding Math and Language: Implicit Ability Stereotypes Are Self-Serving for Boys and Men, but Not for Girls and Women. *Sex Roles* **2011**, *64*, 324–335. [CrossRef]
44. Morrissey, K.; Hallett, D.; Bakhtiar, A.; Fitzpatrick, C. Implicit Math-Gender Stereotype Present in Adults but Not in 8th Grade. *J. Adolesc.* **2019**, *74*, 173–182. [CrossRef]
45. Cvencek, D.; Meltzoff, A.N.; Greenwald, A.G. Math-Gender Stereotypes in Elementary School Children. *Child Dev.* **2011**, *82*, 766–779. [CrossRef]
46. Degner, J.; Mangels, J.; Zander, L. Visualizing Gendered Representations of Male and Female Teachers Using a Reverse Correlation Paradigm. *Soc. Psychol.* **2019**, *50*, 233–251. [CrossRef]
47. Kessels, U.; Rau, M.; Hannover, B. What Goes Well with Physics? Measuring and Altering the Image of Science. *Br. J. Educ. Psychol.* **2006**, *76*, 761–780. [CrossRef]
48. Nosek, B.A.; Banaji, M.R.; Greenwald, A.G. Math = Male, Me = Female, Therefore Math $\neq$ Me. *J. Pers. Soc. Psychol.* **2002**, *83*, 44–59. [CrossRef]
49. Cvencek, D.; Kapur, M.; Meltzoff, A.N. Math Achievement, Stereotypes, and Math Self-Concepts among Elementary-School Students in Singapore. *Learn. Instr.* **2015**, *39*, 1–10. [CrossRef]
50. Dunlap, S.T.; Barth, J.M. Career Stereotypes and Identities: Implicit Beliefs and Major Choice for College Women and Men in STEM and Female-Dominated Fields. *Sex Roles* **2019**, *81*, 548–560. [CrossRef]
51. Lane, K.A.; Goh, J.X.; Driver-Linn, E. Implicit Science Stereotypes Mediate the Relationship between Gender and Academic Participation. *Sex Roles* **2012**, *66*, 220–234. [CrossRef]
52. Cvencek, D.; Brečić, R.; Gaćeša, D.; Meltzoff, A.N. Development of Math Attitudes and Math Self-Concepts: Gender Differences, Implicit–Explicit Dissociations, and Relations to Math Achievement. *Child Dev.* **2021**, *92*, e940–e956. [CrossRef]
53. Potvin, P.; Hasni, A. Interest, Motivation and Attitude towards Science and Technology at K-12 Levels: A Systematic Review of 12 Years of Educational Research. *Stud. Sci. Educ.* **2014**, *50*, 85–129. [CrossRef]
54. Prieto-Rodriguez, E.; Sincock, K.; Blackmore, K. STEM Initiatives Matter: Results from a Systematic Review of Secondary School Interventions for Girls. *Int. J. Sci. Educ.* **2020**, *42*, 1144–1161. [CrossRef]
55. Stoeger, H.; Heilemann, M.; Debatin, T.; Hopp, M.D.S.; Schirner, S.; Ziegler, A. Nine Years of Online Mentoring for Secondary School Girls in STEM: An Empirical Comparison of Three Mentoring Formats. *Ann. N. Y. Acad. Sci.* **2021**, *1483*, 153–173. [CrossRef]
56. Greenwald, A.G.; McGhee, D.E.; Schwartz, J.L.K. Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *J. Pers. Soc. Psychol.* **1998**, *74*, 1464–1480. [CrossRef]
57. Sriram, N.; Greenwald, A.G. The Brief Implicit Association Test. *Exp. Psychol.* **2009**, *56*, 283–294. [CrossRef] [PubMed]
58. Bar-Anan, Y.; Nosek, B.A. A Comparative Investigation of Seven Indirect Attitude Measures. *Behav. Res. Methods* **2014**, *46*, 668–688. [CrossRef]
59. Gawronski, B.; Hahn, A. Implicit Measures: Procedures, Use, and Interpretation. In *Measurement in Social Psychology*; Blanton, H., LaCroix, J.M., Webster, G.D., Eds.; Taylor & Francis: New York, NY, USA, 2019; pp. 29–55, ISBN 978-0-429-45292-5.
60. Ladewig, A.; Keller, M.; Klusmann, U. Sense of Belonging as an Important Factor in the Pursuit of Physics: Does It Also Matter for Female Participants of the German Physics Olympiad? *Front. Psychol.* **2020**, *11*, 548781. [CrossRef]
61. Cheryan, S.; Plaut, V.C.; Davies, P.G.; Steele, C.M. Ambient Belonging: How Stereotypical Cues Impact Gender Participation in Computer Science. *J. Pers. Soc. Psychol.* **2009**, *97*, 1045–1060. [CrossRef]
62. Good, C.; Rattan, A.; Dweck, C.S. Why Do Women Opt out? Sense of Belonging and Women's Representation in Mathematics. *J. Pers. Soc. Psychol.* **2012**, *102*, 700–717. [CrossRef] [PubMed]

63. Lewis, K.L.; Stout, J.G.; Finkelstein, N.D.; Pollock, S.J.; Miyake, A.; Cohen, G.L.; Ito, T.A. Fitting in to Move Forward: Belonging, Gender, and Persistence in the Physical Sciences, Technology, Engineering, and Mathematics (PSTEM). *Psychol. Women Q.* **2017**, *41*, 420–436. [CrossRef]
64. Master, A.; Cheryan, S.; Meltzoff, A.N. Computing Whether She Belongs: Stereotypes Undermine Girls' Interest and Sense of Belonging in Computer Science. *J. Educ. Psychol.* **2016**, *108*, 424–437. [CrossRef]
65. Höhne, E.; Zander, L. Belonging Uncertainty as Predictor of Dropout Intentions among First-Semester Students of the Computer Sciences. *Z. Erzieh.* **2019**, *22*, 1099–1119. [CrossRef]
66. Stoet, G. PsyToolkit: A Software Package for Programming Psychological Experiments Using Linux. *Behav. Res. Methods* **2010**, *42*, 1096–1104. [CrossRef] [PubMed]
67. Stoet, G. PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teach. Psychol.* **2017**, *44*, 24–31. [CrossRef]
68. Nosek, B.A.; Bar-Anan, Y.; Sriram, N.; Axt, J.; Greenwald, A.G. Understanding and Using the Brief Implicit Association Test: Recommended Scoring Procedures. *PLoS ONE* **2014**, *9*, e110938. [CrossRef]
69. Schönbrodt, F.D.; Perugini, M. At What Sample Size Do Correlations Stabilize? *J. Pers.* **2013**, *47*, 609–612. [CrossRef]
70. Stoeger, H.; Duan, X.; Schirner, S.; Greindl, T.; Ziegler, A. The Effectiveness of a One-Year Online Mentoring Program for Girls in STEM. *Comput. Educ.* **2013**, *69*, 408–418. [CrossRef]
71. Stoeger, H.; Schirner, S.; Laemmle, L.; Obergriesser, S.; Heilemann, M.; Ziegler, A. A Contextual Perspective on Talented Female Participants and Their Development in Extracurricular STEM Programs. *Ann. N. Y. Acad. Sci.* **2016**, *1377*, 53–66. [CrossRef] [PubMed]
72. Dweck, C.S. *Self-Theories: Their Role in Motivation, Personality, and Development*; Essays in Social Psychology; Psychology Press: Philadelphia, PA, USA, 1999; ISBN 978-0-86377-570-3.
73. Stout, J.G.; Dasgupta, N.; Hunsinger, M.; McManus, M.A. STEMing the Tide: Using Ingroup Experts to Inoculate Women's Self-Concept in Science, Technology, Engineering, and Mathematics (STEM). *J. Pers. Soc. Psychol.* **2011**, *100*, 255–270. [CrossRef]
74. Kuchynka, S.L.; Reifsteck, T.V.; Gates, A.E.; Rivera, L.M. Which STEM Relationships Promote Science Identities, Attitudes, and Social Belonging? A Longitudinal Investigation with High School Students from Underrepresented Groups. *Soc. Psychol. Educ.* **2022**, *25*, 819–843. [CrossRef]
75. Young, D.M.; Rudman, L.A.; Buettner, H.M.; McLean, M.C. The Influence of Female Role Models on Women's Implicit Science Cognitions. *Psychol. Women Q.* **2013**, *37*, 283–292. [CrossRef]
76. Gawronski, B.; Bodenhausen, G.V. The Associative–Propositional Evaluation Model: Theory, Evidence, and Open Questions. *Adv. Exp. Soc. Psychol.* **2011**, *44*, 59–127. [CrossRef]
77. Lai, C.K.; Skinner, A.L.; Cooley, E.; Murrar, S.; Brauer, M.; Devos, T.; Calanchini, J.; Xiao, Y.J.; Pedram, C.; Marshburn, C.K.; et al. Reducing Implicit Racial Preferences: II. Intervention Effectiveness across Time. *J. Exp. Psychol. Gen.* **2016**, *145*, 1001–1016. [CrossRef]
78. Cheryan, S.; Ziegler, S.A.; Montoya, A.K.; Jiang, L. Why Are Some STEM Fields More Gender Balanced than Others? *Psychol. Bull.* **2017**, *143*, 1–35. [CrossRef]
79. Leslie, S.-J.; Cimpian, A.; Meyer, M.; Freeland, E. Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines. *Science* **2015**, *347*, 262–265. [CrossRef] [PubMed]
80. Starr, C.R.; Leaper, C. Undergraduates' pSTEM Identity and Motivation in Relation to Gender- and Race-Based Perceived Representation, Stereotyped Beliefs, and Implicit Associations. *Group Process. Intergroup Relat.* **2022**, 136843022211282. [CrossRef]
81. Starr, C.R.; Anderson, B.R.; Green, K.A. "I'm a Computer Scientist!": Virtual Reality Experience Influences Stereotype Threat and STEM Motivation Among Undergraduate Women. *J. Sci. Educ. Technol.* **2019**, *28*, 493–507. [CrossRef]

education sciences

Article

What Matters for Boys Does Not Necessarily Matter for Girls: Gender-Specific Relations between Perceived Self-Determination, Engagement, and Performance in School Mathematics

Sarah Isabelle Hofer [1,*], Frank Reinhold [2], Dilan Hulaj [1], Marco Koch [3] and Jörg-Henrik Heine [4]

[1] Faculty of Psychology and Educational Sciences, Ludwig-Maximilians-University, 80802 Munich, Germany
[2] Institute for Mathematics Education, University of Education, 79117 Freiburg, Germany
[3] Individual Differences & Psychodiagnostics, Saarland University, 66123 Saarbrücken, Germany
[4] TUM School of Social Sciences and Technology, Technical University, 80333 Munich, Germany
* Correspondence: sarah.hofer@psy.lmu.de

Abstract: While math performance does not seem to differ systematically between males and females, it is one of the subjects that is consistently perceived as "male" with girls regularly reporting lower levels of motivation and less positive attitudes than boys. This study aimed to uncover gender-specific relations between perceived self-determination, engagement, and performance in school mathematics that might help to better understand this discrepancy. In an online study, we hence assessed perceived competence and autonomy support, social relatedness, cognitive and behavioral engagement, math performance as well as sustained attention as a basic cognitive prerequisite in a sample of N = 221 Seventh-Grade students from southern Germany (M_{age} = 12.84 years, SD_{age} = 0.55, $N_{females}$ = 115). As expected, we found no gender differences in math performance. In multiple group path analyses, perceived autonomy support was the most consistent predictor of cognitive and behavioral engagement for both girls and boys. While it did not affect math performance directly, we found significant indirect effects via cognitive engagement for girls, and via behavioral engagement for boys, whereas competence support in the math classroom, which female students perceived as significantly lower than male students, negatively predicted only girls' performance, sustained attention explained a considerable part of boys' math performance. Girls seem to experience competence support less often than boys, and if they do, we assume it to be in response to low performance rather than to encourage high competence and nurture talent. Our results suggest promising avenues for future research and implications for math classrooms.

Keywords: math performance; gender differences; engagement; self-determination; teacher support

Citation: Hofer, S.I.; Reinhold, F.; Hulaj, D.; Koch, M.; Heine, J.-H. What Matters for Boys Does Not Necessarily Matter for Girls: Gender-Specific Relations between Perceived Self-Determination, Engagement, and Performance in School Mathematics. *Educ. Sci.* **2022**, *12*, 775. https://doi.org/10.3390/educsci12110775

Academic Editor: Mike Joy

Received: 4 August 2022
Accepted: 27 October 2022
Published: 1 November 2022

1. Introduction

Understanding and explaining gender differences in preferences and performance in STEM fields are complex [1–4]. While researchers still argue about the relative importance of biological and socio-cultural factors, most agree that the STEM gender gap is a product of both nature and nurture [1,3,5–7]. In this study, we focus on the core subject of math, due to its role as a gatekeeper to educational success in many other subjects at school and beyond [8,9]. Although gender disparities in math appear to have decreased in more recent times and can be considered less pronounced than in engineering, physics, or computer science [10,11], math is still perceived as a male subject [12,13] requiring innate talent or brilliance that is stereotypically more strongly attributed to men than to women [14]. Gender-specific nature x nurture-influences in the domain of math seem to be particularly evident in motivational-affective variables such as attitudes, choices, self-concept, or interest that consistently show significant gender differences to the detriment of females [12,15–17]. The results on performance, however, are less consistent. While many studies do not find any gender differences [11,17–19], the direction of effects, if any, also varies across countries (for a cross-national meta-analysis, see [16]; for recent results from China rather indicating a small female advantage, see [20]; for an international perspective based on PISA data, see [21]), and the frequently reported advantage of girls in terms of school grades seems to be smallest or not existing in math [22,23]. Although motivational theories usually posit generally valid causal relations between different motivational-affective constructs and academic performance, research on the existence and extent of gender differences in math motivation, where differences seem to exist, and performance, where girls and boys appear to be more alike than different (or, at least, results show a large variance across studies and nations), indicates that these relations might differ systematically between girls and boys.

Motivation is considered to affect performance via increased engagement [24–26]. In this study, we hence not only analyze whether girls indeed show lower levels of motivation in math classrooms and similar levels of math performance compared to boys, but also investigate whether boys and girls differ in how motivation translates into engagement and math performance. Gender differences in these paths could be part of an explanation for the above-mentioned discrepancy between math motivation and performance among female students. The underlying cognitive capacity to engage, i.e., sustained attention, might also play an important role in predicting girls' and boys' engagement and math performance and is therefore controlled for as well [27,28].

Different from other motivational variables, such as interest or self-concept, that may emerge from diverse internal and external factors [29–31], perceived self-determination during math instruction directly reflects a student's perception of a very specific situation. Perceived self-determination (i.e., perceived competence support, autonomy support, and social relatedness) during math instruction can be expected to immediately affect the energy level in students' behavior and the intensity of their engagement in the math classroom. In the remainder of the introduction, we hence introduce self-determination theory and briefly describe cognitive and behavioral engagement as well as the concept of sustained attention (as control variable), whenever possible focusing on their role in the context of math performance and summarizing the current state of research on gender differences on these variables. Before we present our study, we synthesize research investigating relations between these variables.

1.1. Perceived Support for Self-Determination in the Classroom

Self-determination theory considers the satisfaction of three basic psychological needs—the needs of competence, autonomy, and social relatedness—as an important motive for behavior [32]. In school contexts, perceived self-determination is higher, if a situation is experienced as conducive to need satisfaction [33,34]. One common way to investigate needs satisfaction in the math classroom is by assessing (i) how students perceive the level of information provided about their own competence or progress during their math lessons (competence support), (ii) how autonomous they feel during math lessons in school (autonomy support), and (iii) how they perceive the connectedness to other students in the math classroom (social relatedness); [32,35].

Self-determination theory is considered important for understanding the underlying processes of (math-related) motivation and achievement, as students' perceived support for self-determination influences the way students learn [36,37]. In addition to evidence that autonomy support can positively affect later motivation and achievement in math [38–40], a review article more generally underpins the positive relation between classroom instruction that supports the satisfaction of those basic needs, and students' motivation, engagement, and achievement [33]. In a recent study by Szulawski et al. [41] who analyzed the influence of basic psychological needs on performance when external incentives are present, perceived competence support appeared to be the strongest positive predictor of task performance. In their literature review, Guay et al. [42] found evidence for the positive effects of perceived self-determination on performance and learning—again, in particular, for autonomous motivation that is fostered by autonomy support in the classroom. Focusing on gender disparities in science classes, Patall et al. [43] reported lower perceived autonomy support and need satisfaction for female compared to male high school students. Although not addressing perceived competence support, there is broad evidence for a male advantage in terms of the related constructs of math self-efficacy, self-esteem, and self-concept [19,44,45]. Finally, corresponding to widespread math-gender stereotypes [46–48], teachers might be biased in their attribution and evaluation of female vs. male students' performance and behavior—and accordingly, act differently towards girls and boys in the class [20,49]. Gender-related differences in teacher-created learning environments should be reflected in, on average, lower levels of female students' perceived support for self-determination during math instruction.

1.2. Cognitive and Behavioral Engagement

Students' classroom engagement is an empirically validated predictor of academic achievement in math [50–53] that is usually understood as a multifaceted construct, consisting of behavioral, cognitive, and, sometimes, emotional aspects [54]. In the present study, we focus on the most prominent and stable facets of cognitive and behavioral engagement.

Cognitive engagement can be understood as students' "desire to go beyond the requirements and [their] preference for challenge [including] flexibility in problem solving, [and] preference for hard work" [54] (pp. 63–64). Following this conceptualization, cognitive engagement describes the extent of cognitive effort invested by the student—i.e., whether deep learning strategies and adequate cognitive strategies for comprehension are used when learning math [55]. Behavioral engagement, by contrast, represents students' observable behavior during math lessons rather than psychological and cognitive processes—i.e., positive conduct, effort, attention, involvement, persistence, and active participation in classroom and learning activities [54,56].

Regarding gender-specific differences, there is broad evidence that girls seem to be more engaged than boys during classroom learning [57–62]. Looking specifically at math and science and the individual facets of engagement, Fredricks and colleagues [24] likewise reported a female advantage in cognitive and behavioral engagement in middle and high school students—although other studies have suggested an advantage for boys [51,63] or no significant gender differences [64].

1.3. Sustained Attention

We are also interested in students' sustained attention due to its important role in ultimately enabling and constraining engagement and performance [28,65,66]. Sustained attention enables learners to maintain focus on a task for a longer time [27]. High levels of sustained attention require vigilance as well as executive functions such as response inhibition and distractor suppression [67]. Sustained attention is hence important to stay on-task, show on-task behavior, and resist distractions [68,69]—especially during learning that involves complex processing of information [28,65]. Accordingly, sustained attention has been shown to moderate the relationship between intellectual potential and performance (including math grades) at school [66]. Although the basic cognitive function of sustained attention is heritable [70], it can also be affected by nurture during an individual's development [71]. Existing results on gender effects on sustained attention are mixed. While some studies do not find any differences [72], there is evidence for male or female advantages on subcomponents, including vigilance and inhibitory control [73,74]. Riley and colleagues [75], using a gradual-onset continuous performance task, found less variance and faster performance together with more commission errors for males and more omission errors among females. Currently, however, there is no indication to expect gender differences on an overall measure of sustained attention.

1.4. Interrelations between these Bariables and the Role of Gender

While a few studies have looked at interrelations between those constructs of interest to our study, researchers have just started to analyze potential differential effects of students' gender. Simpkins and colleagues [76], for instance, investigated how children's motivational beliefs predict changes in their behavioral engagement in sports, instrumental music, and math. Despite domain-specific mean differences between boys and girls, no significant gender differences were found for the relations between the constructs over time, suggesting similar socialization processes [76]. Fredricks et al. [24] likewise analyzed associations between middle and high school students' motivational beliefs (including utility and attainment values as well as expectancy beliefs), social relatedness support from peers and teachers, and characteristics of instruction with engagement in math and science classrooms. While their quantitative analyses indicated significant relations between the motivational and contextual variables and engagement that were similar for boys and girls, there was some evidence for gender differences derived from student interviews. A personal relationship with their teacher (cf., perceived social relatedness) seemed to be more important to be engaged during instruction for girls than for boys. While teacher support was significantly related to engagement for boys and girls in the quantitative analyses as well, there was a moderating effect of gender with respect to teacher social support, which was associated with higher behavioral engagement in math for girls substantiating the conclusions from the interview. Although many students (irrespective of gender) reported that their participation and engagement in science and math classes also depended on their perceived competence as demonstrated in front of their teachers and peers, female students were more likely to also talk about disengagement and frustration in the face of content perceived as too challenging—particularly in math classes; and they more often mentioned that they preferred to keep silent rather than participate for fear of looking stupid. To sum up, social teacher support (cf. perceived social relatedness) hence seems to be particularly important for girls in math and more so for behavioral engagement. Perceived competence (support) appears to be critical for both girls' and boys' engagement, although girls might have less confidence in their abilities and be more susceptible to external evaluation. Moreover, there is evidence that perceived social relatedness (or classroom emotional climate) is directly and indirectly—via engagement—related to academic achievement [77,78]. In Reyes et al.'s [78] study, the effects were robust across gender. In line with the findings on perceived competence in Fredericks et al.'s [24] study but without focusing on gender effects, individuals, in general, seem to be more likely to engage cognitively, which is associated with higher performance, when they feel competent in a specific content area [26]. Similarly, Shernoff et al. [79] expect autonomy support and appropriate challenge (i.e., competence support) to lead to more engagement. Patall et al. [43] also did not investigate gender effects on relations between variables but analyzed whether male and female high school students differ in perceived autonomy support in the classroom, in need satisfaction (a joint measure of all three sources of need satisfaction), and in engagement (a joint measure of different types of engagement) in science classes and whether gender differences in engagement can be explained by gender differences in perceived autonomy support and need satisfaction. In line with their hypotheses, compared to boys, girls indicated lower autonomy support and need satisfaction and these gender differences accounted for female students' lower engagement in science courses.

To conclude, most studies that examine relations between variables we are interested in do not focus on the differential effects of gender on these relations; and those that do indicate more similarities in the patterns of relations than differences. In order to better understand what matters for girls and boys to explain math performance, we model and compare the paths between perceived self-determination support, engagement, and performance for male and female students. Unlike other studies, we differentiate between perceived competence support, autonomy support, and social relatedness, as well as between behavioral and cognitive engagement—and we additionally consider sustained attention as a basic cognitive prerequisite.

2. The Present Study

Students' engagement in the classroom can be expected to strongly depend on their perception of the learning environment: Do they feel encouraged and supported to develop competence and learn autonomously? Do they feel like they belong? Perceived autonomy support, competence support, and social relatedness may motivate students to both cognitively and behaviorally engage with the content to be learned [80–82]. Higher engagement, in turn, has been shown to be associated with learning and higher performance [78,83,84]. As motivational indicators of the perceived quality of instruction and provided learning opportunities, perceived autonomy support, competence support, and social relatedness can also be expected to directly affect performance and school success [41,78,85]. Finally, students' sustained attention may play an important role in ultimately enabling and constraining engagement and performance, as well as performance via engagement [66]. While male and female students either do not differ in terms of math performance [17] or boys rather outperform girls on standardized tests [86,87]), girls have been repeatedly reported to show, on average, lower levels of motivation [17,45,88] and higher engagement in the classroom than boys [58,62]. These findings suggest gender differences in the prediction of math performance. To conclude, based on potential differences in gender-specific external (teachers) and internal attributions, relations between perceived autonomy support, competence support, and social relatedness, engagement and performance in school mathematics can be expected to differ between girls and boys. In this study, we hence test the following hypotheses derived from existing research stating (1) a male advantage in terms of perceived self-determination in the math classroom, a female advantage in terms of engagement in the math classroom, and no gender differences in terms of math performance and sustained attention, and (2) predict math performance by modeling the direct paths from the three self-determination variables and the indirect path via cognitive and behavioral engagement comparing the resulting models for girls and boys. In an additional analysis, we further include the direct and indirect (via engagement) paths from students' sustained attention to math performance. The second part of this study is explorative in nature because existing research has not accumulated enough evidence yet to derive specific hypotheses regarding gender differences in the relations between the study variables. Based on online survey data from German Seventh-Grade secondary school students, *t*-tests are used to test the hypotheses of the first part of this study. For the second part, multiple group path analyses with gender as a group variable allow us to test for gender differences in the relations between the study variables and, hence, in the prediction of math performance.

3. Methods

3.1. Sample and Procedure

The sample investigated in this study has also been analyzed in another study, however, focusing on a different set of variables and addressing research questions related to remote schooling during the COVID-19 pandemic [89]. We hence examine a convenience sample of $N = 221$ Seventh-Grade students from Germany ($M_{age} = 12.84$ years, $SD_{age} = 0.55$), consisting of $n = 106$ male and $n = 115$ female students ($n = 2$ students who did not indicate male or female gender had to be excluded for the purpose of this study). Students and their parents from 18 schools in Bavaria, Germany, were informed about the study by their math teachers who belonged to a pool of teachers we had already cooperated with in the context of other projects. All students whose parents provided written consent received a link to our online survey in July 2020 to be completed in the next weeks until the end of the school year as voluntary homework assignment. The survey was run via a secure and established tool (Unipark; https://www.unipark.com/en/, accessed on 1 October 2022), conforming with the General Data Protection Regulation. The survey could be accessed from both desktop or laptop computers as well as touchscreen devices such as tablets or smartphones. The students could click through the survey at their own pace. The last page of the survey was linked to the web-based version of the sustained attention test, which is described below in more detail. The whole assessment was intended to be completed in approximately 30 min. After the end of the school year, in August 2020, we closed the survey and obtained the data. Altogether, $N = 421$ people accessed the survey, however, only $N = 223$ students produced valid data sets. The remaining $N = 198$ participants did not finish the survey. The majority stopped in the first third of the survey. Because participation in the survey was not obligatory, the number of students from one class varied considerably—from one student to 23 students at the maximum. Due to the large variation and mostly small number of students per class, it was not possible to consider the classroom structure in our analyses.

Although all $N = 221$ students completed the full set of survey items, several students did not finish the sustained attention test attached at the end of the survey, resulting in invalid measures. While the main analysis is based on the whole sample and the full set of survey items, we also report additional analyses based on the reduced sample of $N = 149$ ($M_{\text{age}} = 12.78$ years, $SD_{\text{age}} = 0.46$; $n = 72$ male and $n = 77$ female students) who also completed the sustained attention test. Since students not finishing the sustained attention test cannot be considered missing at random but rather likely share some common characteristics, the reduced sample must be conceived of as an even more selective sample than the full sample. To take this into account, we compare the model results between the different samples.

Online data collection took place at home on a voluntary basis and only for those students whose parents provided written consent after both parents and students had been informed about the study. In line with the ethics committee of the concerned institution, there was no compelling need for an ethics approval for this project.

3.2. Instruments and Scales

The instruments and scales, which have already been described in [89], are summarized in the following sections. Basic demographic information including age, gender, or school type, was assessed via self-report in the online survey.

Math performance. With Seventh-Graders in focus of the present study, we assessed math performance with 12 items representing basic knowledge of fractions. This content represents the core content of Grade Six in the present curriculum for all participants—i.e., content that should already have been taught to all the students at the time of the study. Basic fraction knowledge is operationalized with both conceptual knowledge items (e.g., students needed to name the fraction depicted in a pie chart with non-equal parts) as well as procedural knowledge items (e.g., students needed to divide 8/35 by 4/15). The instrument's reliability in the sample was estimated as Cronbach's $\alpha = 0.89$ and McDonald's $\omega = 0.91$.

Engagement. The two scales (cognitive and behavioral engagement) originate from a questionnaire by Wang and colleagues [90]. Students were prompted to think about "normal math classes" when answering the following items. Nine items on *cognitive engagement* ($\alpha = 0.78$ and $\omega = 0.83$) assess whether deep learning strategies and adequate cognitive strategies for comprehension are used (sample item: When I do not understand something in math, I try to clarify it). Eight items focusing on *behavioral engagement* ($\alpha = 0.77$ and $\omega = 0.85$) address involvement and active participation in classroom and learning activities related to math (sample item: I stay focused in math). Both scales are based on a 4-point Likert scale from 1 = "*almost never*" to 4 = "*almost always*".

Competence, autonomy, and social relatedness support. The following scales were adopted from Prenzel and Drechsel [91] and adapted to school students and math instruction. Just as with the engagement scales, students were prompted to think about "normal math classes" when answering the items. We used 4-point Likert plural from 1 = "*do not agree at all*" to 4 = "*totally agree*". The scale *perceived competence support* ($\alpha = 0.73$ and $\omega = 0.77$) assesses the students' perception of the awareness, communication, and appreciation of competence in their math lessons (five items; sample item: In math, I am informed about my individual progress). The scale *perceived autonomy support* ($\alpha = 0.64$ and $\omega = 0.75$) focuses on students' perception of the degree of autonomy they have in their math lessons (six items; sample item: In math, I have the opportunity to try out new things myself), and finally, the scale *perceived social relatedness* (α = 0.80 and ω = 0.84) addresses students' perception of the social climate and, in particular, the level of their own integration in the math classroom context (five items; sample item: In math, I feel like I belong).

Sustained attention. The attention swiping task (AST; [92]) was developed as a test of sustained visual attention that can be administered on any mobile device. Throughout the test, participants are presented with rows of nine stimuli (pictorial flowers) which are constructed based on two dichotomous dimensions, resulting in four different stimulus categories (see Figure 1A). The instructions are presented self-paced and the participants are required to remember two stimulus categories (the targets) which they will need to push toward the upper third of the screen. All stimuli which do not meet these criteria (the distractors) must be pushed toward the lower third. To further increase the requirements for participants' sustained attention, the original test was modified in a way that the background color would switch randomly indicating a rule change (i.e., the direction into which stimuli were to be pushed was inverted). After reading the instructions, participants first received a practice row of nine stimuli without any rule changes. After completing this row, they received a practice row with the rules being switched for all nine stimuli. Finally, participants received a row where the rule switched. During the practice rows, participants received visual feedback on their performance (see Figure 1B). To provide all participants with comparable prerequisites before starting the real test, each time they committed more than three mistakes per practice row, they would be presented with another practice row. This process was repeated up to three times. Once the real test started, no more feedback was provided, and participants were required to work on the task for three minutes as fast and conscientiously as possible. In order to constantly remind participants of the time limit, a progress bar was presented at the upper border of the screen, indicating how much time had passed already. Reactions to each item were collected as one of four categories (i.e., hits, omissions, mistakes, and dismissals). For further analyses, a measure of sustained attention was computed by subtracting the number of mistakes and omissions from the hits. This score was used to correct for correct responses that resulted from inattentive guessing. Due to the novelty of this test instrument, there are no norming samples yet, and the resulting score has to be interpreted relative to the study sample. In order to estimate the reliability of sustained attention, the split-half reliability has been used. To avoid overestimation, this split was carried out by separating the items at the 90 s mark (i.e., after half of the test time has passed), resulting in a reliability of r_{tt} = 0.86.

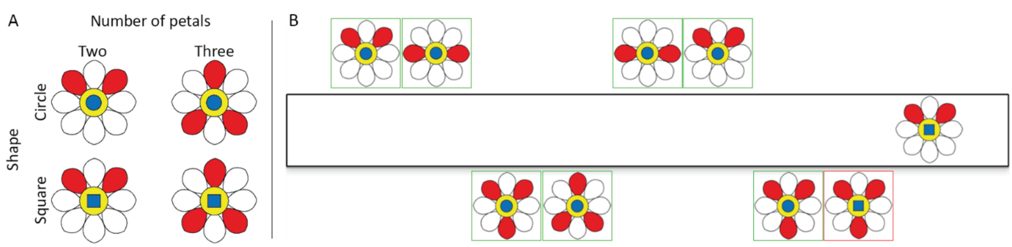


Figure 1. Stimuli and Practice Row of the AST. Note. In (**A**), all possible combinations of stimuli can be seen. The stimuli vary by the shape in the center and the number of colored petals. For each stimulus, four variants exist to prevent participants from memorizing and thus relying too much on visual comparison. In (**B**), an example from a practice trial is presented. The rule was to move stimuli with a circle and two-colored petals as well as stimuli with a square and three colored petals to the upper part of the screen. The first seven stimuli were categorized correctly and have thus been highlighted with a green box. The eighth stimulus should have also been moved to the upper part; thus, it has been marked red and the participant is required to correct their response before continuing with the last stimulus.

3.3. Statistical Analysis

While gender-specific basic correlational analyses on the study variables, descriptive statistics, and *t*-tests between female and male students were run on IBM© SPSS© Statistics Version 27, we used R-4.0.2 for data compilation and all other analyses. The multiple group path analyses and regressions were conducted with the package 'lavaan' version 0.6-7 [93]. In order to protect against potential errors due to not normally distributed variables, maximum likelihood estimation with robust standard errors and a Satorra-Bentler scaled test statistic were used. In the main analysis, both cognitive and behavioral engagement were regressed on perceived autonomy support, perceived competence support, and perceived social relatedness. Math performance was regressed on cognitive and behavioral engagement as well as on perceived autonomy support, perceived competence support, and perceived social relatedness. We also calculated the indirect paths from perceived autonomy support, perceived competence support, and perceived social relatedness via cognitive or behavioral engagement on math performance. The model was estimated separately for female and male students. The model with gender-specific estimates was compared to a joint model that did not distinguish between male and female students to test for an overall effect of gender. Differences in terms of coefficients between females and males were tested by comparing the model with all parameters estimated freely for each gender to nested models with single coefficients being constrained to be equal between female and male students. The nested models' fit to the data was contrasted using the scaled chi-squared difference test [94]. A significant chi-square indicates a significant gender difference on the coefficient set to be equal between female and male students.

3.4. Transparency and Openness

We report how we collected our sample, all data exclusions, and all measures in the study. All data and analysis code are available by emailing the corresponding author. Research materials are available at https://doi.org/10.1007/s10212-021-00590-w (accessed on 1 October 2022). This study's design and its analysis were not pre-registered.

4. Results

In the following section, we first present the results of gender-specific basic correlational analyses on the study variables, descriptive statistics, and *t*-tests between female and male students. The results of the main multiple group path analysis based on the full sample are summarized next, before we focus on the extended analysis including sustained attention based on the reduced sample of $N = 149$ students.

4.1. Basic Gender Differences

Table 1 contains the correlations between all study variables separately for girls and boys. Gender-specific descriptive statistics are presented in Table 2, together with the results of the *t*-tests.

When looking at the correlations in Table 1, we recognize gender differences in the correlations of the different study variables with math performance and, albeit to a smaller extent, with sustained attention—with less significant bivariate correlations on the part of the girls. Significant gender mean-differences exist only in terms of competence support with higher perceived support on the part of male students and in terms of behavioral engagement with girls indicating higher manifestations (Table 2). For cognitive engagement, the same tendency was evident but not significant. The girls' sustained attention also appeared to be slightly higher, however, again not reaching significance based on an α-level of 0.05.

Table 1. Correlations between all Study Variables for Girls and Boys.

	1	2	3	4	5	6	7
Girls (n = 115)							
1. Autonomy support	-						
2. Competence support	0.45 ***	-					
3. Social relatedness	0.54 ***	0.46 ***	-				
4. Cognitive engagement	0.45 ***	0.31 **	0.34 ***	-			
5. Behavioral engagement	0.38 ***	0.25 **	0.32 ***	0.75 ***	-		
6. Math performance	0.18	−0.03	0.15	0.51 ***	0.33 ***	-	
7. Sustained attention [a]	0.06	0.02	0.09	0.19	0.18	0.28 *	-
Boys (n = 106)							
1. Autonomy support	-						
2. Competence support	0.53 **	-					
3. Social relatedness	0.65 **	0.58 **	-				
4. Cognitive engagement	0.47 **	0.34 **	0.43 **	-			
5. Behavioral engagement	0.48 **	0.34 **	0.50 **	0.76 **	-		
6. Math performance	0.39 **	0.38 **	0.53 **	0.54 **	0.56 **	-	
7. Sustained attention [a]	0.30 *	0.20	0.48 **	0.50 **	0.48 **	0.70 **	-

[a] Sample size for sustained attention is $n = 77$ for girls and $n = 72$ for boys. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Table 2. Gender-specific Descriptive Statistics and t-Test Results.

Variables		Girls (n = 115)		Boys (n = 106)		t	95% CI [LL; UL]
	Range	M	SD	M	SD		
Autonomy support	1–4	2.80	0.46	2.83	0.53	−0.47	[−0.16; 0.10]
Competence support	1–4	2.97	0.50	3.14	0.60	−2.25 *	[−0.31; −0.02]
Social relatedness	1–4	3.04	0.55	3.03	0.66	0.10	[−0.15; 0.17]
Cognitive engagement	1–4	3.15	0.49	3.02	0.53	1.90	[−0.01; 0.27]
Behavioral engagement	1–4	3.08	0.50	2.94	0.53	2.09 *	[0.01; 0.28]
Math performance	0–12	8.75	3.34	8.53	3.80	0.46	[−0.73; 1.17]
Sustained attention [a]	-	12.45	23.89	4.07	28.53	1.94	[−0.17; 16.94]

Note. LL = lower limit; UL = upper limit. [a] Sample size for sustained attention is $n = 77$ for girls and $n = 72$ for boys. * $p < 0.05$.

4.2. Main Multiple Group Path Analysis

In order to test for an overall effect of gender, the model with gender-specific estimates was compared to a joint model that did not distinguish between male and female students. The latter fitted the data significantly worse (chi-squared difference = 36.27, $p < 0.001$), indicating overall gender differences. The results of the main multiple group path analysis are summarized in Table 3. Perceived autonomy support was a significant predictor for both cognitive and behavioral engagement for female and male students. Perceived social relatedness was the only other significant predictor and only for behavioral engagement for male students—although this coefficient did not differ significantly between male and female students (see Table 3). To conclude, there were no considerable gender differences in the prediction of cognitive and behavioral engagement based on the three self-determination variables.

Cognitive engagement, in turn, predicted math performance for both girls and boys. The predictor, however, was significantly stronger for female students. Behavioral engagement, by contrast, was no significant predictor for girls' math performance at all, whereas boys' math performance could be regressed on behavioral engagement, indicating a significant difference between female and male students. Social relatedness again appeared to be slightly more important for boys than for girls, being a significant predictor of only boys' math performance (but no significant gender differences). Importantly, perceived competence support turned out to be a significant negative predictor of female students' math performance, while showing no significant association with male students' performance, reflecting a significant effect of gender. Among the indirect paths to math performance, the path from autonomy support via cognitive engagement to math performance was the only one reaching significance—and only for female students (male students: $p = 0.055$). The path coefficient, however, did not significantly differ between gender. For female students, the model explained 36.77%, and for male students, 39.51% of the variance in math performance. To sum up, despite a similar proportion of variance in math performance explained, the paths to math performance showed considerable differences between female and male students. Figure 2 depicts all significant paths for female (left panel) and male (right panel) students.

Table 3. Results of the Main Multiple Group Path Analysis.

	Main Model									
	Girls (n = 115)				Boys (n = 106)				Gender Differences	
Model Path	**β**	***SE***	***z***	***p***	**β**	***SE***	***z***	***p***	**Δchi²**	***p***
Aut. sup. → Cog. eng.	0.37	0.11	3.34	<0.001	0.31	0.11	2.77	0.006	0.14	0.713
Comp. sup. → Cog. eng.	0.1	0.09	1.13	0.26	0.06	0.1	0.63	0.53	-	-
Soc. relat. → Cog. eng.	0.09	0.09	1.02	0.31	0.15	0.11	1.46	0.143	-	-
Aut. sup. → Beh. eng.	0.29	0.12	2.34	0.019	0.26	0.1	2.5	0.013	0.03	0.866
Comp. sup. → Beh. eng.	0.07	0.09	0.74	0.461	0.02	0.09	0.16	0.872	-	-
Soc. relat. → Beh. eng.	0.13	0.12	1.12	0.262	0.26	0.08	3.07	0.002	0.72	0.395
Cog. eng. → Math	4.52	0.63	7.22	<0.001	1.77	0.65	2.74	0.006	4.49	0.034
Beh. eng. → Math	−0.85	0.57	−1.49	0.135	1.62	0.64	2.53	0.011	4.23	0.04
Aut. sup. → Math	−0.11	0.79	−0.13	0.893	−0.64	0.73	−0.88	0.38	-	-
Comp. sup. → Math	−1.50	0.6	−2.49	0.013	0.49	0.59	0.83	0.405	5.65	0.017
Soc. relat. → Math	0.43	0.71	0.61	0.539	1.85	0.68	2.71	0.007	2.11	0.146
Aut. sup. → Cog. eng. → Math	1.67	0.57	2.94	0.003	0.55	0.29	1.92	0.055	4.99	0.083
Comp. sup. → Cog. eng. → Math	0.44	0.41	1.09	0.277	0.11	0.17	0.63	0.527	-	-
Soc. relat. → Cog. eng. → Math	0.42	0.41	1.02	0.305	0.27	0.23	1.2	0.232	-	-
Aut. sup. → Beh. eng. → Math	−0.24	0.17	−1.43	0.154	0.42	0.24	1.73	0.084	-	-
Comp. sup. → Beh. eng. → Math	−0.06	0.08	−0.69	0.49	0.02	0.15	0.16	0.872	-	-
Soc. relat. → Beh. eng. → Math	−0.11	0.14	−0.79	0.429	0.42	0.22	1.88	0.059	-	-
R^2										
Cognitive engagement	0.23				0.25					
Behavioral engagement	0.16				0.29					
Math performance	0.37				0.4					

Note. Δchi² is based on scaled chi-squared difference tests; the test was performed only for significant predictors.

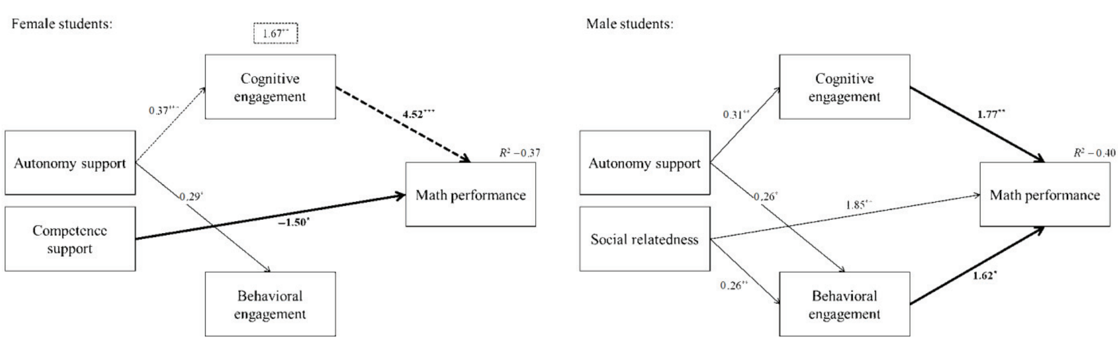


Figure 2. Significant Paths for Female (**Left** Panel) and Male (**Right** Panel) Students Based on the Main Model in the Full Sample. Note. Bold values and paths indicate significant gender differences. Dashed lines represent indirect paths and indirect path coefficients are shown in dashed boxes. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

4.3. Extended Multiple Group Path Analysis including Sustained Attention Based on Reduced Sample

In the following section, we refer to the reduced sample of only those students who also completed the sustained attention test that can be considered as an objective, behavioral measure of students' cognitive potential to engage in the math classroom. We added this predictor to the multiple group path analysis described in the previous part. To be able to estimate the effect of the reduced sample and disentangle it from the effect of including sustained attention as a predictor, we ran the main multiple group path analysis without sustained attention (described in the previous section) on the reduced sample. These results are listed in Table 4, together with the results of the extended multiple group path analysis described in the next paragraph. Importantly, the coefficients resulting from the main multiple group path analysis on the reduced sample did not differ significantly from those reported in Table 3 (i.e., we set each coefficient in the reduced sample-model to the corresponding value obtained with the full sample and compared the restricted and unrestricted reduced sample-models using scaled chi-squared difference tests; all $p > 0.126$). Although these findings suggest rather negligible distorting effects of the reduced sample, they still must be considered when interpreting the results reported in the following paragraph.

Table 4. Results of the Extended Multiple Group Path Analysis and the Main Multiple Group Path Analysis on the Reduced Sample.

	Extended Model (Reduced Sample)										**Main Model (Reduced Sample)**							
	Girls (n = 77)				**Boys (n = 72)**				**Gender Diff.**		**Girls (n = 77)**				**Boys (n = 72)**			
Model Path	**β**	***SE***	***z***	***p***	**β**	***SE***	***z***	***p***	**Δchi²**	***p***	**β**	***SE***	***z***	***p***	**β**	***SE***	***z***	***p***
Aut. sup. → Cog. eng.	0.32	0.12	2.61	0.009	0.36	0.11	3.32	<0.001	0.08	0.774	0.32	0.12	2.56	0.010	0.37	0.12	3.25	0.001
Comp. sup. → Cog. eng.	0.13	0.09	1.45	0.146	0.14	0.11	1.22	0.223	-	-	0.13	0.09	1.37	0.169	0.07	0.11	0.67	0.500
Soc. relat. → Cog. eng.	0.14	0.09	1.59	0.112	−0.07	0.14	−0.49	0.623	-	-	0.16	0.09	1.73	0.083	0.12	0.13	0.89	0.376
Sust. attent. → Cog. eng.	0.003	0.002	1.60	0.109	0.01	0.002	3.95	<0.001	2.84	0.092	-	-	-	-	-	-	-	-
Aut. sup. → Beh. eng.	0.18	0.12	1.49	0.136	0.36	0.10	3.52	<0.001	1.27	0.260	0.18	0.12	1.48	0.140	0.37	0.11	3.39	0.001
Comp. sup. → Beh. eng.	0.11	0.09	1.13	0.258	−0.002	0.13	−0.02	0.986	-	-	0.10	0.09	1.05	0.294	−0.06	0.13	−0.46	0.645
Soc. relat. → Beh. eng.	0.23	0.10	2.24	0.025	0.06	0.11	.54	0.591	1.05	0.306	0.24	0.11	2.24	0.025	0.22	0.11	1.92	0.055
Sust. attent. → Beh. eng.	0.003	0.002	1.50	0.134	0.01	0.002	2.81	0.005	1.64	0.200	-	-	-	-	-	-	-	-
Cog. eng. → Math	2.96	0.76	3.92	<0.001	−0.71	0.57	−1.26	0.206	9.37	0.002	3.18	0.81	3.94	<0.001	0.47	0.73	0.65	0.518
Beh. eng. → Math	−1.20	0.83	−1.44	0.149	1.81	0.65	2.80	0.005	5.25	0.022	−1.08	0.87	−1.24	0.214	2.28	0.76	3.00	0.003
Aut. sup. → Math	0.09	0.74	0.12	0.904	−0.33	0.74	−0.45	0.655	-	-	0.03	0.77	0.04	0.972	−0.86	0.92	−0.93	0.351
Comp. sup. → Math	−1.58	0.71	−2.23	0.026	0.99	0.60	1.67	0.096	7.89	0.005	−1.69	0.72	−2.34	0.019	0.35	0.73	0.48	0.630
Soc. relat. → Math	1.22	0.56	2.18	0.029	1.00	0.67	1.49	0.136	0.06	0.801	1.28	0.56	2.30	0.022	2.41	0.72	3.36	0.001
Sust. attent. → Math	0.03	0.01	1.95	0.051	0.07	0.02	4.19	<0.001	4.35	0.037	-	-	-	-	-	-	-	-
Aut. sup. → Cog. eng. → Math	0.94	0.44	2.12	0.034	−0.26	0.21	−1.23	0.220	9.38	0.009	1.01	0.47	2.14	0.032	0.18	0.28	0.63	0.530
Comp. sup. → Cog. eng. → Math	0.39	0.28	1.40	0.162	−0.10	0.11	−0.87	0.386	-	-	0.40	0.31	1.29	0.197	0.03	0.06	0.53	0.598
Soc. relat. → Cog. eng. → Math	0.42	0.30	1.43	0.154	0.05	0.11	0.46	0.642	-	-	0.50	0.33	1.52	0.129	0.06	0.12	0.48	0.633
Sust. attent. → Cog. eng. → Math	0.01	0.01	1.45	0.148	−0.01	0.005	−1.18	0.237	-	-	-	-	-	-	-	-	-	-
Aut. sup. → Beh. eng. → Math	−0.21	0.21	−1.00	0.317	.66	0.32	2.06	0.040	6.91	0.032	−0.19	0.21	−0.92	0.359	0.85	0.43	1.98	0.048
Comp. sup. → Beh. eng. → Math	−0.13	0.15	−0.83	0.408	−0.004	0.24	−0.02	0.986	-	-	−0.11	0.14	−0.76	0.446	−0.13	0.30	−0.44	0.663
Soc. relat. → Beh. eng. → Math	−0.27	0.22	−1.24	0.214	0.11	0.20	0.54	0.591	-	-	−0.26	0.23	−1.11	0.267	0.50	0.32	1.57	0.117
Sust. attent. → Beh. eng. → Math	−0.003	0.003	−1.14	0.254	0.01	0.01	1.79	0.073	-	-	-	-	-	-	-	-	-	-
R^2																		
Cognitive engagement	0.33				0.42						0.30				0.30			
Behavioral engagement	0.28				0.38						0.26				0.30			
Math performance	0.33				0.62						0.29				0.43			

Note. Δchi² is based on scaled chi-squared difference tests; the test was performed only for significant predictors.

For the extended model in the reduced sample, we again found a significant overall effect of gender (chi-squared difference = 38.86, $p = 0.002$). The new variable sustained attention was a significant predictor for cognitive and behavioral engagement as well as for math performance—but only for male students (with $p = 0.051$ for the prediction of girls' math performance). While the coefficients did not significantly differ between boys and girls when predicting both engagement variables, the path from sustained attention to math performance differed significantly between genders.

Similar to the main analysis, perceived autonomy support significantly predicted cognitive engagement for both boys and girls. Within the group of the girls, however, autonomy support did no longer predict behavioral engagement, but social relatedness did, while in the group of the boys, social relatedness was no longer a significant predictor of behavioral engagement. These slight differences to the results of the main multiple group path analysis based on the full sample might at least partly reflect the effects of the more selective sample, since they are also recognizable when looking at the results of the reduced sample-model without sustained attention (right part of Table 4). In line with the results of the main multiple group path analysis based on the full sample, there were no considerable gender differences in the prediction of cognitive and behavioral engagement.

The results on the prediction of math performance also strongly resembled the results of the main analysis. Cognitive engagement predicted math performance only for girls, representing a significant gender difference. Behavioral engagement, by contrast, was no significant predictor for girls' math performance, whereas boys' math performance could be regressed on behavioral engagement, indicating a significant difference between female and male students. Perceived social relatedness was a significant predictor of only female students' math performance, but we found no significant gender differences on this coefficient. Again, perceived competence support turned out to be a significant negative predictor of female students' math performance, while showing no significant association with male students' performance, reflecting a significant effect of gender.

Among the indirect paths to math performance, the path from perceived autonomy support via cognitive engagement to math performance was the only one reaching significance for female students and significantly differed between boys and girls. In the group of male students, the indirect path from autonomy support via behavioral engagement to math performance was significant and significantly differed between boys and girls. For female students, the extended model explained 32.50%, and for male students, 61.89% of the variance in math performance. To sum up, with sustained attention included, the prediction of male students' math performance improved considerably, while this behavioral measure of students' basic cognitive potential seemed to be less relevant for girls' performance. Figure 3 depicts all significant paths for female (left panel) and male (right panel) students.

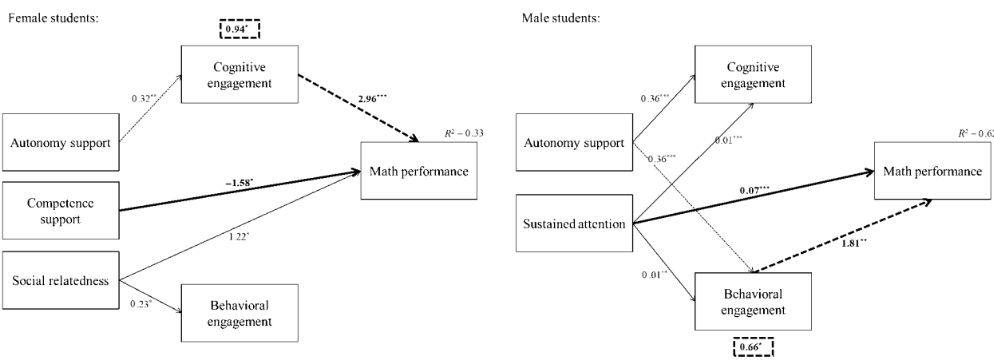


Figure 3. Significant Paths for Female (**Left** Panel) and Male (**Right** Panel) Students Based on the Extended Model in the Reduced Sample. Note. Bold values and paths indicate significant gender differences. Dashed lines represent indirect paths and indirect path coefficients are shown in dashed boxes. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

5. Discussion

While there is now widespread agreement that boys and girls do not consistently differ in their math proficiency, several studies document gender differences in math motivation as well as persistent gender-STEM stereotypes at the societal level. Thus, the question arises if different mechanisms contribute to girls' and boys' knowledge development in math. How do female and male students' (potentially different) perceptions of support in the math classroom influence their engagement and math performance and how does engagement itself affect math performance when controlling for sustained attention as a basic cognitive prerequisite? As expected, we did not find a gender difference in math performance. Regarding the motivational variables, the only significant gender difference consists of a male advantage in perceived competence support. Moreover, in line with the literature, female students reported significantly higher behavioral engagement than male students. Aside from these mean-level differences, however, what matters for math performance indeed seems to differ markedly between female and male students. Whereas perceived competence support in the math classroom and cognitive engagement seem to play a more important role for girls, behavioral engagement and, in particular, sustained attention explain a considerable part of boys' math performance. We propose that these disparate prediction paths epitomize the core of the pertaining mismatch between girls and math. In the following sections, we mainly focus on those paths that turned out to be consistent predictors in the models in both the full and the reduced sample and introduce our idea of underlying mechanisms.

5.1. The Paths Predicting Math Performance

In this study, we only look at influences on math performance that directly refer to what happens during math instruction in the classroom, thereby excluding all variables with diffuse reference and origin (e.g., self-concept, interest, anxiety). This narrow focus allows us to derive assumptions that can be tested on the classroom level (as further illustrated in the next section). We accordingly expected autonomy support, competence support and social relatedness as experienced in the math classroom to influence students' cognitive and behavioral engagement in the math classroom as well as math performance directly. Both types of engagement, in turn, were assumed to affect math performance. Sustained attention, as an objective behavioral measure of students' capacity to engage in learning during math instruction was hence hypothesized to predict both cognitive and behavioral engagement and math performance directly.

In line with existing research emphasizing the role of autonomy for learning [33,38,39,42,95], perceived autonomy support was the most consistent predictor of cognitive and behavioral engagement. There were no significant gender differences in their prediction. The extent to which male and female students are given the freedom to explore content and to work according to their own needs (autonomy support) seems to stimulate them to deeply engage with, elaborate and reflect on the content (cognitive engagement) and perseveringly and diligently participate in learning activities (behavioral engagement). Because we analyzed cross-sectional data, we cannot preclude inverse influences: The higher the students' engagement during class, the more autonomy a teacher may be able and willing to grant. However, there is existing research supporting our interpretation [42,79]. Interestingly, autonomy support indeed seems to take effect on performance via engagement. While it did not affect math performance directly, we found significant indirect effects via cognitive engagement for girls (in both the main and the extended model) and via behavioral engagement for boys (in the extended model only)—with significant gender differences.

These two divergent indirect paths reflect pronounced gender differences in the direct paths from engagement to performance, with much stronger effects of cognitive engagement on performance for girls and of behavioral engagement on performance for boys. Female students, on the mean-level, indicated higher behavioral engagement than male students, suggesting that, for girls, behavioral engagement is more the rule rather than game-changing [24,58,62]. However, the more female students reported to deeply engage with, elaborate on, and reflect on the content addressed in math classes (cognitive engagement), the higher their performance. This relation holds for boys in the main model as well, but cognitive engagement ceases to predict performance in the extended model with sustained attention included. This might suggest that the part of cognitive engagement that predicted the boys' performance in the main model largely reflected their level of sustained attention during math classes. The part of cognitive engagement that predicted the girls' performance, by contrast, showed less overlap with the basic cognitive prerequisite of sustained attention that was no significant predictor for female students at all. For them, strategic knowledge about cognitive elaboration, as well as control and reflection strategies, which can be considered to be less strongly dependent on sustained attention, seems to be of central importance. While some studies, focusing on learning strategy use do not find differences between girls and boys in science classes, language, or math [96,97], Ruffing et al. [98], investigating college students, found females to apply most learning strategies more frequently than males. These strategies could be related to learning discipline and conscientiousness. However, they incrementally predicted academic performance beyond general cognitive ability independent of students' gender. Our results might have resembled Ruffing et al.'s [98] findings if we had not distinguished between cognitive and behavioral engagement.

Sustained attention turned out to be a strong predictor of cognitive engagement, behavioral engagement, and math performance, but only for male students ($\Delta R^2 = 0.19$ for the comparison between the main and the extended model with sustained attention for boys in the reduced sample). The path from sustained attention to math performance showed a significant gender effect (stronger effect for male students). While it is perfectly in line with our expectations that students' ability to concentrate on a task and stay focused for a longer period of time predicts both engagement in a learning situation and its outcome (i.e., performance; [28,65]), the relative irrelevance of sustained attention in the female sample (which is presumably also reflected in the considerably lower proportion of variance explained, $R^2 = 0.33$, as compared to $R^2 = 0.62$ for male students) is surprising. In line with the female advantage in terms of behavioral engagement and the more important role of cognitive engagement for girls' math performance, boys', on the average, less conscientious and disciplined behavior and strategy use might increase the importance of sustained attention for learning. At the same time, it seems that something in the math classroom keeps female students from relying on their cognitive potential and investing it in learning (for similar results in physics, see [99])—as explicated in the next section.

This is the right moment to turn to another unexpected finding: the negative path from perceived competence support to math performance for girls (representing a significant gender difference). Since it is highly unlikely that higher competence support leads to lower math performance, we assume a bidirectional or even inverse relationship in this case: Female students with lower performance might receive more competence support from their math teachers—they get more time to practice, are praised more often, are more often informed about their progress and what they still can improve and are more often given credit for difficult tasks. Overall, female students seem to receive less competence support than male students, as indicated by the significant mean-level difference, and while the bivariate correlation between competence support and performance is significant and positive for male students, it is negative, albeit not significant, for female students. Girls hence seem to experience this kind of care in the math classroom less often than boys and if they do, it tends to be in response to low performance—rather than to encourage high competence and nurture talent. Such experiences in the math classroom might be the product of gender-STEM stereotypes that associate math proficiency with innate ability or brilliance that you either have or do not have and that, in turn, is more strongly associated with male than female students [12–14] and teachers, probably unintentionally, acting accordingly [100–103]. Relatively independent of female students' level of sustained attention, they might consequently not feel encouraged to invest their cognitive potential in math [19,44,45] and highly capable girls might lack chances to exploit their potential. This mechanism might partly explain the weak connection between sustained attention and math performance for girls. To build math knowledge, female students acquire and apply helpful cognitive learning strategies, while naturally behaving conscientiously and putting in the work, which seem both to be easier in autonomy-supportive classrooms [79]. Figure 4 summarizes our findings including underlying assumptions that have to be examined in future studies.

Although perceived social relatedness appeared to be positively associated with both girls' and boys' math performance and behavioral engagement, the significance of these relations varied across the different samples and models analyzed. We found no evidence of significant gender differences regarding this variable. Nevertheless, our findings are compatible with existing research that has demonstrated the importance of social relatedness in the math classroom, especially for behavioral engagement and female students [24], as well as its direct and indirect—via engagement—association with academic achievement [77,78].

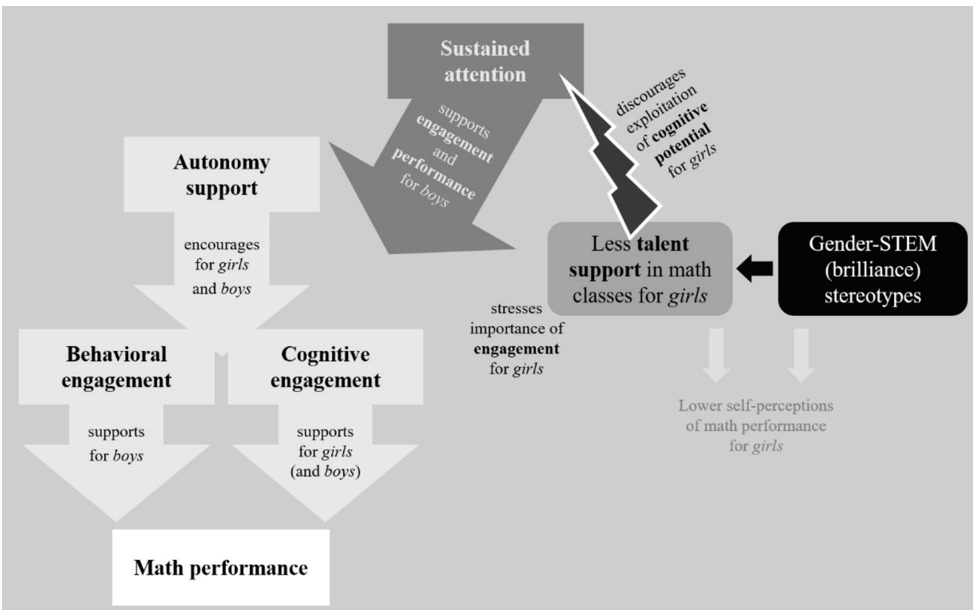


Figure 4. Gendered Pathways to Math Performance.

5.2. Limitations and Implications for Research

In particular, the hypothesized relations on the right half of Figure 4 are not derived from the present data and have to be evaluated in future studies. The idea that female students do not exploit their cognitive potential in math classrooms, which is empirically mainly based on the weaker direct and indirect influence of sustained attention on math performance among girls compared to boys, should be substantiated using other measures of cognitive potential, including intelligence measures. In addition, given the cross-sectional design of this study and the exploratory nature of our research questions on gender-specific relations, we are aware that all interpretations in terms of causality and directed effects must be considered with caution. Longitudinal studies with several measurement points of the same variables would be suited to substantiate the present findings based on the path analyses. Although we deliberately focus on students' perceptions of support in math classrooms, instead of assessing more objective instructional indicators, it would be worthwhile to contrast girls' and boys' subjective evaluations of teacher support with an objective assessment. Based on the present study, we do not know whether girls and boys are treated differently by their math teachers or whether they might interpret similar teacher behavior in different ways, maybe based on their expectations [104,105] and attributional tendencies [106,107]. In addition, experimental designs that vary the level of teacher support over a specific period of time or that implement motivational-affective interventions such as attributional retraining (for a meta-analysis on this topic, see [108,109]) could help to better understand underlying mechanisms. In the present study, we used a math performance test that is closely tied to the topics covered in the German secondary school curriculum and based on problem types typically encountered during math instruction and exams. Despite this test's high ecological validity, the present findings should be confirmed in more controlled pre-posttest design learning settings that allow direct relationships to be established between perceived support for self-determination, engagement, and performance (i.e., learning).

Finally, participants in our study could opt in or out on a voluntary basis. Accordingly, we must assume selection bias in our sample: Students completing the survey might differ systematically from other students, presumably, especially in terms of higher manifestations on variables positively associated with perseverance and compliance [110,111]. Our mean-level findings and gender-specific relations between variables may hence only apply to this specific sample and look different in a less committed group of students. However, a comparison between the overall sample and the even more selective sample of students who completed the sustained attention test revealed no substantial differences in model results—which could be interpreted as the first indication of robust effects that still need to be verified in further samples.

6. Conclusions

For researchers and practitioners alike, knowing more about the correlates of engagement and performance in the math classroom and the extent of gender differences in that regard is instrumental [76]. One central finding of this study that applies to both female and male students, is the importance of autonomy supportive instruction for student engagement. Although perceived autonomy support does not seem to affect math performance directly, we provide evidence for its indirect influence via engagement. If students are allowed to make autonomous decisions (at least to some degree) and are supported to explore content and to work according to their own needs, they seem to be more likely to deeply engage with, elaborate and reflect on the content (cognitive engagement) and perseveringly and diligently participate in learning activities (behavioral engagement).

Because behavioral engagement seems to be particularly important for boys' math performance and boys tend to show lower behavioral engagement than girls, math teachers might especially focus on creating calm and distraction-free learning environments to help increase their male students' perseverance and diligence.

The finding that girls perceive less competence support than boys together with its negative relation with math performance only in the group of female students, suggests either considerable gender differences in the perception of teacher behavior or considerable differences in teacher behavior dependent on student gender (or a mixture of both). Although we cannot disentangle the underlying mechanisms based on the existing data, we assume that lower-performing girls receive more time to practice, are praised more often, are more often informed about their progress and what they can still improve and are more often given credit for difficult tasks, while this kind of competence support might often be withheld from higher-performing female students. This does not seem to be the case in the group of the male students. Making math teachers aware of the risk of discriminative behavior involving missing out on nurturing female students with high cognitive potential (i.e., sustained attention as a proxy) and the implicit confirmation of math-gender stereotypes, could be one possible way to intervene. In addition to helping more female students to perform at a high level in math, such interventions could also be key to changing girls' self-perceptions, motivation, and attitudes toward this fundamental subject.

Author Contributions: Conceptualization, S.I.H.; methodology, S.I.H., F.R., M.K. and J.-H.H.; investigation, S.I.H.; resources, S.I.H.; data curation, D.H.; writing—original draft preparation, S.I.H.; writing—review and editing, S.I.H., F.R., D.H., M.K. and J.-H.H.; visualization, S.I.H.; supervision, S.I.H.; project administration, S.I.H.; funding acquisition, S.I.H. All authors have read and agreed to the published version of the manuscript.

Funding: Sarah I. Hofer was supported by a Jacobs Foundation Research Fellowship (number: 2018 128809).

Institutional Review Board Statement: In line with the ethics committee of the concerned institution, there was no compelling need for an ethics approval for this project.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: We report how we collected our sample, all data exclusions, and all measures in this study. All data and analysis code are available by emailing the corresponding author. Research materials are available at https://doi.org/10.1007/s10212-021-00590-w (accessed on 1 October 2022). This study's design and its analysis were not pre-registered.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ceci, S.J.; Ginther, D.K.; Kahn, S.; Williams, W.M. Women in academic science: A changing landscape. *Psychol. Sci. Public Interest J. Am. Psychol. Soc.* **2014**, *15*, 75–141. [CrossRef] [PubMed]
2. Halpern, D.F. It's complicated-in fact, it's complex: Explaining the gender gap in academic achievement in science and math. *Psychol. Sci. Public Interest J. Am. Psychol. Soc.* **2014**, *15*, 72–74. [CrossRef] [PubMed]
3. Halpern, D.F.; Benbow, C.P.; Geary, D.C.; Gur, R.C.; Hyde, J.S.; Gernsbacher, M.A. The science of sex differences in science and math. *Psychol. Sci. Public Interest J. Am. Psychol. Soc.* **2007**, *8*, 1–51. [CrossRef]
4. Hyde, J.S. The gender similarities hypothesis. *Am. Psychol.* **2005**, *60*, 581–592. [CrossRef]
5. Stewart-Williams, S.; Halsey, L.G. Men, women, and STEM: Why the differences and what should be done? *Eur. J. Personal.* **2021**, *35*, 3–39. [CrossRef]
6. Stoet, G.; Geary, D.C. The gender-equality paradox in science, technology, engineering, and mathematics education. *Psychol. Sci.* **2018**, *29*, 581–593. [CrossRef]
7. Wang, M.T.; Degol, J. Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Dev. Rev.* **2013**, *33*, 304–340. [CrossRef]
8. Jones, J.I. An overview of employment and wages in science, technology, engineering, and math (STEM) groups. *Beyond Numbers Employ. Unempl.* **2014**, *3*, 1–4.
9. Stone, J.R.; Alfeld, C.; Pearson, D. Rigor and Relevance: Enhancing High School Students' Math Skills Through Career and Technical Education. *Am. Educ. Res. J.* **2008**, *45*, 767–795. [CrossRef]
10. Cimpian, J.R.; Kim, T.H.; McDermott, Z.T. Understanding persistent gender gaps in STEM. *Science* **2020**, *368*, 1317–1319. [CrossRef]
11. Hyde, J.S.; Lindberg, S.M.; Linn, M.C.; Ellis, A.B.; Williams, C.C. Gender similarities characterize math performance. *Science* **2008**, *321*, 494–495. [CrossRef] [PubMed]
12. Nosek, B.A.; Banaji, M.R.; Greenwald, A.G. Math = Male, Me = Female, therefore Math $\neq$ Me. *J. Personal. Soc. Psychol.* **2002**, *83*, 44–59. [CrossRef]
13. Smith, J.L. The interplay among stereotypes, performance-avoidance goals, and women's math performance expectations. *Sex Roles* **2006**, *54*, 287–296. [CrossRef]
14. Leslie, S.-J.; Cimpian, A.; Meyer, M.; Freeland, E. Expectations of brilliance underlie gender distributions across academic disciplines. *Science* **2015**, *347*, 262–265. [CrossRef]
15. Ceci, S.J.; Williams, W.M. Sex differences in math-intensive fields. *Curr. Dir. Psychol. Sci.* **2010**, *19*, 275–279. [CrossRef] [PubMed]
16. Else-Quest, N.M.; Hyde, J.S.; Linn, M.C. Cross-national patterns of gender differences in math: A meta-analysis. *Psychol. Bull.* **2010**, *136*, 103–127. [CrossRef] [PubMed]
17. Rodríguez, S.; Regueiro, B.; Piñeiro, I.; Estévez, I.; Valle, A. Gender differences in math motivation: Differential effects on performance in primary education. *Front. Psychol.* **2020**, *10*, 3050. [CrossRef]
18. Lindberg, S.M.; Hyde, J.S.; Petersen, J.L.; Linn, M.C. New trends in gender and math performance: A meta-analysis. *Psychol. Bull.* **2010**, *136*, 1123–1135. [CrossRef]
19. Sewasew, D.; Schroeders, U.; Schiefer, I.M.; Weirich, S.; Artelt, C. Development of sex differences in math achievement, self-concept, and interest from grade 5 to 7. *Contemp. Educ. Psychol.* **2018**, *54*, 55–65. [CrossRef]
20. Li, M.; Zhang, Y.; Liu, H.; Hao, Y. Gender differences in math achievement in Beijing: A meta-analysis. *Br. J. Educ. Psychol.* **2018**, *88*, 566–583. [CrossRef]
21. Reinhold, F.; Reiss, K.; Diedrich, J.; Hofer, S.I.; Heinze, A. Mathematische Kompetenz in PISA 2018—Aktueller Stand Und Entwicklung [Math Competence in PISA 2018—Current Status and Development]. In *PISA 2018. Grundbildung im Internationalen Vergleich*; Reiss, K., Weis, M., Klieme, E., Köller, O., Eds.; Waxmann: Münster, Germany, 2019; pp. 187–210.
22. Downey, D.B.; Vogt Yuan, A.S. Sex differences in school performance during high school: Puzzling patterns and possible explanations. *Sociol. Q.* **2005**, *46*, 299–321. [CrossRef]
23. Voyer, D.; Voyer, S.D. Gender differences in scholastic achievement: A meta-analysis. *Psychol. Bull.* **2014**, *140*, 1174–1204. [CrossRef] [PubMed]

24. Fredricks, J.A.; Hofkens, T.; Wang, M.-T.; Mortenson, E.; Scott, P. Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *J. Res. Sci. Teach.* **2018**, *55*, 271–298. [CrossRef]
25. Schiefele, U. The role of interest in motivation and learning. In *Intelligence and Personality: Bridging the Gap in Theory and Measurement*; Collins, J.M., Messick, S., Eds.; Erlbaum: Mahwah, NJ, USA, 2001; pp. 163–194.
26. Wigfield, A.; Eccles, J.S. The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In *Development of Achievement Motivation*; Wigfield, A., Eccles, J.S., Eds.; Academic Press: Cambridge, MA, USA, 2002; pp. 91–120.
27. Blotenberg, I.; Schmidt-Atzert, L. Towards a process model of sustained attention tests. *J. Intell.* **2019**, *7*, 3. [CrossRef]
28. Schweizer, K.; Moosbrugger, H. Attention and working memory as predictors of intelligence. *Intelligence* **2004**, *32*, 329–347. [CrossRef]
29. Barbuto, J.E.; Scholl, R.W. Motivation sources inventory: Development and validation of new scales to measure an integrative taxonomy of motivation. *Psychol. Rep.* **1998**, *82*, 1011–1022. [CrossRef]
30. Eccles, J.S. Studying the development of learning and task motivation. *Learn. Instr.* **2005**, *15*, 161–171. [CrossRef]
31. Wigfield, A.; Eccles, J.S.; Fredricks, J.A.; Simpkins, S.; Roeser, R.W.; Schiefele, U. Development of achievement motivation and engagement. In *Handbook of Child Psychology and Developmental Science*; Lerner, R.M., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2015; pp. 1–44. ISBN 978-1-118-96341-8.
32. Ryan, R.M.; Deci, E.L. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **2000**, *55*, 68–78. [CrossRef]
33. Stroet, K.; Opdenakker, M.-C.; Minnaert, A. Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educ. Res. Rev.* **2013**, *9*, 65–87. [CrossRef]
34. Spearman, J.; Watt, H.M.G. Perception shapes experience: The influence of actual and perceived classroom environment dimensions on girls' motivations for science. *Learn. Environ. Res.* **2013**, *16*, 217–238. [CrossRef]
35. Deci, E.L.; Ryan, R.M. Motivation, personality, and development within embedded social contexts: An overview of self-determination theory. In *The Oxford Handbook of Human Motivation*; Ryan, R.M., Ed.; Oxford University Press: New York, NY, USA, 2012; pp. 84–108. ISBN 978-0-19-539982-0.
36. Black, A.E.; Deci, E.L. The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Sci. Educ.* **2000**, *84*, 740–756. [CrossRef]
37. Ryan, R.M.; Grolnick, W.S. Origins and pawns in the classroom: Self-report and projective assessments of individual differences in children's perceptions. *J. Personal. Soc. Psychol.* **1986**, *50*, 550–558. [CrossRef]
38. Froiland, J.M.; Davison, M.L.; Worrell, F.C. Aloha teachers: Teacher autonomy support promotes native Hawaiian and Pacific Islander students' motivation, school belonging, course-taking and math achievement. *Soc. Psychol. Educ.* **2016**, *19*, 879–894. [CrossRef]
39. Gutiérrez, M.; Sancho, P.; Galiana, L.; Tomás, J.M. Autonomy support, psychological needs satisfaction, school engagement and academic success: A mediation model. *Univ. Psychol.* **2018**, *17*, 1. [CrossRef]
40. Wei, X.; Wagner, M.; Christiano, E.R.A.; Shattuck, P.; Yu, J.W. Special Education Services Received by Students with Autism Spectrum Disorders from Preschool through High School. *J. Spec. Educ.* **2014**, *48*, 167–179. [CrossRef]
41. Szulawski, M.; Kaźmierczak, I.; Prusik, M. Is self-determination good for your effectiveness? A study of factors which influence performance within self-determination theory. *PLoS ONE* **2021**, *16*, e0256558. [CrossRef]
42. Guay, F.; Ratelle, C.F.; Chanal, J. Optimal learning in optimal contexts: The role of self-determination in education. *Can. Psychol. Psychol. Can.* **2008**, *49*, 233–240. [CrossRef]
43. Patall, E.A.; Steingut, R.R.; Freeman, J.L.; Pituch, K.A.; Vasquez, A.C. Gender disparities in students' motivational experiences in high school science classrooms. *Sci. Educ.* **2018**, *102*, 951–977. [CrossRef]
44. Syzmanowicz, A.; Furnham, A. Gender differences in self-estimates of general, mathematical, spatial and verbal intelligence: Four meta analyses. *Learn. Individ. Differ.* **2011**, *21*, 493–504. [CrossRef]
45. Zander, L.; Höhne, E.; Harms, S.; Pfost, M.; Hornsey, M.J. When grades are high but self-efficacy is low: Unpacking the confidence gap between girls and boys in math. *Front. Psychol.* **2020**, *11*, 552355. [CrossRef]
46. Cvencek, D.; Meltzoff, A.N.; Greenwald, A.G. Math-gender stereotypes in elementary school children. *Child Dev.* **2011**, *82*, 766–779. [CrossRef] [PubMed]
47. Keller, C. Effect of teachers' stereotyping on students' stereotyping of math as a male domain. *J. Soc. Psychol.* **2001**, *141*, 165–173. [CrossRef] [PubMed]
48. Song, J.; Zuo, B.; Wen, F.; Yan, L. Math-gender stereotypes and career intentions: An application of expectancy–value theory. *Br. J. Guid. Couns.* **2017**, *45*, 328–340. [CrossRef]
49. Robinson-Cimpian, J.P.; Lubienski, S.T.; Ganley, C.M.; Copur-Gencturk, Y. Teachers' perceptions of students' math proficiency may exacerbate early gender gaps in achievement. *Dev. Psychol.* **2014**, *50*, 1262–1281. [CrossRef] [PubMed]
50. Appleton, J.J.; Christenson, S.L.; Kim, D.; Reschly, A.L. Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *J. Sch. Psychol.* **2006**, *44*, 427–445. [CrossRef]
51. Barkatsas, A.T.; Kasimatis, K.; Gialamas, V. Learning secondary mathematics with technology: Exploring the complex interrelationship between students' attitudes, engagement, gender and achievement. *Comput. Educ.* **2009**, *52*, 562–570. [CrossRef]

52. Fung, F.; Tan, C.Y.; Chen, G. Student engagement and mathematics achievement: Unraveling main and interactive effects. *Psychol. Sch.* **2018**, *55*, 815–831. [CrossRef]
53. Skinner, E.A.; Furrer, C.; Marchand, G.; Kindermann, T. Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *J. Educ. Psychol.* **2008**, *100*, 765. [CrossRef]
54. Fredricks, J.A.; Blumenfeld, P.C.; Paris, A.H. School engagement: Potential of the concept, state of the evidence. *Rev. Educ. Res.* **2004**, *74*, 59–109. [CrossRef]
55. Greene, B.A. Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educ. Psychol.* **2015**, *50*, 14–30. [CrossRef]
56. Fredricks, J.A.; McColskey, W. The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In *Handbook of Research on Student Engagement*; Christenson, S.L., Reschly, A.L., Wylie, C., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 763–782. ISBN 978-1-4614-2017-0.
57. Cooper, K.S. Eliciting engagement in the high school classroom: A mixed-methods examination of teaching practices. *Am. Educ. Res. J.* **2014**, *51*, 363–402. [CrossRef]
58. Johnson, M.K.; Crosnoe, R.; Elder, G.H. Students' attachment and academic engagement: The role of race and ethnicity. *Sociol. Educ.* **2001**, *74*, 318–340. [CrossRef]
59. Lam, S.; Jimerson, S.; Kikas, E.; Cefai, C.; Veiga, F.H.; Nelson, B.; Hatzichristou, C.; Polychroni, F.; Basnett, J.; Duck, R. Do girls and boys perceive themselves as equally engaged in school? The results of an international study from 12 countries. *J. Sch. Psychol.* **2012**, *50*, 77–94. [CrossRef] [PubMed]
60. Lamote, C.; Speybroeck, S.; Van Den Noortgate, W.; Van Damme, J. Different pathways towards dropout: The role of engagement in early school leaving. *Oxf. Rev. Educ.* **2013**, *39*, 739–760. [CrossRef]
61. Lietaert, S.; Roorda, D.; Laevers, F.; Verschueren, K.; De Fraine, B. The gender gap in student engagement: The role of teachers' autonomy support, structure, and involvement. *Br. J. Educ. Psychol.* **2015**, *85*, 498–518. [CrossRef]
62. Marks, H.M. Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *Am. Educ. Res. J.* **2000**, *37*, 153–184. [CrossRef]
63. Crombie, G.; Pyke, S.W.; Silverthorn, N.; Jones, A.; Piccinin, S. Students' perceptions of their classroom participation and instructor as a function of gender and context. *J. High. Educ.* **2003**, *74*, 51–76. [CrossRef]
64. Reinhold, F.; Strohmaier, A.; Hoch, S.; Reiss, K.; Böheim, R.; Seidel, T. Process Data from Electronic Textbooks Indicate Students' Classroom Engagement. *Learn. Individ. Differ.* **2020**, *83–84*, 101934. [CrossRef]
65. Oakes, L.M.; Kannass, K.N.; Shaddy, D.J. Developmental changes in endogenous control of attention: The role of target familiarity on infants' distraction latency. *Child Dev.* **2002**, *73*, 1644–1655. [CrossRef]
66. Steinmayr, R.; Ziegler, M.; Träuble, B. Do intelligence and sustained attention interact in predicting academic achievement? *Learn. Individ. Differ.* **2010**, *20*, 14–18. [CrossRef]
67. Sarter, M.; Givens, B.; Bruno, J.P. The cognitive neuroscience of sustained attention: Where top-down meets bottom-up. *Brain Res. Rev.* **2001**, *35*, 146–160. [CrossRef]
68. Axelrod, M.I.; Zhe, E.J.; Haugen, K.A.; Klein, J.A. Self-management of on-task homework behavior: A promising strategy for adolescents with attention and behavior problems. *Sch. Psychol. Rev.* **2009**, *38*, 325–333. [CrossRef]
69. Bryan, T.; Burstein, K.; Bryan, J. Students with learning disabilities: Homework problems and promising practices. *Educ. Psychol.* **2001**, *36*, 167–180. [CrossRef]
70. Fan, J.; Wu, Y.; Fossella, J.A.; Posner, M.I. Assessing the heritability of attentional networks. *BMC Neurosci.* **2001**, *2*, 14. [CrossRef] [PubMed]
71. Banz, B.C.; Wu, J.; Crowley, M.J.; Potenza, M.N.; Mayes, L.C. Gender-related differences in inhibitory control and sustained attention among adolescents with prenatal cocaine exposure. *Yale J. Biol. Med.* **2016**, *89*, 143–151.
72. Chan, R.C.K. A further study on the sustained attention response to task (SART): The effect of age, gender and education. *Brain Inj.* **2001**, *15*, 819–829. [CrossRef]
73. Blatter, K.; Graw, P.; Münch, M.; Knoblauch, V.; Wirz-Justice, A.; Cajochen, C. Gender and age differences in psychomotor vigilance performance under differential sleep pressure conditions. *Behav. Brain Res.* **2006**, *168*, 312–317. [CrossRef]
74. Yuan, J.; He, Y.; Qinglin, Z.; Chen, A.; Li, H. Gender differences in behavioral inhibitory control: ERP evidence from a two-choice oddball task. *Psychophysiology* **2008**, *45*, 986–993. [CrossRef]
75. Riley, E.; Okabe, H.; Germine, L.; Wilmer, J.; Esterman, M.; DeGutis, J. Gender differences in sustained attentional control relate to gender inequality across countries. *PLoS ONE* **2016**, *11*, e0165100. [CrossRef]
76. Simpkins, S.D.; Davis-Kean, P.E.; Eccles, J.S. Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Dev. Psychol.* **2006**, *42*, 70–83. [CrossRef]
77. King, R.B. Sense of relatedness boosts engagement, achievement, and well-being: A latent growth model study. *Contemp. Educ. Psychol.* **2015**, *42*, 26–38. [CrossRef]
78. Reyes, M.R.; Brackett, M.A.; Rivers, S.E.; White, M.; Salovey, P. Classroom emotional climate, student engagement, and academic achievement. *J. Educ. Psychol.* **2012**, *104*, 700–712. [CrossRef]
79. Shernoff, D.J.; Csikszentmihalyi, M.; Shneider, B.; Shernoff, E.S. Student engagement in high school classrooms from the perspective of flow theory. *Sch. Psychol. Q.* **2003**, *18*, 158–176. [CrossRef]

80. Chiu, T.K.F. Applying the self-determination theory (SDT) to explain student engagement in online learning during the COVID-19 pandemic. *J. Res. Technol. Educ.* **2022**, *54*, S14–S30. [CrossRef]
81. Park, S.; Holloway, S.D.; Arendtsz, A.; Bempechat, J.; Li, J. What makes students engaged in learning? A time-use study of within-and between-individual predictors of emotional engagement in low-performing high schools. *J. Youth Adolesc.* **2012**, *41*, 390–401. [CrossRef]
82. Reeve, J.; Jang, H.; Carrell, D.; Jeon, S.; Barch, J. Enhancing students' engagement by increasing teachers' autonomy support. *Motiv. Emot.* **2004**, *28*, 147–169. [CrossRef]
83. Carini, R.M.; Kuh, G.D.; Klein, S.P. Student engagement and student learning: Testing the linkages. *Res. High. Educ.* **2006**, *47*, 1–32. [CrossRef]
84. Nguyen, C.-D. Scaffolding student engagement with written corrective feedback: Transforming feedback sessions into learning affordances. *Lang. Teach. Res.* **2021**. [CrossRef]
85. Vansteenkiste, M.; Simons, J.; Lens, W.; Sheldon, K.M.; Deci, E.L. Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *J. Personal. Soc. Psychol.* **2004**, *87*, 246–260. [CrossRef]
86. Reilly, D.; Neumann, D.L.; Andrews, G. Sex differences in math and science achievement: A meta-analysis of National Assessment of Educational Progress assessments. *J. Educ. Psychol.* **2015**, *107*, 645–662. [CrossRef]
87. Reinhold, F.; Hofer, S.; Berkowitz, M.; Strohmaier, A.; Scheuerer, S.; Loch, F.; Vogel-Heuser, B.; Reiss, K. The Role of Spatial, Verbal, Numerical, and General Reasoning Abilities in Complex Word Problem Solving for Young Female and Male Adults. *Math. Educ. Res. J.* **2020**. [CrossRef]
88. Dias, M.A.; Carvalho, P.S.; Ventura, D.R. How to Study the Doppler Effect with Audacity Software. *Phys. Educ.* **2016**, *51*, 035002. [CrossRef]
89. Hofer, S.I.; Reinhold, F.; Koch, M. Students Home Alone—Profiles of Internal and External Conditions Associated with Mathematics Learning from Home. *Eur. J. Psychol. Educ.* **2022**. [CrossRef]
90. Wang, M.T.; Fredricks, J.A.; Ye, F.; Hofkens, T.L.; Linn, J.S. The math and science engagement scales: Scale development, validation, and psychometric properties. *Learn. Instr.* **2016**, *43*, 16–26. [CrossRef]
91. Prenzel, M. Mehrdimensionale Bildungsziele Im Mathematikunterricht Und Ihr Zusammenhang Mit Den Basisdimensionen Der Unterrichtsqualität. Multi-Dimens. *Educ. Goals Math. Classr. Their Relatsh. Instr. Qual.* **2016**, *44*, 211–225. [CrossRef]
92. Koch, M.; Möller, C.; Spinath, F.M. Are You Swiping, or Just Marking? Exploring the Feasibility of Psychological Testing on Mobile Devices. *Psychol. Test Assess. Model.* **2021**, *63*, 507–524.
93. Rosseel, Y. Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *J. Stat. Softw.* **2012**, *48*, 1–36. [CrossRef]
94. Satorra, A.; Bentler, P.M. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* **2001**, *66*, 507–514. [CrossRef]
95. Wei, D.; Zhang, D.; He, J.; Bobis, J. The impact of perceived teachers' autonomy support on students' mathematics achievement: Evidences based on latent growth curve modelling. *Eur. J. Psychol. Educ.* **2019**, *35*, 703–725. [CrossRef]
96. Meece, J.L.; Jones, M.G. Gender differences in motivation and strategy use in science: Are girls rote learners? *J. Res. Sci. Teach.* **1996**, *33*, 393–406. [CrossRef]
97. Metallidou, P.; Vlachou, A. Motivational beliefs, cognitive engagement, and achievement in language and math in elementary school children. *Int. J. Psychol.* **2007**, *42*, 2–15. [CrossRef] [PubMed]
98. Ruffing, S.; Wach, F.-S.; Spinath, F.M.; Brünken, R.; Karbach, J. Learning strategies and general cognitive ability as predictors of gender-specific academic achievement. *Front. Psychol.* **2015**, *6*, 1238. [CrossRef] [PubMed]
99. Hofer, S.I.; Stern, E. Underachievement in Physics: When Intelligent Girls Fail. *Learn. Individ. Differ.* **2016**, *51*, 119–131. [CrossRef]
100. Hofer, S.I. Studying Gender Bias in Physics Grading: The Role of Teaching Experience and Country. *Int. J. Sci. Educ.* **2015**, *37*, 2879–2905. [CrossRef]
101. McCullough, L. Women in physics: A review. *Phys. Teach.* **2002**, *40*, 86–91. [CrossRef]
102. Meece, J.L.; Glienke, B.B.; Burg, S. Gender and motivation. *J. Sch. Psychol.* **2006**, *44*, 351–373. [CrossRef]
103. Taasoobshirazi, G.; Carr, M. Gender differences in science: An expertise perspective. *Educ. Psychol. Rev.* **2008**, *20*, 149–169. [CrossRef]
104. Frenzel, A.C.; Pekrun, R.; Goetz, T. Girls and mathematics—A "hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *Eur. J. Psychol. Educ.* **2007**, *22*, 497–514. [CrossRef]
105. Goetz, T.; Bieg, M.; Lüdtke, O.; Pekrun, R.; Hall, N.C. Do girls really experience more anxiety in mathematics? *Psychol. Sci.* **2013**, *24*, 2079–2087. [CrossRef]
106. Lohbeck, A.; Grube, D.; Moschner, B. Academic self-concept and causal attributions for success and failure amongst elementary school children. *Int. J. Early Years Educ.* **2017**, *25*, 190–203. [CrossRef]
107. Mok, M.M.C.; Kennedy, K.J.; Moore, P.J. Academic attribution of secondary students: Gender, year level and achievement level. *Educ. Psychol.* **2011**, *31*, 87–104. [CrossRef]
108. Lesperance, K.; Hofer, S.; Retelsdorf, J.; Holzberger, D. Reducing Gender Differences in Student Motivational-affective Factors: A Meta-analysis of School-based Interventions. *Br. J. Educ. Psychol.* **2022**, bjep.12512. [CrossRef] [PubMed]

109. Ziegler, A.; Stoeger, H. Evaluation of an attributional retraining (modeling technique) to reduce gender differences in chemistry instruction. *High Abil. Stud.* **2010**, *15*, 63–83. [CrossRef]
110. Hernán, M.; Hernández-Díaz, S.; Robins, J. A structural approach to selection bias. *Epidemiology* **2004**, *15*, 615–625. [CrossRef] [PubMed]
111. Smart, R.G. Subject selection bias in psychological research. *Can. Psychol. Psychol. Can.* **1966**, *7*, 115–121. [CrossRef]

education sciences

MDPI

Article

Career Profiles of University Students: How STEM Students Distinguish Regarding Interests, Prestige and Sextype

Divan Mouton [1,*], Florian G. Hartmann [2] and Bernhard Ertl [1]

[1] Learning and Teaching with Media, Institute of Education, Department of Human Sciences, Universität der Bundeswehr München, 85579 Neubiberg, Germany; bernhard.ertl@unibw.de

[2] Division of Psychological Assessment, Department of Psychology, Paris Lodron Universität Salzburg, 5020 Salzburg, Austria; florian.hartmann@plus.ac.at

* Correspondence: divan.mouton@unibw.de

Abstract: Various factors may be important while individuals develop vocational aspirations. Although occupations that fit one's personal interests appear to be attractive, contextual factors may repel groups, such as young girls, to develop towards areas such as STEM. Especially, the sextype of STEM occupations, that is often considered as male, could limit STEM career choice of young girls. This study investigates career profiles of n = 9277 German university freshmen based on interests, prestige, and sextype. Eleven latent profiles were found. Five profiles can be characterized by their prestige levels and two further by their sextype. Certain profiles are significantly associated with study outcomes and study satisfaction, which allows to identify at-risk profiles. Of note, especially female students in STEM subjects with a low proportion of females distributed widely across the 11 profiles. The implications of this study suggest that career choice profiles according to Gottfredson's framework can help to identify at-risk students.

Keywords: occupational aspirations; career profiles; latent profile analysis; STEM; higher education

Citation: Mouton, D.; Hartmann, F.G.; Ertl, B. Career Profiles of University Students: How STEM Students Distinguish Regarding Interests, Prestige and Sextype. *Educ. Sci.* **2023**, *13*, 324. https://doi.org/10.3390/educsci13030324

Academic Editor: James Albright

Received: 8 February 2023
Revised: 9 March 2023
Accepted: 18 March 2023
Published: 22 March 2023

1. Introduction

Although an individual's career decision may depend on different factors, several theories emphasize the impact of an individual's interests in this context, e.g., [1–3]. Often, concepts such as expectancies, values, and prior experiences are further considered for explaining career decisions, e.g., [4,5]. Such theories primarily focus on the individual and implicitly describe the "pull" factors that are important for the decision for developing *towards* a specific career path. If we, however, consider females' decisions for or against STEM (science, technology, engineering, and mathematics), we see that many of them *rule out* such study areas because of occupational aspects, especially because they consider it as too male [5]. Thus, the *sextype* of occupations, as pointed out by Gottfredson (1981), seems to be an important obstacle for female students that prevents them from developing towards STEM careers. Another factor working simultaneously may be the *prestige* of an occupation: it may be considered as too low and thereby serve as an obstacle for students choosing an occupation, although more in the context of male students' decision against social occupations rather than in STEM. Such occupational aspects, that often serve as "repel" factors that set boundaries for individuals for not going into a specific career, received less attention in research, besides in research building on Gottfredson's [6] theory that considers interests, prestige, and sextype as dimensions for career decisions. However, these boundary conditions that repel individuals from some occupational areas may not affect all students similarly: several females choose a STEM career path despite such occupations being considered as "male". This means that there are groups of individuals that may rather follow their interests while ignoring the sextype of an occupation (or perceiving the sextype differently). In other words, Gottfredson's boundary dimensions do not have the same effect on career choices for everyone. Rather, there are different groups of individuals who deal differently with repellent factors in the sense of Gottfredson. This paper aims to focus such groups and therefore distinguishes latent profiles according to the dimensions proposed by Gottfredson [6]: interest, prestige, and sextype. It focuses on university freshmen as they just made the first step from a rather general education in school towards a narrower and more occupation-specific one at university. After identifying latent profiles, the paper analyzes how far a respective profile membership can be characterized by background variables at study entry, either by study subject choice or individual orientations for getting a better idea of the students in the respective profiles. In a third step, the study focuses on study outcomes of the students in the different profiles to validate how far students that find themselves in at-risk profiles according to [1,2] and how far they show disadvantages with respect to study outcomes and satisfaction. Using data from a large-scale panel study allows to longitudinally compare students' outcomes and to focus on six different fields of study for comparison, including STEM fields with a medium (STEM-M) and a low proportion of females (STEM-L). Of particular interest is the question of the extent to which female STEM students form their own profile or are distributed across different profiles.

1.1. Gottfredson's Theory and Related Constructs

A student's vocational interests are an important predictor of their later vocational behaviour and, according to Holland's [2] person–environment fit theory, students should aspire to an occupation that matches their interests. Holland's recommendation is supported by a huge body of research that indicates favourable outcomes for interest congruence (fit) in relation to satisfaction, performance, and persistence [7–11]. Vocational interests, however, develop through the repeated interactions of the person and their occupational perceptions of the environment, including their perceptions about their outcomes, with their eventual performance or achievements [3,4]. Gottfredson's [6,12,13] theory postulated two perceptions of an occupational environment that are fundamental to the development of interest, namely prestige and sextype.

Gottfredson's Theory. Gottfredson's [6,12,13] developmental theory of occupational aspirations expands on [2] person–environment fit by positing that, along with their field of interest, students also base their career choices on their aspired occupation's level of prestige and its perceived masculine–feminine traits, known as sextype. Both dimensions have shown evidence of shaping student's efficacy expectations, their interest development, and their career choices, both in the early stages of development and shortly before a career decision is required [14,15]. These three strongly interrelated dimensions [16] are emphasized in the latter of the four-stage process of *circumscription* which emphasizes the development of both a person's self-concept and their cognitive map of occupations and causes them ultimately to develop an occupational aspiration by successively eliminating occupations that they do not find compatible. While children in the first stage (3–5 years) simply recognize what an occupation is, the second stage (6–8 years) emphasizes the morally central aspects of *sex roles* and produces rigid vocational preferences for same-sex adult behavior and the rejection of cross-sextyped activities, depending on their development. In the third stage (9–13 years), young students make *social evaluations* of both unacceptably low status occupations as well as high prestige careers and the amount of effort attaining one would require. Finally, as adolescents (stage four, 14+ years), students explore their *internal and unique self* to a greater degree in order to identify what types of work are compatible to their field of interest. By the end of the circumscription process, a student's occupational choice is grounded within its *social space,* which is a bounded area on the cognitive map of occupations according to two axes—prestige *level* (high vs. low) and *sextype* (masculinity vs. femininity)—within which a person can discover where their *interests* fit into the world of work. Once an individual is inclined towards going to a specific occupation, then the reality whether this occupation is accessible or not comes to the fore. This outlines strategic and sometimes difficult choices during a *compromise* process which are, according to Gottfredson [13], based on the field of interest, prestige level, and sextype of each career alternative. The following outlines each of these dimensions in greater depth.

Field of Interest. Gottfredson [6] (p. 548) made parallels with her assessment of compatibility and Holland's [2] person–environment fit (P-E fit) called interest congruence, although she notes the indirect nature of such kind of measure. Holland [2] defines vocational interests as trait-like characteristics that can best be described using six different interest dimensions (RIASEC): *Realistic* interests, *Investigative* interests, *Artistic* interests, *Social* interests, *Enterprising* interests, and *Conventional* interests. The RIASEC types can be used to describe the interest profiles of individuals and their academic or occupational environments. As a result, a match can be determined by comparing the individual and environmental RIASEC profiles. Holland's [2] *congruence hypothesis* claims that the fit between a person and his or her environment predicts academic and occupational outcomes such as performance and persistence [9–11]. Holland [2] proposes a hexagonal structure of the six interest types and this arrangement is basis for many algorithms for calculating congruence. Although this hexagonal structure has been shown to be particularly evident in U.S. samples [17], it can be taken as given and gender invariant for German university samples [18]. Appropriately, RIASEC interests have been and still are extensively investigated in Germany (e.g., [19–21]).

Prestige. Although definitions differ, prestige, or social status, is a complex and multifaceted concept associated with an occupation's socioeconomic status, level of education, difficultly, responsibility [17], as well as the aspects of the effort and skill required from the occupation [22]. Despite its complex nature, people in developed countries have remarkably similar interpretations of prestige [13] (p. 91), evidenced by various standardized prestige and social status scales that strongly correlate [23,24]. In principle, an occupation with a higher prestige is regarded as more desirable due to the social recognition and financial rewards that they provide, but they also may require a greater amount of effort and ability to attain [13], which is an aspect that is also taken up by rational choice theory [25]. A less challenging but still desirable aspect to prestige is to minimize the risk of downward social mobility by, at minimum, aspiring to return to the same status as their parents [26]. The choice of university degree has shown to be greatly influenced by parents, particularly their socio-economic status, which also significantly interacts with the student's gender [27].

Sextype. The perceived gender identity of an occupation or study area influences young children's perceptions about sex-typical vocational behaviour [13]. It is, however, noteworthy to point out that distinguishing and measuring the sextype of a wide range of occupations is less straightforward than it seems. Technology-related subjects such as engineering, computer sciences, and physics are obviously perceived as prototypically male areas [5] that persistently present a traditional gender gap [28]. They show a strong over-representation of males with usually more than three males for every one female, and for some engineering areas, the ratio is even as low as six to one [29]. Consequently, such areas are widely *perceived* as having a male gender identity that is also coherent with their male-over-represented sex *proportion*. Even though some STEM areas present balanced sex proportions (between 30% and 70% female students), such as mathematics, chemistry, and geography, they are still perceived as belonging to a male stereotyped science area [30]. Study programs in the life sciences have progressively shifted to a proportion of female students around 70% or more [31]. These observable shifts in sex proportion in the life sciences over time may also be reshaping societal perceptions of sextype in related areas. In this line, the proportion of females in technical areas raises if these include an aspect of life sciences. For example, typical computer science studies present around an 18% proportion of female students, but medical and bioinformatics present 46% and 48% female proportions, respectively. This effect, however, may interact with student's attitudes towards stereotypically traditional gender roles, whether masculine or feminine, or more balanced egalitarian roles [32]. Nonetheless, career choices appear to be particularly affected if the sex proportion in the occupation exceeds the 30-70% balance [33], which in our study specifically relates to specific STEM-related study clusters with a low female proportion [19].

Of note, Gottfredson [13] discussed different configurations of these three dimensions. Furthermore, such configurations imply that there may be different manifest or latent profiles of students' career choices. Since people do not always have the opportunity to choose an occupation that perfectly fits with regard to all three dimensions, they may have to compromise regarding one or more occupational characteristics. According to Gottfredson [12], a severe threat of the 'right' sextype is related to higher costs than a severe threat to the desired level of prestige or to the fulfillment of one's interests. However, previous studies revealed mixed results regarding the compromise behavior and the weighting of the three dimensions. For example, Armstrong und Crombie [34] confirmed Gottfredson's compromise hypothesis, while the study by Hesketh et al. [35] indicates that, under compromise condition, interests were more important for the attractiveness of a job than prestige or sextype. In fact, and of course, there are women who enter typically male domains. One reason for this may be that there are some individuals who weight the fulfillment of their interests or a prestigious position more strongly than the 'right' sextype when choosing a career.

1.2. A Person-Centered Approach to Career Choice

An appropriate method to reveal groups of individuals that deal differently with the dimensions of sextype, prestige, and interests when aspiring a career is a *person-centered approach*, such as a latent profile analysis. A person-centered approach distinguishes itself from a *variable-centered approach* which suggests that all members of a population show a similar configuration regarding a set of variables, usually applied in factor analyses or structural equation modelling [36]. In the context of our study, a variable-centered approach would assume that the whole population of students would be more concerned about, e.g., their gender-identity than about their prestige or interests. Such approaches correspond to some previous studies who found results that rank prestige as more important in conditions where it was not theoretically expected, e.g., [37–39]. This is different when a *person-centered approach* is used. Here, the assumption of a homogenous population is relaxed by suggesting that there can be different subpopulations of students with distinct profiles [40,41]. As a result, a person-centered approach could allow the possibility that some students are less concerned about the gender typicality and more concerned with a higher prestige occupation, while other groups could prioritize interests over gender typicality and/or prestige.

We found two studies that applied a person-centered approach to examine career choices based on Gottfredson's theory [42,43]. Both studies found four types of career profiles; however, neither study assessed an empirical measure of vocational interest. Furthermore, Ryu and Jeong [43], also excluded sextype from their investigation because of contextual reasons. Our study expands this work by including all three of Gottfredson's career choice dimensions within a large-scale sample with the goal of analyzing the extent to which different subpopulations of students differentially prioritize the three career choice dimensions in the context of their study decision.

While there is a lack of person-centered research, this study does contribute to previous research that aimed at investigating career choices using vocational interests, especially regarding fields with unbalanced gender distributions (e.g., [19,44–48]).

1.3. Characteristics of Study Outcomes Related to Gottfredson's Dimensions

Focusing on the *interest* dimension, students with a poor person–environment fit should show lower persistence, performance, and satisfaction according to Holland [2] and the supporting literature [8–11]; they may therefore be considered as at-risk students. If these findings are extrapolated onto a sample of university freshmen, this would imply that latent profiles that are characterized by a notably lower orientation to their field of interest may show poorer study outcomes such as lower completion rates, lower grades, less satisfaction with their studies [7], and furthermore higher intentions to dropout.

Students characterized by a *prestige* orientation may be characterized by a greater importance for status maintenance [26] to reach the same prestige level their parents already have. As discussed in [13,22,25], there is a relationship between effort and prestige. Students who aspire to prestigious careers may therefore experience consequences related to high effort requirements, such as challenges with study load or higher chances of failure. Prestige is also already reflected in different study subjects that can be characterized by differences in prestige such as occupations [49]; they may inherently distinguish the prestige of some of the latent profiles.

Finally, profiles that indicate a high *sextype* orientation may be characterized by a strongly traditional gender role attitude as opposed to an egalitarian one [50]. As sextype-oriented profiles may indicate a high level of compromise [1], they may show indications of lower study outcomes.

2. Research Questions

As summarized above, previous research on Gottfredson's [6,12,13] career choice dimensions revealed mixed results: while some results suggest that people weigh matching sextype most heavily in career choices, other findings suggest that interests may play a more important role. In general, mixed results can be an indicator that different subpopulations exist. In the current study, this would mean that there are different groups of students who use different career choice strategies that may be related to different study outcomes. Therefore, the current study aims to investigate the following research questions:

RQ1: How Far Can Different Profiles of Career Choice Based on Gottfredson's Dimensions Be Distinguished within a Population of University Freshmen?

To examine RQ1, we use a person-centered approach, which is exploratory in nature and does not allow for the formulation of precise hypotheses. However, this approach still leaves some room for theoretically framed expectations [40,51]. For example, there may be students who do not prioritize their interests in order to pursue a prestigious or gender-typical career [6,12,13], while other students may focus on finding a job that fit their interests [2].

While the latent profile may reflect Gottfredson's career choice dimensions, it is not a given that they will be distinctive in relation to aspects relevant at study entry. For making the latent profiles better identifiable, research question 2 analyzes how far further variables important at study entry may characterize the latent profiles:

RQ2: How Far Are Different Career Profiles Related to Background and Early Study Variables?

Due to the exploratory nature of *RQ1,* and an a priori unknown characteristic about the latent profiles, our study cannot make any explicit hypotheses regarding what different background or early study variables might characterize a specific latent profile. We can, however, derive assumptions for specific profile characteristics and therefore assume students with a higher aspirational prestige to expect poorer study outcomes because of the greater efforts that will likely be required to attain a high prestige occupation ([13]; assumption A. 1). In addition, we assume students in profiles with a higher prestige to find it more important to maintain their parents' status than those with a lower prestige ([26]; A. 2). For students that aspire to occupations with a lower prestige, we would also assume a lower rating of the chances for getting a good job (A. 3) according to rational choice theory [25].

In the framework of RQ2, it is examined how students from different study subjects distribute across different profiles. Of special interest in the current study is the question to what extent female STEM-L students constitute own profiles or in which profiles they can be found. In light of Gottfredson's theory, female STEM-L students can be seen as a group of individuals who overcome the 'wrong' sextype, which may be grounded in especially strong interests in STEM fields and/or a particularly pronounced striving for a prestigious position. This may either lead to a distinct latent profile or these females are intermixed with individuals of other latent profiles.

RQ2 already raises aspects such as the study outcome expectation at study entry that may characterize several profiles. RQ3 now takes a longitudinal perspective and analyzes study outcomes and study satisfaction longitudinally to validate how far profiles that show pattern of at-risk students are indeed characterized by lower outcomes.

RQ3: How Far Are the Different Career Profiles Related to Different Aspects of Study Outcomes?

Several meta-studies indicate that interest congruence predicts favourable outcomes such as satisfaction ([8]; A. 4), performance ([9–11]; A. 5), and persistence ([9–11]; A. 6), which suggests that profiles that are characterized by interest congruence will likely have better study outcomes. Regarding students who aspire a gender-atypical career, because it is more congruent to his or her interests, there is, however, the question if they are more likely to successfully complete their degree [9] or if they are affected by phenomena that incur with having an untypical sextype ([1]; A. 7a/b). Regarding prestige-characterized profiles, we assume students with high prestige to have poorer study outcomes due to increased difficulty required to attain their aspired occupation ([6,17]; A. 8). Furthermore, we assume students who aspire to high same-sex occupations to show more traditional gender role attitudes while those who aspire to low same-sex occupations are more egalitarian ([50]; A. 9).

3. Method

3.1. Sample

A sample of $n = 9277$ German university students from the National Education Panel Study (NEPS; SC5:15.0.0; [52]) was analyzed. Their age was between 18 and 27 years at study entry with a mean of 20.1 years ($sd = 1.8$). These students started in winter semester 2010/2011 and were from six study clusters, namely STEM-L, STEM-M, education, language, medicine, and economics [19]. Moreover, 61% of the 9277 students were female. It is noteworthy to point out that NEPS oversampled teaching education students [53] (p. 13), and therefore the data set comprises of an oversample of female students.

3.2. Measures

As NEPS is a longitudinal panel study, the measures may come from different time points. Variables for generating the latent profiles (RQ1) and for characterizing profiles (RQ2) were surveyed directly at study entry, while validation variables (RQ3) were surveyed later in the course of study. For RQ3, episode data of the first study program in which a student was enrolled were analyzed with respect to its study outcome. If a student, e.g., enrolled for electrical engineering, the episode data indicate if the student finished this program successfully as well as the final grade of this program.

3.2.1. Variables for Defining the Latent Profiles

The three career choice dimensions were generated in relation to each student's occupational aspiration at the study entry (wave 1). Occupational aspirations are acquired by NEPS by asking the open-ended question: "*Regardless of how your degree course is going, what job would you most like to have at some point?*". The open text answers were codified by NEPS into different occupational classification systems, such as ISCO-08 [54] and KldB2010 [55], and furthermore classified according to standardized prestige scales such as the Standard International Occupational Prestige Scale (SIOPS, [56]) and the Magnitude Prestige Scale (MPS, [24]).

Interest *congruence* builds on Holland's [2] RIASEC dimensions, as measured using the 18-item adult version of the Interest Inventory Life Span (IILS-II; [57]), and the O*net [58] interest classification of the student's occupational aspiration. A congruency measure based on the Euclidean distance between the individual and the occupational interest vectors was generated as described in [19]. It should be noted that shorter vectors indicate higher congruence.

Occupational ***Prestige*** of students' aspirations is already classified in NEPS, and this study uses the classification using the German-based Magnitude Prestige Scale (MPS; [24]).

Sextype is estimated by the Same-Sex Proportion (SSP), which is generated using the proportion of each sex employed within their aspired occupation as classified within a taxonomy of occupations, similar to Beavis' [59] concept of "sex composition". Therefore, the 5-digit German system to classify occupations [55] is used to match female proportions from the Federal Statistical Office [60] onto the occupational aspirations available in NEPS. The study specifically applies SSP from the year 2010 as it corresponds with the start of university for the NEPS SC5 cohort.

3.2.2. Variables for Characterizing the Latent Profiles

Two types of variables are used to characterize profile membership:

1. *Demographic variables*, consisting of ***sex*** and six ***study clusters*** (namely STEM-L, STEM-M, education, language, medicine, and economics; [19], as well as information about whether their studies are ***teaching orientated*** [61].
2. *Background variables available at study entry*; these include ***study outcome expectations*** (wave 1), ***chances for getting a good job*** (wave 1), and the ***importance of status maintenance of parents*** (wave 1).

The descriptive statistics, exemplary items, as well as references for the respective scales can be found in Supplementary Table S1.

3.2.3. Variables for Validating the Latent Profiles

For validating how far latent profile membership provides higher risks for a student, his or her study outcomes were analyzed longitudinally. This comprises of variables that were surveyed in later waves of the panel as well as episode data that refer to the outcomes of the first study program for which a student enrolled.

1. *Longitudinal variables that were available in later waves* include the ***intention to dropout*** (wave 2), three dimensions of ***study satisfaction*** (wave 3; related to study content, study conditions, and coping with study burdens), and ***gender role attitudes*** (wave 4). Their descriptive statistics, exemplary items, and references for the respective scales can also be found in Supplementary Table S1.
2. *Study episode data*. The NEPS dataset provides information about study episodes [53]. For example, a student can indicate the start and end dates of any study episode, e.g., a start date for his or her initial bachelor studies. Such episodes are sometimes open, which means without an end date, e.g., if a student is part of the panel attrition group.

Study Outcomes will focus on the *successful completion* or the *failure* of the first study episode up to eight years after the start, specifically up to wave 14. A very small number of students indicated that they ended their studies in the first term, or that they explicitly indicated that they did not finish their studies. Due to the marginal number of such cases, these two outcomes were not observed in the analysis.

Final Grades. If students successfully reach the end of a study episode then a final grade for this episode is usually provided. To ensure comparability, grades are z-standardized within each study subject. Our study therefore compares the z-scores for students who successfully completed their first study episode in their respective subject, i.e., engineering. It should be noted that the German grading system indicates better grades as lower values.

3.3. Analyses and Procedure

3.3.1. Description of Profiles: Latent Profile Analysis

A latent profile analysis (LPA) is a person-centered approach that assumes that it is possible to identity distinct latent subgroups based on their shared scores on certain indicators [36]. This study refers to the analysis as an LPA and not a latent class analysis (LCA) because of the continuous nature of the indicator variables [36] (p. 7). The three career choice dimensions, namely the (1) interest congruence, (2) occupational prestige, and (3) SSP within the aspired occupation, are used as input variables for the LPA. For calculating the LPA, the three input variables were z-standardized. In addition, interest congruence was inverted so that higher values indicate a higher congruence. Of note, all the charts and descriptions of the LPA in this paper use the raw values of the dimensions to allow readers to have a better estimation of the profile characteristics.

The latent profile analysis was run using MPlus version 8.2 [62] to derive alternative latent profile solutions ranging from three to twenty profiles (see Supplementary Table S2). The choice of latent profile configurations is guided by several statistical characteristics, especially by the replication of the best log-likelihood value in combination with testing improvements of the model fit over the *k*-1 model using the Vuong–Lo–Mendell–Rubin likelihood ratio test (LRT) and the bootstrapped likelihood ratio test (BLRT) [63]. In addition to statistical comparisons of alternative latent profile configurations, a review of each latent profile configuration helped to decide on the final solution. Based on the gender distribution in the sample, we furthermore validated the results of the gender inclusive latent profile analysis with two further latent profiles analyses that were separated for male and female students and comprised the same number of latent profiles like the gender inclusive one.

3.3.2. Characterization and Validation of the Profiles

Demographic and background variables, listed in Section 3.2.2, are used to explore assumptions that characterize members of a specific career profile. Career profiles are validated using variables listed in Section 3.2.3. Although some of the validation variables have a substantial percentage of missing responses due to the longitudinal panel attrition, their inclusion provides valuable evidence of each profile's characteristics. Means and confidence intervals were used to analyze continuous variables for significant differences for all the validation variables except for successful and failed study outcomes, which are analyzed using a chi-square test.

4. Results

Before answering the research questions, we will first refer to the descriptive statistics (see Supplementary Tables S1 and S3–S5) and the correlations between the variables. A correlation matrix amongst interest congruence, prestige, and female proportions in their aspired occupations (adapted to generate the SSP), as well as other characterizing variables, can be found in Supplementary Table S3. The three career choice dimensions present significant but small correlations with one another. Furthermore, the sample distribution by study cluster and sex can be found in Supplementary Table S6.

4.1. Latent Profile Analysis

As a result of the LPA, an eleven-profile solution was chosen because of the LRT based on the *k*-1 model test, a satisfactory entropy, and a sound replication of the best log-likelihood value that indicates a stable solution (see Supplementary Table S2). Upon further inspection, the eleven-profile solution remained superior based on its well-differentiated profiles and evenly distributed sample.

To further validate the eleven career profiles, this study investigates whether LPAs using sex-separated samples, which are set to find the same number of profiles (eleven) in order to correspond with the latent profiles from the whole sample, come to similar results. Crosstabulations reveal that most profiles share 100% of their cases, while three female and four male profiles share upwards of 90% of their cases with whole sample profiles. Only one profile from either sex, related to sextype-characterized profiles, shares around 80% of their cases with a corresponding profile from the whole sample. This suggests that other than the sextype characterized profiles, the majority of the eleven profiles could be interpreted without undue concern for sex differences (See Supplementary Tables S7–S9).

Due to the comparably high number of profiles and the large sample sizes that allow even small differences to be significant, and because furthermore the LPA aims at revealing homogenous classes (with small standard deviations) that lead to high effect sizes of the profile differences, we added a gray range around the sample mean to indicate less distinguished profiles for a better interpretation of the differences in Figure 1. For congruence and prestige, the gray area indicates values within +/− 1 SD of the sample mean (congruence 0.49/1.14; prestige 163/109). For SSP, the gray neutral area is defined theoretically [33] as a male/female sex proportion between 30 and 70% which centers around a balanced 50% value [33]. Of note, one must keep in mind that the SSP of an occupation and the proportion of females in a profile are different measures. Each profile's mean score was used to characterize it as either low or high if it falls outside of the gray range (see Supplementary Figure S1 for a zoomed-in version). The following section describes the latent profiles by the most characterizing aspects; more detailed information can be found in Supplementary Table S10. All profiles were within one standard deviation of the mean with respect to congruence.

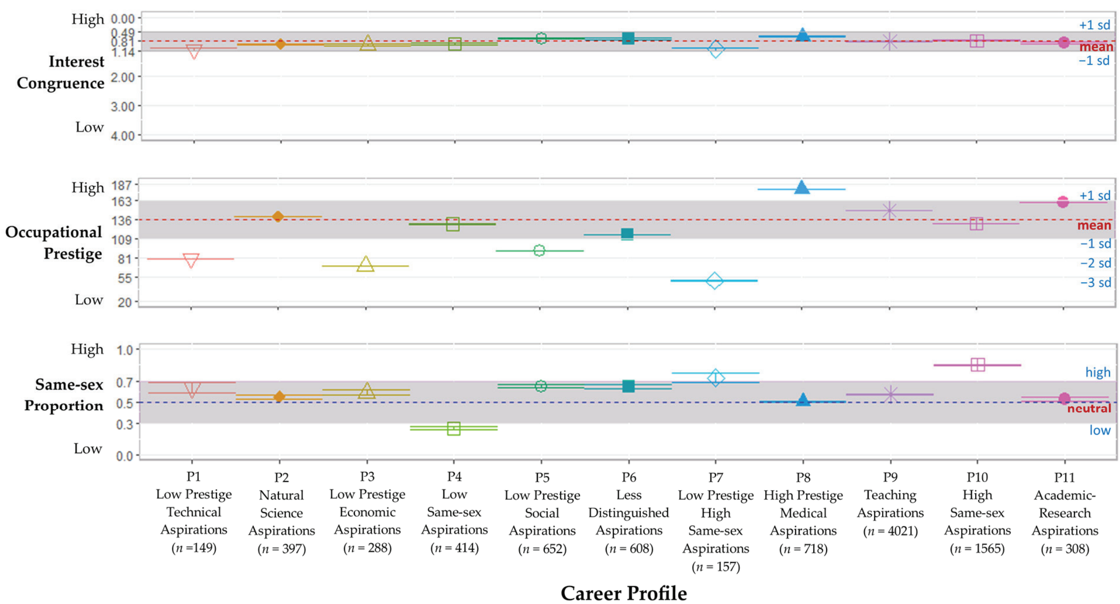


Figure 1. The means and confidence intervals of career profiles across the three career choice dimensions. *Note*: n = 9277. Each dimension is displayed in its full and unstandardized scale. Congruence ranges from 0 (perfect) to ~4 (poor), prestige ranges from ~20 (low) to ~187 (high), SSP in the aspired occupation ranges from 0 (low) to 1 (high).

The largest profile (P9; *Teaching Aspirations*; n = 4021) accounts for 43% of the sample and consists of 78% female students. The means of all three career choice dimensions (interests, prestige, and sextype) are within the less distinguishable gray range (see Figure 1 and Supplementary Table S10 for the values; see Supplementary Tables S11–S13 for effect sizes). Looking into students' occupational aspirations, we found that more than 80% of P9 profile aspire to become teachers (see Supplementary Table S14).

4.1.1. Prestige Characterized Profiles

While all profiles were within one standard deviation from the sample mean with respect to interest congruence, we see five profiles that are clearly distinguished by their prestige: one with a distinctly high prestige, one with a distinctly low prestige, and three further profiles that can also be distinguished by their low prestige but require further distinguishing criterion.

The *P8* profile (*High Prestige Medical Aspirations*; *n* = 718, 8%) consists of 68% females. Besides their high prestige, P8 also present above average congruence and a neutral SSP. Participants of this profile mainly aspire to work in medical occupations.

In contrast, students of the *P7* profile (*Low Prestige High Same-Sex Proportion Aspirations*; *n* = 157; 2%) are characterized by a very low prestige. The profile consists of 26% female students and is further characterized by a high SSP and a congruence that is clearly below average. As their SSP factor suggests, these students aspire to gender-typical professions. Male students mainly aspire to become machine operators and work in mechatronics, while the females aim towards business organization and strategy supervisors.

The *P1* profile (*Low Prestige Technical Aspirations*; *n* = 149, 2%) is one of the three low prestige profiles which need to be distinguished in more detail. Although their prestige is quite low, it is not as low as P7. It consists of 38% female students and is further characterized by an interest congruence clearly below average and an SSP that tends to be high. This profile mainly aspires toward technical professions either in the domain of engineering or machine operations for males, or medical technicians and other skilled technical professions for females.

The *P3* profile (*Low Prestige Economic Aspirations*; *n* = 288; 3%) consists of 58% female students. The SSP of this group is slightly tending towards high and the interest congruence slightly below average. This profile aspires to rather economics-oriented occupations such as media designing and advertising, technical design, banking, and business administration.

The *P5* profile (*Low Prestige Social Aspirations*; *n* = 652; 7%) consists of 73% female students. The SSP of this group's aspiration is clearly towards the higher range and their interest congruence slightly above average. Participants of this group mainly aspire to become social workers.

4.1.2. SSP Characterized Profiles

Besides these, the LPA also revealed two SSP characterized profiles.

The *P4* profile (*Low Same-sex Proportion Aspirations*; *n* = 414, 4%) consists of 88% female students. Both their prestige and congruence are slightly below average, but their highly atypical SSP distinguishes their aspirations. Female students mainly aspire to work in technical and economic occupations such as construction and business consulting, while male students aspire to become primary school teachers.

The *P10* profile (*High Same-Sex Proportion Aspirations*; *n* = 1565; 17%) consists of 31% females. The congruence of this group's participants is slightly above average while their prestige slightly below. In this profile, female students mainly aspire to become primary school teachers, while males aspire to technical occupations such as machine building and operation.

4.1.3. Profiles with Less Distinguishable Career Choice Dimensions

Like the *P9* profile, the remaining profiles are within the gray range on all three career choice dimensions and therefore are only distinguished by their profile's main occupational aspirations.

The *P2* profile (*Natural Science Aspirations*; *n* = 397; 4%) consists of 53% female students. Both males and female students of this profile mainly aspire to occupations related to the natural sciences such as chemistry, physics, and biology.

The *P11* profile (*Academic/Research Aspirations*; *n* = 308; 3%) consists of 50% female students. Participants of this profile mainly aspire to become university researchers/lecturers, with some aspiring to be dentists or veterinarians.

The *P6* profile (*Less Distinguished Aspirations*; *n* = 608, 7%) consists of 43% female students. The aspirations are partially of two career paths related to computer sciences and language. Males mainly aspire to computer sciences-related occupations with a small portion aspiring to journalism, while the females aspire to work in language-related occupations such as journalism, copy editors, and translators and a small portion aspire to become computer scientists.

4.2. Characterization of the Profiles

The demographic and background variables that characterize the profiles include gender, study cluster, teaching orientation, as well as the importance of maintaining their parent's status. In addition, two outcome-related characterizing variables are also included, namely study outcome expectations and chances of getting a good job.

4.2.1. Sex

The chi-square test found a significant and moderate effect size for sex in association with the career profiles (χ^2 (10) = 1301.216, $p < 0.001$, Contingency Coeff = 0.351, Cramer's V = 0.375). Specifically, males were significantly associated with six profiles, namely P10 (*High SSP*), P6 (*Less Distinguished*), P7 (*Low Prestige High SSP*), P1 *(Low Prestige Technical),* P11 (*Academic/Research*), and P2 (*Natural Science*), in order of the highest significant standardized residuals (see Supplementary Figure S2 and Supplementary Table S15). Conversely, females were significantly associated with P9 (*Teaching*), P4 (*Low SSP*), P5 (*Low Prestige Social*), and P8 (*High Prestige Medical*) profiles.

4.2.2. Study Clusters

The chi-square test results indicate a significant association between study clusters and career profiles (χ^2 (50) = 12,267.61, $p < 0.001$). Due to the larger number of levels involved in the analysis, an effect size cannot be calculated. In addition, several cells have less than five cases per cell. This, however, does not violate the expected frequency assumption because the minimum expected frequency of 9.96 well exceeds the recommended number [64] (p. 935). The results suggest that students in STEM-L are significantly and positively associated with several profiles including P10 (*High SSP*), P7 (*Low Prestige High SSP*), P6 (*Less Distinguished*), P1 *(Low Prestige Technical),* P4 (*Low SSP*), P3 (*Low Prestige Economic*), P11 (*Academic/Research*), and P2 (*Natural Science*), while STEM-M students are significantly associated with P2 (*Natural Science*), P9 (*Teaching*), and P11 (*Academic/Research*). The results further indicated that medicine students are particularly associated with P8 (*High Prestige Medical*) and to a lesser extent P11 (*Academic/Research*), some of whom aspire to be dentists (see Figure 2 and Supplementary Table S16). Economics students are significantly associated with P3 (*Low Prestige Economic*), P4 (*Low SSP*), and P5 (*Low Prestige Social*), while education students are significantly associated with P5 (*Low Prestige Social*) and P9 (*Teaching*). Finally, language students are significantly associated with P9 (*Teaching*).

4.2.3. Study Orientation Teaching

The chi-square test indicates a strong and significant effect size for teaching vs. non-teaching orientation in association with the different career profiles (χ^2 (10) = 4635.636, $p < 0.001$. Contingency Coeff = 0.578, Cramer's V = 0.707). Students that are teaching orientated are significantly associated with the P9 (*Teaching*) profile which consists of 43% of the sample (see Supplementary Table S17).

4.2.4. Outcome Expectation at Study Entry

Two profiles indicate better study outcome expectations, namely P8 (*High Prestige Medical*) and P11 (*Academic/Research*), while, by contrast, P1 *(Low Prestige Technical)*, P3 (*Low Prestige Economic*), and P7 (*Low Prestige High Same-Sex*) indicate lower study outcome expectations that are significantly below the sample mean (see Figure 3a, Supplementary Figure S3a, and Supplementary Table S18).

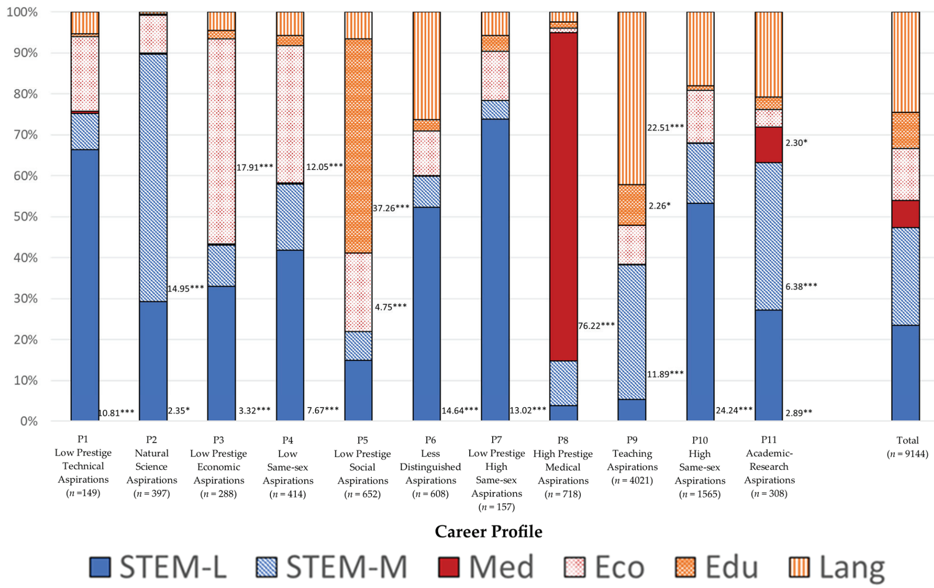


Figure 2. Bar graph of the study cluster distribution amongst career profiles with significance based of chi-square test z-scores. *Note*: $n = 9277$; χ^2 (50) = 12,267.61, $p < 0.001$. Minimum expected frequency: 9.957961; positive and significant standardized residuals, measured in z-scores, are indicated adjacent to the corresponding group. Significant standardized residuals are indicated as * $p < 0.05$ if $z > \pm 1.96$; ** $p < 0.01$ if $z > \pm 2.58$; *** $p < 0.001$ if $z > \pm 3.29$.STEM-L = STEM studies with a female proportion less than 30%; STEM-M = STEM studies with a sex proportion between 30 and 70%; Med = medicine; Eco = economics; Edu = education; Lang = language.

4.2.5. Chances for Getting a Good Job

Most notably, P8 (*High Prestige Medical*) indicates significantly and above average career prospects from studying their degree, followed by P1 *(Low Prestige Technical)*, P3 (*Low Prestige Economic*), P7 (*Low Prestige High SSP*), P10 (*High SSP*), and P11 (*Academic/Research*), who all indicate chances that are significantly above the sample mean. By contrast, P5 (*Low Prestige Social*), P9 (*Teaching*), and P6 (*Less Distinguished*) rate their career prospects as significantly below that of the sample mean (see Figure 3b, Supplementary Figure S3b, and Supplementary Table S18).

4.2.6. Status Maintenance of Parents

The results indicated that P2 (*Natural Science*), P3 (*Low Prestige Economic*), P4 (*Low SSP*), P6 (*Less Distinguished*), and P11 (*Academic/Research*) care more about maintaining the status of their parents than the sample mean, while P8 (*High Prestige Medical*) and P9 (*Teaching*) take a view of maintaining their parents' status that is below that of the sample mean (see Figure 3c, Supplementary Figure S3c and Supplementary Table S18).

4.3. Validation of the Profiles

In the following, we investigate the profiles in relation to variables that indicate study outcomes longitudinally and that therefore allow to estimate how far a specific profile could be considered as an at-risk profile. Therefore, we will first focus on the study outcomes from episode data (Figure 4) before we investigate variables of later survey waves that may already have been indicators for success and failure (Figure 5; see Supplementary Figure S4 for a zoomed-in version of Figure 5 and Supplementary Table S19 for means and standard deviations of each validation variable). For a comprehensive comparison of the effect sizes for each validation variable amongst the career profiles, see Supplementary Tables S20–S25. For an overview of the significance of both characterization and validation variables across the career profiles, see Table 1.

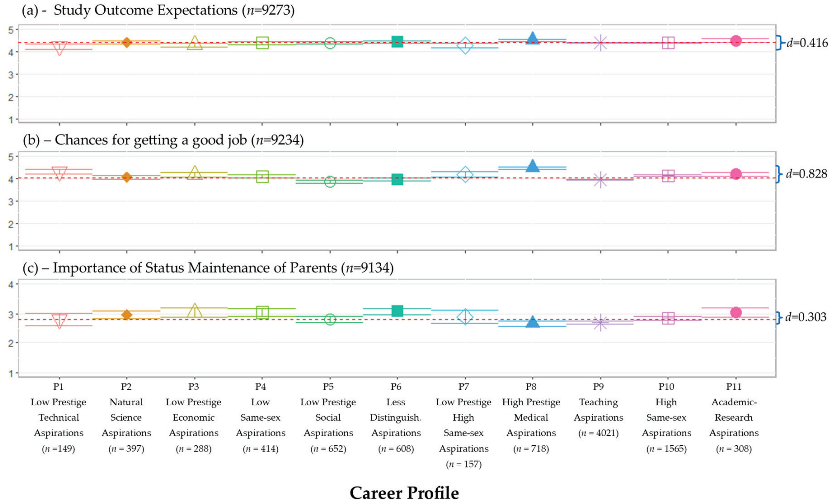


Figure 3. Variables for the characterization of the career profiles (full-scale version). *Note*: sample sizes vary according to availability in the NEPS dataset. Sample means are illustrated as the dotted red line. The effect size (Cohen's *d*) is illustrating the range between the highest and the lowest mean score across the career profiles for each variable. For the zoomed-in version, see Supplementary Figure S3.

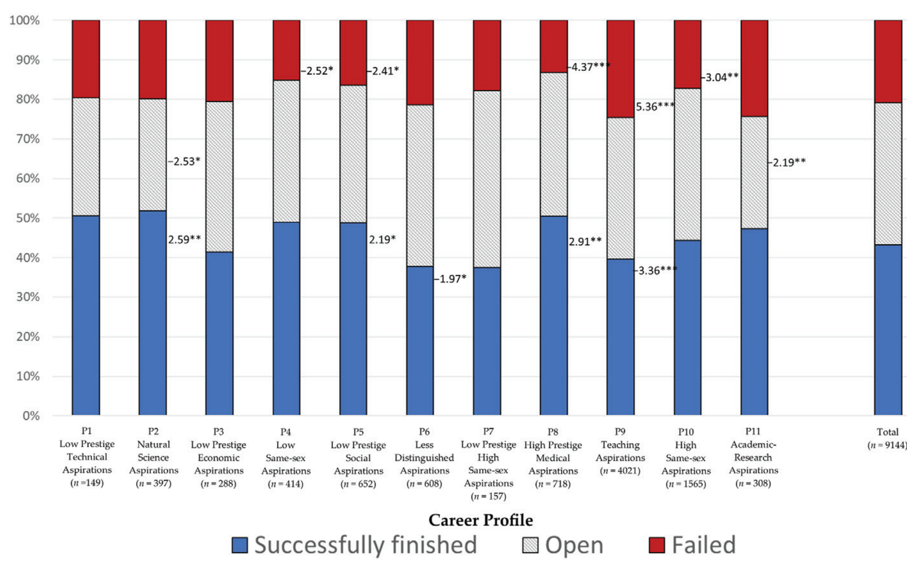


Figure 4. Study outcomes of career profiles with significance based on chi-square test z-scores (n = 9144). Note: χ^2 (20) = 138.297, $p < 0.001$. Contingency Coeff = 0.122, Cramer's V = 0.087; * $p < 0.05$ if $z > +1.96$; ** $p < 0.01$ if $z > +2.58$; *** $p < 0.001$ if $z > +3.29$.

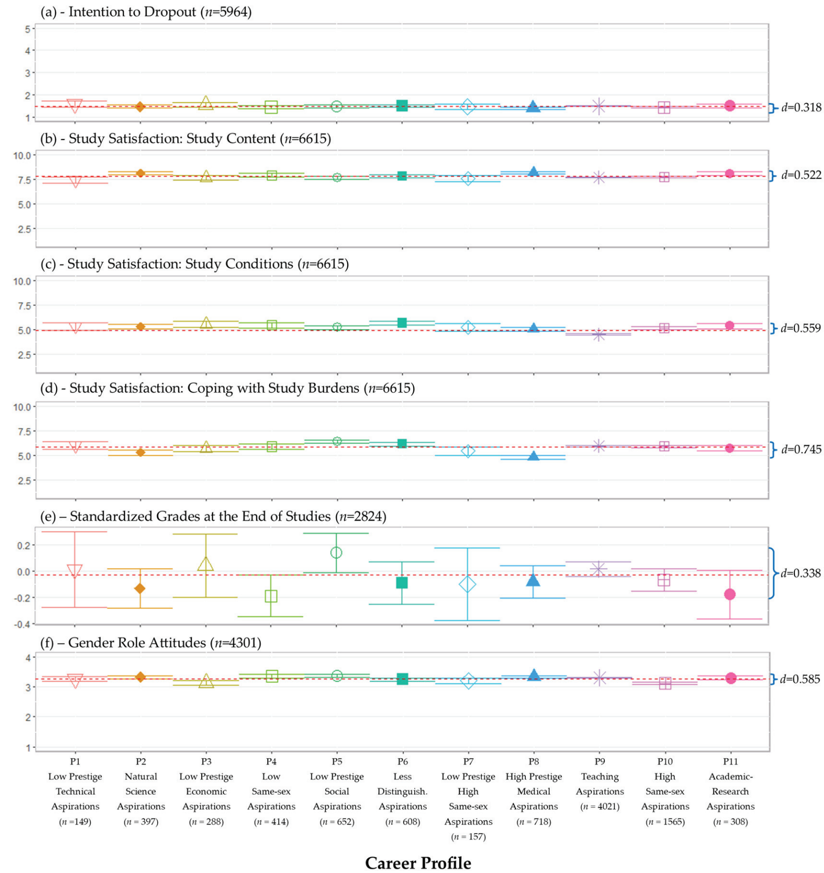


Figure 5. Variables for the validation of the career profiles (full-scale version). *Note:* sample sizes vary according to availability in the NEPS dataset. Sample means are illustrated as the dotted red line. The effect size (Cohen's *d*) is illustrated by the range between the highest and the lowest mean score across the career profiles for each variable. For (5**e**), lower values indicate better grades. For (5**f**), lower values indicate traditional gender role attitudes while higher values indicate egalitarian gender role attitudes (see Supplementary Figure S4 for a zoomed-in version).

Regarding ***study outcomes*** (see Figure 4; Supplementary Table S26 for the chi-square test contingency table), the chi-square analysis presents a significant χ^2 (20) = 138.297, $p < 0.001$, but a small effect size (Contingency Coeff = 0.122, Cramer's V = 0.087), suggesting that the results should be cautiously interpreted due to weak associations. Nonetheless, significant differences in study outcomes were observed.

In the following, we will present the most salient results of Figures 4 and 5 as well as results that relate to our assumptions regarding the research questions.

The *Teaching Aspirations* (P9) profile is both more likely to fail and less likely to successfully finish their degree, which indicates poor study outcomes (see Figure 4). This is reflected in the satisfaction with the study conditions value that is observably the lowest while other variables were located around the sample mean (see Figure 5c and Supplementary Figure S4c). The second profile less likely to reach a successful study outcome is P6 (*Less Distinguished*) that, however, indicates a high satisfaction with study conditions and a good coping with study burden.

Table 1. Summary of Career Profiles across Characterization and Validation Variables.

	Demographic Variables			*Background Variables at Study Entry*			*Longitudinal Variables in Later Waves*					*Study Episode Data*	
	Sex *	**Study Cluster ***	**Teach. Orient. ***	**Study Outcome Expect.**	**Chances for Getting a Good Job**	**Status Maint. of Parents**	**Intention to Dropout**	**Study Content Satisfact.**	**Study Cond. Satisfact.**	**Coping w/Burdens Satisfact.**	**Gender Role Attitudes**	**Final Study Grades**	**Study Outcomes *A**
P1—Low Prestige Technical	Male	STEM-L	No	Below	Above	-	-	-	-	-	-	-	-
P2—Natural Science	Male	STEM-L STEM-M	No	-	-	Above	-	Above	Above	Below	-	-	+Success
P3—Low Prestige Economic	-	STEM-L ECO	No	Below	Above	Above	-	-	Above	-	Trad.	-	-
P4—Low Same-sex	Female	STEM-L ECO	No	-	-	Above	-	-	Above	-	Egal.	-	-Fail
P5—Low Prestige Social	Female	ECO EDU	No	-	Below	-	-	-	Above	Above	Egal.	Above	+Success -Fail
P6—Less Distinguish.	Male	STEM-L	No	-	Below	Above	-	-	Above	Above	-	-	-Fail
P7—Low Prestige High Same-sex	Male	STEM-L	No	Below	Above	-	-	-	-	-	-	-	-
P8—High Prestige Medical	Female	MED	No	Above	Clearly Above	Below	Below	Above	-	Clearly Below	Egal.	-	+Success -Fail
P9—Teaching	Female	STEM-M EDU LANG	Yes	-	Below	Below		Below	Below	Above	Egal.	-	-Success +Fail
P10—High Same-sex	Male	STEM-L	No	-	Above	-	-	-	Above	-	Trad.	-	-Fail
P11—Academic-Research	Male	STEM-L STEM-M MED	No	Above	Above	Above	-	Above	Above	-	-	-	-

Note: * = Variables analysed using chi square tests are only summarized by significant associations. A = Study outcomes also indicate the direction of the association using a '+' for increased and a '-' for a decreased association. Above/Below = Mean scores are indicated as either significantly above or below the sample mean scores. Clearly Above/Below = Mean scores that are indicated as either significantly above or below any other profile's confidence interval. Gender role attitudes are indicated as either significantly traditional (Trad.) or egalitarian (Egal.). Teach. Orient. = Teaching orientation. Study outcome expect. = Study outcome expectations. Status maint. of parents = Importance of status maintenance with parents. Study content satisfact. = Satisfaction with study content. Study cond. satisfact. = Satisfaction with study conditions. Coping w/ burdens. satisfact. = Satisfaction related to coping with study burdens. * STEM-L = STEM studies with a female proportion less than 30%; STEM-M = STEM studies with a sex proportion between 30–70%, MED = Medicine, ECO = Economics, EDU = Education, LANG = Language. For references regarding significance see the following tables or figures: "Sex" = Supplementary Table S15; "Study Cluster" = Supplementary Table S16; "Teaching Orientation" = Supplementary Table S17; "Study outcome expectations", "Chances for getting a good job"," Importance of status maintenance with parents" = Supplementary Table S18 & Supplementary Figure S3; " Intention to dropout", "Satisfaction with study content", "Satisfaction with study conditions", "Satisfaction related to coping with study burdens", " Final study grades" = Supplementary Table S19 & Supplementary Figure S4; " Study Outcomes" = Supplementary Table S26.

If we now look to the prestige-characterized profiles, students in the *High Prestige Medical* (P8) are less likely to fail and more likely to successfully finish their studies. Although they show a low intention to dropout and a high satisfaction with the study content, this profile is characterized by the worst coping with the study conditions. This is contrasted by the *Low Prestige Social* (P5) profile that has the same study outcome pattern but shows the best coping with the study conditions, although with the lowest grades at the end of the study. The other three prestige-characterized profiles do not show significant effects regarding study outcomes, although P1 (*Low Prestige Technical*) shows the highest intention to drop out and the lowest satisfaction with the study content. P3 (*Low Prestige Economic*) shows the second highest intention to dropout but with a quite high satisfaction with the study conditions, and P7 (*Low Prestige High SSP*) also one of the lower coping strategies.

Both sextype characterized profiles (P4-*Low SSP* and P10-*High SSP*) are significantly less likely to fail their studies, with P4 showing comparably the best grades at the end of the study.

Students of the *Natural Science* (P2) profile are significantly more likely to successfully finish their studies and less likely to have open study outcomes; they show a high satisfaction with their study content but are the second lowest group with respect to coping with their study burdens. Students of P11 (*Academic/Research*) also are less likely to show open study outcomes. They are comparably satisfied with the study content and are receiving comparably good grades.

Finally, we look at gender roles attitudes. While we conceptually would have preferred to discuss gender role attitudes as an early background variable, this was not possible due to its collection in the later waves. Its inclusion can indicate if students of a profile are more or less inclined to traditional gender roles. Like Figure 5f shows, students with *High SSP* (*P10*) and *Low Prestige Economic* (*P3*) profiles indicate more traditional gender role attitudes while the *Low SSP* (*P4*), the *Low Prestige Social* (*P5*), and *High Prestige Medical* (*P8*) all indicate higher egalitarian gender role attitudes.

5. Summary and Discussion

5.1. Distinguishing Latent Profiles according to Gottfredson's Theory

Reflecting on Gottfredson's theory [6,12,13] and looking overarchingly onto the latent profiles that we found, we can hardly see profiles that are distinguished by their **interests**. Although there were significant differences between the profiles on the interest dimension, all of them centered +/− 1 SD from the sample mean on a relatively high level. This indicates that the individual interests were an important factor for their career choices and that a major compromise [6,12,13], i.e., compromising interests for a higher sextype, did not take place within this sample of students. Of note, students already possess the highest level of school leaving degree and therefore may have not been in need to compromise their interests so much. Additionally, the **prestige** of the occupations that students aspired was relatively high, but here, we could see one profile with an even higher prestige (P8, High Prestige Medical) and four profiles with a clearly lower prestige. We want to emphasize that all low prestige profiles either were related to a lower study outcome expectation (P1 Low Prestige Technical, P3 Low Prestige Economic, and P7 Low Prestige High SSP), a lower estimation of their chances for getting a good job (P5 Low Prestige Social), or a higher intention to dropout (P1 Low Prestige Technical, P3 Low Prestige Economic), which indicates that students within these profiles made some moderate compromises [6,12,13] and might therefore be at-risk students, especially students of P7 (Low Prestige High SSP) who also showed lower values regarding coping with study burdens (Table 1 gives an overview on the respective profile characteristics). Regarding **sextype**, most profiles were within a neutral range between 30 and 70% SSP with one profile (P4 Low SSP) indicating a lower SSP and two profiles, P7 (Low Prestige High SSP) and P10 (High SSP), indicating a higher SSP. While P4 (Low SSP) students showed a quite positive pattern including better grades, students of P7 (Low Prestige High SSP) were already characterized as at-risk students with a moderate level of compromise.

5.2. Identifying At-Risk Students

Combining these observations with the results about study outcomes, we see that two of the potential at-risk profiles (P3 Low Prestige Economic, P7 Low Prestige High SSP) show, although not significant, below average study success rates, with students of P7 (Low Prestige High SSP) showing the lowest proportion of successfully finished study episodes. This highlights how Gottfredson's dimensions and the estimated compromises [6,12,13] can point towards at-risk students that, however, must be validated in a broader context as only two of the four profiles actually show lower study outcomes. Furthermore, two profiles, P6 (Less Distinguished) and P9 (Teaching), also show lower success rates, although they are inconspicuous with respect to Gottfredson's dimensions. For them, a lower estimation of their chances for getting a good job (for both P6 Less Distinguished and P9 Teaching) together with a low satisfaction with the study conditions (P9 Teaching) rather point to effects of expectancy value theories [4] and the socio-cognitive career theory [3] to explain their study outcomes. Consequently, Gottfredson's [6,12,13] concept of compromise can help to identify some groups of at-risk students, but several groups of at-risk students cannot be discovered by compromise pattern, at least in this study.

5.3. Occupational Characteristics

Moving now the focus from the Gottfredson dimensions to the underlying occupational aspirations of the latent profiles, we see that most profiles can be characterized by quite narrow ranges of prestige and SSP of these occupations. Some profiles comprised of students who mostly aspire to the same occupation, e.g., P9 (Teaching) with respect to teaching profiles (about 80%), or P8 (High Prestige Medical) regarding medical professions (also over 80% when including specializations). Furthermore, three of the four low prestige profiles can be distinguished by broad occupational categories, e.g., P1 (Low Prestige Technical), P3 (Low Prestige Economic), and P5 (Low Prestige Social). While such distinction seems theoretically as well as empirically meaningful, it points towards framing conditions of the underlying concepts of Gottfredson's theory [6,12,13]: interest, prestige, and sextype, all three which relate to the same occupation or occupational aspiration. All three have the challenge of measurement with the respective consequences. Using an individual's self-rated estimation of prestige and sextype may be biased and therefore unreliable, especially for individuals who aspire to low prestige occupations, because low socioeconomic status individuals show evidence of systematically rating low prestige occupations as higher when compared to individuals from a higher status background [65] (p. 270). Using, in contrast, a standardized scale of prestige, such as the MPS [24], and the proportions of males and females in an occupation for sextype, e.g., according to official labor statistics data [29], may cause deviations between the calculated and the perceived occurrence. We already discussed this with respect to science domains with a balanced SSP that is still often perceived as male ([30]; see also the introductory part). However, this phenomenon may also apply to life science subjects such as medicine and biology. Although they usually are less perceived as male domains, the corresponding scientist jobs such as medical doctor or biologist may still have a male connotation. This may result from media portrayals, and we can exemplify that by the coverage of virology, epidemiology, and infection biology researchers in the beginning of the COVID-19 pandemic in Germany (https://www.spiegel.de/gesundheit/corona-virus-christian-drosten-ist-nummer-eins-bei-medienpraesenz-von-virologen-a-e3d97148-06db-4b9d-bb5d-511543f7cf43 (accessed on 19 March 2023)): Although the proportion of female students in this area is about 70%, the top five researchers that were mentioned in press reports about COVID-19 were male and their coverage had a proportion of about 85% of all reports. Furthermore, just 2 female researchers could be found in the top 10 (on position 6 and 9) with a coverage of 4% and 2%, respectively. This biased media coverage may show impacts on the perceived sextype of occupations and the respective tolerable boundary [6,12,13]. Such phenomena may also explain mixed results from previous research that arose when researchers tried to test Gottfredson's theory with respect to compromise [37–39]. Of note, the proportion of female students that earned a university access degree raised dramatically during the last decades [66], and consequently, the proportion of females in high prestige occupations is changing. Thus, one must generally acknowledge interaction between occupations and prestige and sextype, also for areas outside STEM.

5.4. Teaching Aspirations

The largest profile (P9, Teaching, 43%) is characterized by teaching aspirations with students of almost each of the six study areas. Although this profile was inconspicuous with respect to most additional analyses, students of this profile showed the lowest satisfaction with their study conditions, a significantly lower success and a higher failure rate, lower estimation of their chances for getting a good job, and a lower importance of maintaining their parents' status. Focusing subject areas, the teaching profile comprised of comparably few STEM and especially few STEM-L students, which highlights the STEM teacher congruence dilemma [61] that suggests that recruiting teachers in STEM is difficult because the type of interest profiles that STEM teachers would benefit from (i.e., high social for teaching and high realistic for STEM) are rare in society.

5.5. Female STEM-L Students

In the theoretical part, we raised the question whether females in STEM-L constitute their own profile since they have to overcome the 'wrong' sextype, which may be grounded in especially strong interests (i.e., a high congruence) and/or a particularly pronounced striving for prestigious occupations. However, the results indicate that women in STEM-L are relatively broadly distributed across the 11 profiles and that the largest proportion of female STEM-L students (36%) are assigned to profile 4 (see Supplementary Table S6). This profile is characterized neither by particularly high congruence nor by particularly high prestige. In addition, 13% of female STEM-L students can be found in P9 (Teaching), while the other half of them is scattered across all profiles with more than 5% of them being present in 8 out of 11 profiles. This finding may indicate that female students in STEM-L, although studying a subject area with a low SSP, rather prefer occupations with a more balanced SSP, which may refer to the tolerable boundary for sextype according to Gottfredson [6]. Considering furthermore that the Low SSP (P4) profile is comprised of a small number of male students of any subject area, while more than 5% of the economics and more than 35% of the STEM-L students in this profile are female, we can support Gottfredson's [13] (p. 105) observation that "women are more willing to perform cross-sextyped work than are men", i.e., sextype is less a barrier for female students than it is for male students.

5.6. Further Characteristics of the Study Areas

In contrast to the P9 (Teaching) profile that includes students of all study areas (besides medicine), most profiles show rather clear study area characteristics (see Supplementary Table S6). This is most obvious for medical students as most fall into P8 (High Prestige Medical) with a small number, between 5% and 10%, of female students going into P11 (Academic/Research). All other profiles were just populated marginally (<5%) by medicine students. This is similar for students of education who, besides being present in P9 (Teaching), could mainly be found in P5 (Low Prestige Social), for students of the languages who could be found in P6 (Less Distinguished) besides in P9 (Teaching) with a considerable number of female students also being in P10 (High SSP), and for STEM-M with students being present in P2 (Natural Science) besides P9 (Teaching) with several students also being present in P10 (High SSP). As with STEM-L, we see quite scattered distributions across the profiles for economics. The P11 (Academic/Research) profile is a sparsely populated one which may reflect the small numbers of students who intend to stay in academia. Such sparse populations can also be found in three of the four low prestige profiles P1 (Low Prestige Technical), P3 (Low Prestige Economics), and P7(Low Prestige High SSP), which indicates that university students usually aspire higher prestige occupations, but this is different for P5 (Low Prestige Social), which may indicate that social aspirations usually show lower prestige levels.

5.7. Evaluation of our Assumptions

After providing a general view on our results, we will now focus on the assumptions that we stated in the context of our research questions. We raised several assumptions about **interest characterized profiles**, especially a higher satisfaction (A. 4), performance (A. 5), and persistence (A. 6). The profiles in this study, however, showed quite similar levels of interest congruence and can hardly be defined by an outstanding interest congruence, although there are significant differences between some of them. Comparing now the three profiles with the highest congruence level (P8, High Prestige Medical; P5, Low Prestige Social; P6, Less Distinguished) and the two with the lowest (P1, Low Prestige Technical; P7, Low Prestige High SSP), we see a higher satisfaction with the study contents for P8 and lower satisfactions for P1 and P7, all with at least medium effect sizes, while for P5 and P6, the results are less distinguishable. Regarding satisfaction with study conditions, all profiles stand out against P9 (Teaching), while there are only small and indifferent effects with respect to the other profiles. Moreover, for coping with study burden, we see that the high interest profile P8 (High Prestige Medical) shows worse coping, while the other two high interest profile (P5, Low Prestige Social; P6, Less Distinguished) show better coping. Consequently, we just have vague evidence for assumption 4 and this is primarily with respect to the satisfaction study content—this, however, seems to fit the connection to interests. Regarding performance (A. 5), we do not find evidence for assumption 5 with the high interest P5 (Low Prestige Social) showing the worst grades of all profiles. Regarding persistence (A. 6), we see the high congruence profiles (P8, High Prestige Medical; P5, Low Prestige Social) with higher success and lower failure rates, but the third high interest profile (P6, Less Distinguished) showed lower success rates similarly to P7 (Low Prestige High SSP). P1 (Low Prestige Technical) was inconspicuous. Thus, for the interest related profiles, these assumptions cannot be confirmed convincingly. Considering previous work of [7,9] that found effects of interest congruence with variable centered approaches, we like to emphasize the observation that all profiles share a narrow range of interests, and the effects of interest congruence may be moderated by prestige- and sextype-related effects that were much more distinguishable in our latent profile analysis. In this context, we want to point towards findings on differences in interest congruence of different subject areas like in [19] and thus we assume that the inconsistencies regarding the assumptions 4–6 may further result of subject- and profile-specific differences.

Regarding **prestige profiles,** assumptions 1, 3, and 8 proposed poorer outcomes of students in high prestige profiles with A. 1 targeting students' outcome expectations, A. 3 targeting the chances for getting a good job, and A. 8 targeting study success and failure. If we now look at profiles with the higher prestige, we just see a higher outcome expectation for P11 (Academic/Research) and P8 (High Prestige Medical), while for the low prestige profiles P1 (Low Prestige Technical), P3 (Low Prestige Economic) and P7 (Low Prestige High SSP), we see a lower study outcome expectation. If we would discuss this finding in the context of Gottfredson's theory about tolerable effort boundaries [6,13], we would rather have expected students in the lower prestige profiles to show a comparable or higher outcome expectation. Focusing on these results from the perspective of expectancy–value theories [4], however, we would argue that students with a lower outcome expectation aspire to lower prestige occupations. Thus, we must emphasize that expectancy–value theory [4] seems to be more appropriate to model an outcome–expectation–prestige relation than Gottfredson's theory about effort boundaries [6] like we proposed in A. 1. Regarding the chances of getting a good job (A. 3), the profile with the highest prestige (P8, High Prestige Medical) and the one low prestige profile (P5, Low Prestige Social) were able to support this assumption. The other low prestige profiles (P1, Low Prestige Technical, P3, Low Prestige Economic, and P7, Low Prestige High SSP) also presented above average ratings on their chances of getting a good job, while a profile with an above average prestige occupation such as teaching rated their chances as below the sample mean. These mixed results may point to job-specific characteristics which none of the theories that we have discussed are able to coherently explain [4,6,13,25]. Focusing on A. 8, we see that the high prestige profile P8 (High Prestige Medical) shows higher chances of success as well as lower chances of failure, however combined with the lowest value for coping with study burden that might indicate the high efforts required for highly prestigious jobs according to Gottfredson [6]. It should be noted that studying medicine in Germany requires excellent grades in the school leaving examination, and it may be possible that a considerable effort is required to attain those grades as well as for passing exams in the medicine studies which can be experienced as a heavy burden. This could provide partial evidence for A. 8. Regarding our assumption in A. 2, that students in profiles with higher prestige emphasize the importance of status maintenance, we found a diverse set of prestige profiles that found maintaining their parent's status as important, while the profile who did not find maintaining their parent's status as important were of a high prestige i.e., P8 (High Prestige Medical). Of note, students of P8 (High Prestige Medical) already have parents in high prestige professions (see Supplementary Figure S5 and Supplementary Table S27). This gives the impression that status maintenance is not a factor that is directly related to aspirational prestige.

Finally, regarding the **sextype characterized profiles**, we discussed how far this is associated with a higher or lower chance for successful study outcomes (A. 7a/b) and gender role attitudes (A. 9). Here, we see that the high (P10, High SSP) and low SSP (P4, Low SSP) aspiration profiles show significant lower rates of failing. We could assume that the effects of an atypical career choice relate to (A. 7a), although they are not above average in the congruence dimension, while for the typical career choice, the "right sextype" factor may apply (A. 7b). Finally, we can confirm A. 9 with Low SSP profiles (P4, Low SSP) comprising of more egalitarian gender role attitudes than the High SSP Profiles (P7, Low prestige High SSP; P10 High SSP). Of note, also P5 (Low prestige Social) shows a rather egalitarian gender role attitude, while P3 (Low prestige Economic) rather a traditional one.

5.8. Limitations

We already discussed several limitations of the study above, such as that interest, prestige, and sextype relate to the same aspiration and therefore somehow depend on each other, which is most obvious for the sextype–prestige combinations. This dependence is a side condition of Gottfredson's theory [1] and especially applies if the prestige and sextype scores are estimated from established scales [23,24] or labor market data [60]. Sextype was originally measured as a self-reported masculine–feminine rating of occupations [6], but this is neither available in our data nor conceptually practical to generate for such a large set of occupations like for the ISCO-08 that comprises of 436 Unit Groups [54]. As already discussed before, using ratings of the participants may rather provide further bias than better estimations for inter-individual comparisons, especially as sex proportions in occupations were changing over time.

While Gottfredson [1] was introducing the concept of the sextype of an occupation as male and female, this sextype per se is less a factor than the relation between the sextype and the sex of the individual. Thus, sextype-oriented vocational choices mean occupations with a high same sex proportion. This implicated for the latent profile analysis to include either the sextype variable in a mixed sex sample that might inform about male and female occupations, to split the sample into males and females for the profile analysis, which would reveal profiles for male and female students, or to use the SSP for the profile analysis. We consider the SSP variable as conceptually similar to Gottfredson's [6] original sextype variable and the respective consequences that it produces similar to previous research [16,59]. However, we also provided a comparison to split male/female LPAs for comparison that, however, revealed quite similar profiles for most of the students.

Another limitation relates to the generalizability of the results outside of Germany. Since previous studies found evidence that the structural validity of the RIASEC interests may depend on the cultural context of an investigation [17], the results of the current study may not be readily transferable to other countries or cultures. This may be reinforced by cultural differences in the prestige of occupations or in (the strength of) gender stereotypes (see, e.g., [67,68]).

A fourth limitation relates to the large sample of teachers in the study. Although NEPS intentionally oversampled teachers [53], the large teacher profile P9 (43% of the sample) indicates challenges for the prestige/sextype analyses of Gottfredson [1]: as both are quite homogeneous for the teachers in this profile, this may have suppressed differences in the interest congruence that were found, e.g., by [61]. Considering phenomena such as the teacher congruence dilemma for STEM teachers, it becomes clear that much greater differentiation between and within occupations may be desirable, but this may provide obstacles for valid estimations of prestige and sextype of such distinguished occupations, especially regarding secondary education teachers who are distinguished by just one category in classifications such as the ISCO-08 [54].

Finally, the study is subject to all phenomena that incur with large-scale panel studies such as panel attrition and missing values in variables of later waves like those that are discussed in [69].

6. Conclusions

The focus of our study was on career profiles of university freshmen, and we applied Gottfredson's [6,12,13] dimensions to create latent profiles. The analysis disclosed several profiles that were distinguished by prestige and/or sextype and some of them could also be classified as at-risk profiles. Yet, the analysis showed that many profiles also represent a latent structure of occupational clusters. Consequently, we see that students from several study areas, especially medicine, education, and languages, only distribute to few profiles. This is different for students of economics and **STEM** and especially for female students in STEM-L that distribute over 8 of the 11 profiles with more than 5% of the subsample size each. Furthermore, we can observe that only about a third of the female STEM-L students can be found in the low SSP profile and almost two thirds divert to other profiles. This is different for male students of whom more than half stays in the respective two high SSP profiles. Of note, only a marginal number of male students (smaller than 5% of each subject sample) was found in the Low SSP (P4) profile. Thus, regarding the topic of the special issue "Sticking with STEM: Who Comes, Who Stays, Who Goes, and Why?" we can conclude that female students in STEM-L, although studying a subject with a male sextype according to Gottfredson [6,12,13], rather seem to estimate this sextype as inappropriate and turn away towards occupations with a more appropriate sextype. Previous studies have made various conclusions which suggest that a good fit in STEM is particularly relevant (e.g., [19,44–48]), which leads us to hypothesize that STEM-L studies might just be seen as a vehicle for female students to come to a more sex-balanced profession.

The latter is an important side condition when looking into the **at-risk** profiles. Here, we identified two low-prestige profiles, one with a high SSP, as being at-risk together with two others that are inconspicuous with respect to interests, prestige, or sextype. For the sample of university students, we would therefore conclude that the identification of at-risk students may be more straightforward in the context of expectancy–value [4] or rational choice theory [25] rather than the interaction of interests, prestige, and sextype as proposed by Gottfredson [13] (p. 103). It seems that if students have crossed the barrier of choosing a low SSP study area, they do this purposefully and consequently show a lower chance of failing.

Sometimes, there is the claim that female students experience disadvantages because they do not go into **prestigious STEM professions**, e.g., [70]. Although STEM professions may provide higher incomes than others (what we did not analyze in this paper), we see the three high prestige profiles P8 (High Prestige Medical), P11 (Academic research), or P9 (Teaching) neither as STEM specific nor with a male over-representation. P2 (Natural Science), the top-prestige STEM profile, is ranked 4th in prestige and shows a quite balanced proportion of male and female students. Thus, if we focus on prestige as a career driver, we would conclude that there are more prestigious options than following a STEM career and consequently, we would consider the claim that women avoid prestigious careers as a myth, at least in the context of university students.

Our study focuses on university students that comprise of a school leaving degree that enables them almost all career options. **Future research** might refine our approach with respect to two directions: One direction may be to look into the labor market and analyze how far such profiles can be replicated in a population of working people; the other direction might look into schools before career decisions are made and analyze respective career profiles there. Considering three points of measurement, e.g., at the ages of 15, 25, and 35 by growth curve modelling or latent transition analyses, could model changes in aspirations over time amongst the three career choice dimensions [71].

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci13030324/s1, Table S1. Means, Standard Deviations and Reliability amongst Characterization and Validation Variables. Table S2. Latent Profile Model Fit Information for the Selection of Career Profiles (*n* = 9277). Table S3. Correlation matrix of Career Choice Dimensions and other Characterization Variables for the whole sample (sex-combined). Table S4. Correlations amongst Career Choice Dimensions and other Characterization Variables for sex-separated samples (upper triangle female; lower triangle male). Table S5. Descriptive Statistics: Sex ratio, Age and Study Fields of Career Profiles (*n* = 9277). Table S6. Descriptive Statistics: Study Cluster and Sex Interaction by Career Profile (*n* = 9277). Table S7. Crosstabulation of Females and Whole Sample Career Profiles. Table S8. Crosstabulation of Whole Sample Career Profiles (P11) with Males Sample Career Profiles. Table S9. Cross-tabulated Distribution of Whole and Sex-Separated Profiles (*n* = 9277). Table S10. Career Profile Labels according to Career Choice Dimension Levels. Table S11. Matrix of Effect Size Comparisons amongst Career Profiles based on Interest Congruence. Table S12. Matrix of Effect Size Comparisons amongst Career Profiles based on the Magnitude Prestige Scale. Table S13. Matrix of Effect Size Comparisons amongst Career Profiles based on Same-Sex Proportion. Table S14. Top Five Occupational Aspirations of Career Profiles by Sex (*n* = 9277). Table S15. Chi Square Test Contingency Table of Gender Distribution in relation to the Career Profiles. Table S16. Chi Square Test Contingency Table of Study Clusters in relation to the Career Profiles. Table S17. Chi Square Test Contingency Table of Study Teaching Orientation in relation to the Career Profiles. Table S18. Descriptive Statistics of Early Background Variables by Career Profiles. Table S19. Descriptive Statistics of Validation Variables by Career Profiles. Table S20. Matrix of Effect Size Comparisons amongst Career Profiles based on Intentions to Dropout. Table S21. Matrix of Effect Size Comparisons amongst Career Profiles based on Satisfaction with Study Content. Table S22. Matrix of Effect Size Comparisons amongst Career Profiles based on Satisfaction with Study Conditions. Table S23. Matrix of Effect Size Comparisons amongst Career Profiles based on Satisfaction with Coping with Study Burdens. Table S24. Matrix of Effect Size Comparisons amongst Career Profiles based on Grades at the End of Studies. Table S25. Matrix of Effect Size Comparisons amongst Career Profiles based on Gender Role Attitudes. Table S26. Chi Square Test Contingency Table of Study Outcomes in relation to the Career Profiles. Table S27. Descriptive Statistics of Career Profile's Occupational Prestige with Parents. Figure S1. The Means and Confidence Intervals of Career Profiles across the three Career Choice Dimensions (Zoomed-in Version). Figure S2. Bargraph of the Gender Distribution of Career Profiles with Significance based of Chi Square Test Z-scores. Figure S3. Variables for the Validation of the Compromise Profiles (Zoomed-in Version). Figure S4. Variables for the Validation of the Compromise Profiles (Zoomed-in Version). Figure S5. Comparison of Career Profile's Occupational Prestige with Parents. References [2,7,19,24,50,72–76] are cited in the supplementary materials.

Author Contributions: Conceptualization, D.M. and B.E.; methodology, D.M. and B.E.; formal analysis, D.M., F.G.H. and B.E.; data curation, D.M.; writing—original draft preparation, D.M and F.G.H.; writing—review and editing, D.M., F.G.H. and B.E.; visualization, D.M.; supervision, B.E.; project administration, B.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG), Grant number ER470/2-1.

Institutional Review Board Statement: All students from this cohort gave informed consent to participate in the panel by providing their phone number for being contacted for telephone interviews after being informed about the purposes of the study. Specific information about the recruitment process can be found in the field report of the study [77].

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study through the National Educational Panel Study (NEPS; see [52]).

Data Availability Statement: Publicly available datasets were analyzed in this study. Data sources used for the analyses were from the cohort of first year students (doi: 10.5157/NEPS:SC5:14.0.0) of the German National Educational Panel Study [52]).

Acknowledgments: This paper uses data from the National Educational Panel Study (NEPS; see [52]). The NEPS is carried out by the Leibniz Institute for Educational Trajectories (LIfBi, Germany) in cooperation with a nationwide network.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gottfredson, L.S. Using Gottfredson's theory of circumscription and compromise in career guidance and counseling. In *Career Development and Counseling. Putting Theory and Research to Work*; Brown, S.D., Lent, R.W., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2005; pp. 71–100.
2. Holland, J.L. *Making Vocational Choices. A Theory of Vocational Personalities and Work Environments*, 3rd ed.; Psychological Assessment Resources; American Psychological Association: Washington, DC, USA, 1997.
3. Lent, R.W.; Brown, S.D.; Hackett, G. Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *J. Vocat. Behav.* **1994**, *45*, 79–122. [CrossRef]
4. Eccles, J.S.; Adler, T.F.; Futterman, R.; Goff, S.B.; Kaczala, C.M.; Meece, J.L.; Midgley, C. Expectancies, values, and academic behaviors. In *Achievement and Achievement Motives*; Spence, J.T., Ed.; Freeman: San Francisco, CA, USA, 1983; pp. 75–146.
5. Hannover, B.; Kessels, U. Self-to-prototype matching as a strategy for making academic choices. Why High School students do not like math and science. *Learn. Instr.* **2004**, *14*, 51–67. [CrossRef]
6. Gottfredson, L.S. Circumscription and compromise: A developmental theory of occupational aspirations. *J. Couns. Psychol.* **1981**, *28*, 545–580. [CrossRef]
7. Ertl, B.; Hartmann, F.G.; Wunderlich, A. Impact of interest congruence on study outcomes. *Front. Psychol.* **2022**, *13*, 816620. [CrossRef]
8. Hoff, K.A.; Song, Q.C.; Wee, C.J.M.; Phan WM, J.; Rounds, J. Interest fit and job satisfaction: A systematic review and meta-analysis. *J. Vocat. Behav.* **2020**, *123*, 103503. [CrossRef]
9. Nye, C.D.; Su, R.; Rounds, J.; Drasgow, F. Vocational interests and performance. A quantitative summary of over 60 years of research. *Perspect. Psychol. Sci.* **2012**, *7*, 384–403. [CrossRef] [PubMed]
10. Nye, C.D.; Su, R.; Rounds, J.; Drasgow, F. Interest congruence and performance: Revisiting recent meta-analytic findings. *J. Vocat. Behav.* **2017**, *98*, 138–151. [CrossRef]
11. Van Iddekinge, C.H.; Roth, P.L.; Putka, D.J.; Lanivich, S.E. Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. *J. Appl. Psychol.* **2011**, *96*, 1167–1194. [CrossRef]
12. Gottfredson, L.S. Gottfredson's theory of circumscription and compromise. In *Career Choice and Development*; Brown, D., Brooks, L., Eds.; Jossey-Bass: San Francisco, CA, USA, 1996; pp. 179–232.
13. Gottfredson, L.S. Gottfredson's theory of circumscription, compromise, and self-creation. In *Career Choice and Development*; Brown, D., Ed.; John Wiley & Sons: Hoboken, NJ, USA, 2002; pp. 85–148.
14. Lapan, R.T.; Hinkelman, J.M.; Adams, A.; Turner, S. Understanding Rural Adolescents' Interests, Values, and Efficacy Expectations. *J. Career Dev.* **1999**, *26*, 107–124. [CrossRef]
15. Helwig, A.A. Gender-Role Stereotyping: Testing Theory with a Longitudinal Sample. *Sex Roles* **1998**, *38*, 403–423. [CrossRef]
16. Einarsdóttir, S.; Rounds, J. Application of Three Dimensions of Vocational Interests to the Strong Interest Inventory. *J. Vocat. Behav.* **2000**, *56*, 363–379. [CrossRef]
17. Tracey, T.J.; Rounds, J. The spherical representation of vocational interests. *J. Vocat. Behav.* **1996**, *48*, 3–41. [CrossRef]
18. Nagy, G.; Trautwein, U.; Lüdtke, O. The structure of vocational interests in Germany: Different methodologies, different conclusions. *J. Vocat. Behav.* **2010**, *76*, 153–169. [CrossRef]
19. Ertl, B.; Hartmann, F.G. The interest profiles and interest congruence of male and female students in STEM and non-STEM fields. *Front. Psychol.* **2019**, *10*, 897. [CrossRef] [PubMed]
20. Ertl, B.; Hartmann, F.G.; Wunderlich, A. Stability of vocational interests during university studies. *J. Individ. Differ.* **2023**. [CrossRef]
21. Stoll, G.; Rieger, S.; Nagengast, B.; Trautwein, U.; Rounds, J. Stability and change in vocational interests after graduation from high school: A six-wave longitudinal study. *J. Personal. Soc. Psychol.* **2021**, *120*, 1091–1116. [CrossRef]
22. Sodano, S.M.; Tracey, T.J. Prestige in interest activity assessment. *J. Vocat. Behav.* **2008**, *73*, 310–317. [CrossRef]

23. Ganzeboom, H.; Treiman, D.J. nternationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations. *Soc. Sci. Res.* **1996**, *25*, 201–239. [CrossRef]
24. Christoph, B. Zur Messung des Berufsprestiges. Aktualisierung der Magnitude-Prestigeskala auf die Berufsklassifikation ISCO88. *ZUMA Nachr.* **2005**, *29*, 79–127. Available online: https://nbn-resolving.org/urn:nbn:de:0168-ssoar-207543 (accessed on 17 March 2023).
25. Stocké, V. Educational Decisions as Rational Choice? An Empirical Test of the Erikson-Jonsson Model for Explaining Educational Attainment. 2008. Available online: http://ub-madoc.bib.uni-mannheim.de/2308 (accessed on 17 March 2023).
26. Stocké, V. Explaining educational decision and effects of families' social class position: An empirical test of the Breen–Goldthorpe model of educational attainment. *Eur. Sociol. Rev.* **2007**, *23*, 505–519. [CrossRef]
27. Olmos-Gómez, M.C.; Luque-Suárez, M.; Becerril-Ruiz, D.; Cuevas-Rincón, J.M. Gender and Socioeconomic Status as Factors of Individual Differences in Pre-University Students' Decision-Making for Careers, with a Focus on Family Influence and Psychosocial Factors. *Int. J. Environ. Res. Public Health* **2021**, *18*, 1344. [CrossRef] [PubMed]
28. Tomperi, P.; Kvivesen, M.; Manshadi, S.; Uteng, S.; Shestova, Y.; Lyash, O.; Lazareva, I.; Lyash, A. Investigation of STEM Subject and Career Aspirations of Lower Secondary School Students in the North Calotte Region of Finland, Norway, and Russia. *Educ. Sci.* **2022**, *12*, 192. [CrossRef]
29. Destatis (Statistisches Bundesamt). Studierende in Mathematik, Informatik, Naturwissenschaft (MINT) und Technik-Fächern. 2021. Available online: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Tabellen/studierende-mint-faechern.html (accessed on 17 August 2021).
30. Miller, D.I.; Eagly, A.H.; Linn, M.C. Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *J. Educ. Psychol.* **2015**, *107*, 631–644. [CrossRef]
31. Braunschweig, L.; Christoph, B.; Schreyer, F. Ingenieur- und Naturwissenschaften: In manchen MINT-Fächern Dominieren Frauen. 2019. Available online: https://www.iab-forum.de/ingenieur-und-naturwissenschaften-in-manchen-mint-faechern-dominieren-frauen (accessed on 30 January 2022).
32. Makarova, E.; Aeschlimann, B.; Herzog, W. The Gender Gap in STEM Fields: The Impact of the Gender Stereotype of Math and Science on Secondary Students' Career Aspirations. *Front. Educ.* **2019**, *4*, 60. [CrossRef]
33. Frome, P.M.; Alfeld, C.J.; Eccles, J.S.; Barber, B.L. Why don't they want a male-dominated job? An investigation of young women who changed their occupational aspirations. *Educ. Res. Eval.* **2006**, *12*, 359–372. [CrossRef]
34. Armstrong, P.I.; Crombie, G. Compromises in Adolescents' Occupational Aspirations and Expectations from Grades 8 to 10. *J. Vocat. Behav.* **2000**, *56*, 82–98. [CrossRef]
35. Hesketh, B.; Durant, C.; Pryor, R. Career compromise: A test of Gottfredson's (1981) theory using a policy-capturing procedure. *J. Vocat. Behav.* **1990**, *36*, 97–108. [CrossRef]
36. Collins, L.M.; Lanza, S.T. Latent Class and Latent Transition Analysis. In *With Applications in the Social, Behavioral, and Health Sciences*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2010.
37. Blanchard, C.A.; Lichtenberg, J.W. Compromise in career decision making: A test of Gottfredson's theory. *J. Vocat. Behav.* **2003**, *62*, 250–271. [CrossRef]
38. Junk, K.E.; Armstrong, P.I. Stability of Career Aspirations: A Longitudinal Test of Gottfredson's Theory. *J. Career Dev.* **2010**, *37*, 579–598. [CrossRef]
39. Wee, S. Compromises in career-related decisions: Examining the role of compromise severity. *J. Couns. Psychol.* **2014**, *61*, 593–604. [CrossRef]
40. Hofmans, J.; Wille, B.; Schreurs, B. Person-centered methods in vocational research. *J. Vocat. Behav.* **2020**, *118*, 103398. [CrossRef]
41. Spurk, D.; Hirschi, A.; Wang, M.; Valero, D.; Kauffeld, S. Latent profile analysis: A review and "how to" guide of its application within vocational behavior research. *J. Vocat. Behav.* **2020**, *120*, 103445. [CrossRef]
42. Berger, J.B.; Ramirez, G.B.; Lyons, S. Past to Present. A Historical Look at Retention. In *College Student Retention. Formula for Student Success*, 2nd ed.; Seidman, A., Ed.; Rowman & Littlefield: Lanham, MD, USA, 2012; pp. 7–34.
43. Ryu, J.; Jeong, J. Career Compromise Types Among University Graduates During the School-to-Work Transition. *Career Dev. Q.* **2021**, *69*, 19–33. [CrossRef]
44. Cardador, M.T.; Damian, R.I.; Wiegand, J.P. Does more mean less? Interest surplus and the gender gap in STEM careers. *J. Career Assess.* **2020**, *29*, 76–97. [CrossRef]
45. Cardador, M.T.; Hill, P.L.; Lopez-Alvarez, G. Examining interests and goals as predictors of gender differences in engineers' pursuit of managerial roles. *J. Eng. Educ.* **2022**, *111*, 852–867. [CrossRef]
46. Schelfhout, S.; Wille, B.; Fonteyne, L.; Roels, E.; Derous, E.; De Fruyt, F.; Duyck, W. How interest fit relates to STEM study choice: Female students fit their choices better. *J. Vocat. Behav.* **2021**, *129*, 103614. [CrossRef]
47. Wicht, A.; Miyamoto, A.; Lechner, C.M. Are girls more ambitious than boys? Vocational interests partly explain gender differences in occupational aspirations. *J. Career Dev.* **2021**, *49*, 551–568. [CrossRef]
48. Wille, E.; Stoll, G.; Gfrörer, T.; Cambria, J.; Nagengast, B.; Trautwein, U. It takes two: Expectancy-value constructs and vocational interests jointly predict STEM major choices. *Contemp. Educ. Psychol.* **2020**, *61*, 101858. [CrossRef]
49. Pfuhl, N. *Untersuchung zur Bestimmung von Typischen Merkmalen des Image von Studienfächern*; Waxmann: Münster, Germany, 2010.
50. Athenstaedt, U. Normative Geschlechtsrollenorientierung: Entwicklung und Validierung eines Fragebogens. *Z. Differ. Diagn. Psychol.* **2000**, *21*, 91–104. [CrossRef]

51. Morin, A.J.S.; Bujacz, A.; Gagné, M. Person-centered methodologies in the organizational sciences: Introduction to the feature topic. *Organ. Res. Methods* **2018**, *21*, 803–813. [CrossRef]
52. Blossfeld, H.-P.; Roßbach, H.-G. (Eds.) *Education as a Lifelong Process: The German National Educational Panel Study (NEPS)*, 2nd ed.; Springer VS: Berlin/Heidelberg, Germany, 2019.
53. FDZ-LIfBi. *Data Manual NEPS Starting Cohort 5–First-Year Students, from Higher Education to the Labor Market, Scientific Use File Version 15.0.0*; Leibniz Institute for Educational Trajectories, National Educational Panel Study: Bamberg, Germany, 2021.
54. International Labour Organization. International Standard Classification of Occupations. Structure, group definitions and correspondence tables. Geneva: International Labour Organization (ILO) (1). 2012. Available online: https://www.ilo.org/public/english/bureau/stat/isco/docs/publication08.pdf (accessed on 20 March 2023).
55. Bundesagentur für Arbeit, Klassifikation der Berufe 2010—Überarbeitete Fassung 2020, Band 2: Definitorischer und Beschreibender Teil. Nürnberg, November 2021. Available online: https://statistik.arbeitsagentur.de/DE/Statischer-Content/Grundlagen/Klassifikationen/Klassifikation-der-Berufe/KldB2010-Fassung2020/Printausgabe-KldB-2010-Fassung2020/Generische-Publikationen/KldB2010-PDF-Version-Band2-Fassung2020.pdf?__blob=publicationFile&v=20 (accessed on 17 March 2023).
56. Treiman, D.J. *Occupational Prestige in Comparitive Perspective*; Academic Press: New York, NY, USA, 1977.
57. von Maurice, J. ICA-D. In *Deutschsprachige Version des Inventory of Children's Activities—Revised (ICA-R, Tracey & Ward, 1998)*; Unpublished Manuscript; Otto-Friedrich-Universität Bamberg: Bamberg, Germany, 2006.
58. O*Net. O*NET®22.3 Database. 2018. Available online: https://www.onetcenter.org/dl_files/database/db_22_3_dictionary.pdf (accessed on 17 March 2023).
59. Beavis, A. Evidence in Support of Gottfredson's Common Cognitive Map of Occupations. *Aust. J. Career Dev.* **2007**, *16*, 38–44. [CrossRef]
60. Destatis [Statistisches Bundesamt]. Erwerbstätige ab 15 Jahren nach Beruf (KldB 2010 5-Steller) in Deutschland * [Employed Persons Aged 15 and over by Occupation (KldB 2010 5-Digit) in Germany *] [xlx]. 2022. Available online: https://ergebnisse.zensus2011.de/auswertungsdb/download?xls=11&tableId=BEV_10_31&locale=DE (accessed on 17 March 2023).
61. Hartmann, F.G.; Mouton, D.; Ertl, B. The Big Six interests of STEM and non-STEM students inside and outside of teacher education. *Teach. Teach. Educ.* **2022**, *112*, 103622. [CrossRef]
62. Muthén, L.K.; Muthén, B.O. *Mplus User's Guide, 8th ed*; Muthén & Muthén: Los Angeles, CA, USA, 2017.
63. Nylund, K.L.; Asparouhov, T.; Muthén, B.O. Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Struct. Equ. Model. Multidiscip. J.* **2007**, *14*, 535–569. [CrossRef]
64. Field, A.P.; Miles, J.; Field, Z. *Discovering Statistics Using R*; Sage: London, UK; Thousand Oaks, CA, USA, 2012.
65. Bernd, W. Concepts and Measurement of Prestige. *Annu. Rev. Sociol.* **1992**, *18*, 253–280. [CrossRef]
66. Ertl, B.; Csanadi, A.; Tarnai, C. Getting closer to the digital divide: An analysis of impacts on digital competencies based on the German PIAAC sample. *Int. J. Educ. Dev.* **2020**, *78*, 102259. [CrossRef]
67. Nosek, B.A.; Smyth, F.L.; Hansen, J.J.; Devos, T.; Lindner, N.M.; Ranganath, K.A.; Smith, C.T.; Olson, K.R.; Chugh, D.; Greenwald, A.G.; et al. Pervasiveness and correlates of implicit attitudes and stereotypes. *Eur. Rev. Soc. Psychol.* **2007**, *18*, 36–88. [CrossRef]
68. Obioma, I.F.; Hentschel, T.; Hernandez Bark, A.S. Gender stereotypes and self-characterizations in Germany and Nigeria: A cross-cultural comparison. *J. Appl. Soc. Psychol.* **2022**, *52*, 764–780. [CrossRef]
69. Ertl, B.; Hartmann, F.G.; Heine, J.-H. Analyzing Large-Scale Studies: Benefits and Challenges [Opinion]. *Front. Psychol.* **2020**, *11*, 577410. [CrossRef]
70. Sáinz, M.; Martínez-Cantos, J.-L.; Rodó-de-Zárate, M.; Romano, M.J.; Arroyo, L.; Fàbregues, S. Young Spanish People's Gendered Representations of People Working in STEM. A Qualitative Study. *Front. Psychol.* **2019**, *10*, 996. [CrossRef]
71. Woo, S.E.; Jebb, A.T.; Tay, L.; Parrigon, S. Putting the "Person" in the Center. Review and Synthesis of Person-Centered Approaches and Methods in Organizational Science. *Organ. Res. Methods* **2018**, *21*, 814–845. [CrossRef]
72. Deutsches Jugendinstitut (DJI), München; Technische Universität Chemnitz-Zwickau. *DJI Foreigner Survey—Young People*; ZA3371 Data File Version 1.0.0; GESIS Data Archive: Cologne, Germany, 2001. [CrossRef]
73. Esser, H.; Stocké, V. *Forschungsprojekt "Bildungsaspirationen, Bezugsgruppen und Bildungsentscheidungen"*; Mannheimer Bildungspanel, 2003.
74. Stocké, V. *Statuserhaltsmotiv*; ZIS—GESIS Leibniz Institute for the Social Sciences: Mannheim, Germany, 2005. [CrossRef]
75. Trautwein, U.; Köller, O.; Watermann, R. Transformation des Sekundarschulsystems und akademische Karrieren—Zusammenfassung, Diskussion und ein Ausblick [Transformation of the secondary school system and academic careers—Summary, discussion and outlook]. In *Wege zur Hochschulreife in Baden-Württemberg*; Köller, O., Watermann, R., Trautwein, U., Lüdtke, O., Eds.; VS Verlag für Sozialwissenschaften: Berlin/Heidelberg, Germany, 2004; pp. 451–471. [CrossRef]
76. Westermann, R.; Heise, E.; Spies, K.; Trautwein, U. Identifikation und Erfassung von Komponenten der Studienzufriedenheit. *Psychol. Erzieh. Unterr.* **1996**, *43*, 1–2.
77. Steinwede, J.; Aust, F. *Methodenbericht NEPS Startkohorte 5: CATI-Haupterhebung. Herbst 2010. B52*; Infas: Bonn, Germany, 2012.

education sciences

Article

"Who's Better at Math, Boys or Girls?": Changes in Adolescents' Math Gender Stereotypes and Their Motivational Beliefs from Early to Late Adolescence

Christine R. Starr [1,2,*], Yannan Gao [3], Charlott Rubach [4], Glona Lee [1], Nayssan Safavian [1], Anna-Lena Dicke [1], Jacquelynne S. Eccles [1] and Sandra D. Simpkins [1]

1 School of Education, University of California, Irvine, CA 92697, USA; glonal@uci.edu (G.L.); nayssan.safavian@uci.edu (N.S.); adicke@uci.edu (A.-L.D.); jseccles@uci.edu (J.S.E.); simpkins@uci.edu (S.D.S.)
2 Department of Educational Psychology, University of Wisconsin, Madison, WI 53706, USA
3 Division of Arts and Science, New York University Shanghai, Shanghai 200122, China; yannang@uci.edu
4 Department of School Pedagogy and Educational Research, Universitat Rostock, 18055 Rostock, Germany; charlott.rubach@uni-rostock.de
* Correspondence: starr3@wisc.edu

Abstract: Though adults tend to endorse the stereotype that boys are better than girls in math, children tend to favor their own gender or be gender egalitarian. When do individuals start endorsing the traditional stereotype that boys are better? Using two longitudinal U.S. datasets that span 1993 to 2011, we examined three questions: (1) What are the developmental changes in adolescents' gender stereotypes about math abilities from early to late adolescence? (2) Do the developmental changes vary based on gender and race/ethnicity? (3) Are adolescents' stereotypes related to their math motivational beliefs? Finally, (4) do these patterns replicate across two datasets that vary in historical time? Adolescents in grades 8/9 and 11 were asked whether girls or boys are better at math (n's = 1186 and 23,340, 49–53% girls, 30–54% White, 13–60% Black, 1–22% Latinx, and 2% to 4% Asian). Early adolescents were more likely to be gender egalitarian or favor their own gender. By late adolescence, adolescents' stereotypes typically shifted towards the traditional stereotype that boys are better. In terms of race/ethnicity, White and Asian adolescents significantly favored boys, whereas Black and Latinx adolescents were more likely to endorse gender egalitarian beliefs. Adolescents' stereotypes were significantly related to their expectancy beliefs, negatively for girls and positively for boys.

Keywords: STEM; math; gender stereotypes; motivational beliefs; expectancy-value beliefs

Citation: Starr, C.R.; Gao, Y.; Rubach, C.; Lee, G.; Safavian, N.; Dicke, A.-L.; Eccles, J.S.; Simpkins, S.D. "Who's Better at Math, Boys or Girls?": Changes in Adolescents' Math Gender Stereotypes and Their Motivational Beliefs from Early to Late Adolescence. *Educ. Sci.* **2023**, *13*, 866. https://doi.org/10.3390/educsci13090866

Academic Editor: Brian M. McSkimming

Received: 14 June 2023
Revised: 17 August 2023
Accepted: 22 August 2023
Published: 25 August 2023

1. Introduction

Psychologists have studied and attempted to minimize gender gaps in STEM for several decades, yet gender gaps persist to this day [1,2]. Gender gaps in math are of particular concern, as math serves as a gateway to numerous STEM educational opportunities [3]. One possible contributor to the persistent gender gaps in math is if individuals' stereotypes have not changed historically. Gendered math ability stereotypes, such as the belief that boys are better at math than girls, exacerbate gender gaps by promoting boys in math while hindering girls, who are already marginalized in math [1,4]. Though women historically have made gains in many STEM fields (including math) [5], U.S. media portrayals have not substantially improved [6]. Furthermore, a recent U.S. study suggests that parents' math ability gender stereotypes have not significantly changed over a 20-year period from 1984 to 2009 [7]. Although the U.S. ranks in the upper third of countries for gender equality (ranking 43 out of 146), it currently ranks below many nations, including most of Europe [8]. Furthermore, the U.S. has a greater gender gap in STEM participation compared to many nations [1]. Examining the potential historical differences in U.S. adolescents' math stereotypes will help us understand whether these gender stereotypes have persisted like parents' stereotypes and potentially continue to contribute to gender differences in adolescents' math motivational beliefs.

Most of the existing research on adolescents' gender ability stereotypes in math focuses on Whites and does not address developmental changes nor historical differences [9,10], although there are notable exceptions, e.g., [9,11]. Furthermore, replication of research findings is important for our field but seldom investigated [12]. In particular, conceptual replication is valuable because it examines whether effects persist across different demographics, measures, and contexts, which speaks to the robustness of the finding [12,13]. To address these gaps in the literature, the present study explores the prevalence of adolescents' math ability gender stereotypes and the correlates across two large U.S. datasets that span early and late adolescence and include racially/ethnically diverse adolescents. Though the two longitudinal datasets differ in some respects, testing these processes across these two datasets allowed us to test the extent to which the findings were replicated during two different historical time periods, namely the 1990s and 2010s. Prior research has found that parents' math gender stereotypes favoring boys [7] and gender differences in math self-concept [2] have not changed during this same period of time. Thus, we wished to explore if adolescent gender stereotypes also have not changed during this time period. Similar endorsement of math gender stereotypes during this time period may help explain why gender differences in math self-concept, in addition to differences in participation in certain STEM fields, remain [1,2].

1.1. Understanding Math Ability Stereotypes through Situated Expectancy-Value Theory and Social Status Theory

Situated expectancy-value theory [14] postulates that gender stereotypes are culturally situated attitudes that influence individuals' beliefs, choices, and performance. According to this theory, adolescents' gender stereotypes about a domain like math influence their subsequent motivational beliefs, which are composed of two primary constructs— expectancies for success (how well someone expects to perform in a domain such as math) and subjective task-values (beliefs about how interesting, useful, or important a task is) [14]. Specifically, if an individual believes their gender is generally good at a domain, they are likely to have higher expectancy and value beliefs in that domain [10,15]. Individuals' motivational beliefs, in turn, are central determinants of their academic and career outcomes throughout life [14]. Thus, adolescents' gender stereotypes about a domain should have implications for their motivational beliefs and subsequent outcomes in that domain.

Situated expectancy-value theory argues that social position factors, including gender and race/ethnicity, influence where one is situated in society; however, it provides less insight into the specific processes or expectations. To examine why the prevalence of individuals' math gender stereotypes may differ by social group, we employed social status theory. Social status theory [16] argues that socially privileged groups (e.g., White individuals, boys) are more likely to endorse traditional gender stereotypes because they have more to gain by upholding current social systems. Based on this theory, male, White, and Asian adolescents (particularly White and Asian boys) have more to gain by upholding traditional gender stereotypes about math because of their strong representation, whereas female, Black, and Latinx adolescents (particularly Black and Latina girls) have less to gain from upholding traditional stereotypes and may be more likely to endorse gender egalitarian or non-traditional beliefs. Although White and Asian girls may not directly benefit from the traditional math gender stereotype that boys have stronger math abilities, they may still be more likely to endorse them than Black and Latinx adolescents. This is because members of their family and community (such as their fathers and brothers) may benefit from traditional math gender stereotypes; recent evidence suggests that Asian and White parents are more likely to hold traditional math gender stereotypes than Black and Latinx parents [7]. As a result, Asian and White parents and other family members may be more likely to transmit these beliefs to their children, including their daughters. Additionally, competency and self-reliance are valued as part of Black women's gender identity [17]. Thus, female, Black, and Latinx adolescents may be less likely to endorse traditional math gender stereotypes, whereas male, Asian, and White adolescents may be more likely to endorse these stereotypes.

Adolescence is an important developmental period to investigate changes in math gender stereotypes because it coincides with cognitive maturation as well as identity development [18–20]. Greater endorsement of traditional math gender stereotypes as adolescents move through early to late adolescence is plausible because adolescents have increasingly sophisticated cognitive abilities to apply stereotypes to their own identity and motivational beliefs [20,21]. Moreover, puberty as well as middle and high school settings, may lead adolescents to pay more attention to gendered social comparisons in math classrooms [22]. Adolescents also have a greater lifetime exposure to stereotypes about STEM than younger children and may be in contexts that highlight traditional stereotypes, like advanced STEM courses [22,23]. Thus, adolescents' traditional stereotypes may become more prevalent with age due to developmental processes and as contexts draw attention to gender and math ability [24].

1.2. Empirical Evidence on the Prevalence and Differences in Adolescents' Math Ability Gender Stereotypes

Developmental changes. Prior research suggests that individuals' beliefs may shift to more traditional math gender stereotypes during adolescence. Two U.S. studies among racially/ethnically diverse youth suggest that early adolescents have gender egalitarian views—the belief that both genders are equally good at math—compared to children [25] and late adolescents [10]. One of these studies charted the changes in students' math gender stereotypes longitudinally; using one of the datasets used in the present study, the researchers found that U.S. adolescents' gender-stereotyped beliefs became more traditional across high school, and by the eleventh grade, both girls and boys were significantly more likely to believe males are better at math [10]. These patterns align with cross-sectional research on majority-White French and German high school and college students [26,27]. Though several studies suggest that the traditional stereotype that boys are better than girls at math is more prominent in late adolescence, there are some inconsistencies in the literature. One study suggests traditional stereotype beliefs may emerge as early as the ninth grade among majority-White gifted youth in the UK [28], whereas another UK study and a French Canadian study (both majority White) found that adolescents (grades 9–12) and late adolescents (grade 10) hold gender egalitarian beliefs [9,29]. Thus, research conducted among diverse samples in the U.S. has largely found that early adolescents hold gender egalitarian views, which switch to more traditional views in late adolescence [10]. These differences were replicated in some cross-sectional non-U.S. studies [26,27] but not in other studies, potentially because the latter studies had, on average, younger adolescents (e.g., grade 10 versus grade 11) [10,26,29]. Thus, the present study aimed to expand upon the prior literature by examining the replication of changes in adolescent math gender stereotypes in two diverse, longitudinal U.S. datasets that included both early and late adolescents.

Gender differences. Findings from several studies suggest that boys typically endorse traditional gender math stereotypes more often than girls [9,30]. This is potentially because, according to social status theory, boys have more to gain from endorsing these beliefs compared to girls and thus may be more likely to endorse traditional math gender stereotypes [16]. Alternatively, some parents and media may socialize girls not to endorse traditional gender stereotypes. Prior research finds that parents of girls are less likely to endorse traditional math gender stereotypes [7], and parents may be more careful about endorsing traditional gender stereotypes around their daughters. Relatedly, although the media as a whole still largely portrays sexist stereotypes [31], there are some programs and media targeted towards girls that endorse the belief girls can be anything, including being good at math. Thus, endorsing traditional math stereotypes may be more common among boys and may be particularly pronounced among boys in late adolescence, given the prevalence of traditional math stereotypes at that time [27,32,33].

Racial/Ethnic differences. Social status theory suggests that members of groups underrepresented in STEM (i.e., Black and Latinx adolescents) will be less likely to endorse traditional stereotypes in that domain compared to members of groups overrepresented in STEM (i.e., Asian and White adolescents). Findings among Black and White Americans support this assertion [16,34]. For example, one study about gender stereotypes among fourth-, sixth-, and eighth-grade Black and White adolescents found that girls and Black adolescents were less likely to endorse traditional gender stereotypes than their peers [10]. Black girls were the least likely to endorse traditional gender ability stereotypes in math and science, whereas White boys were the most likely to endorse the stereotype [16]. Relatedly, one study among Black seventh and eighth graders found that adolescents, on average, endorsed non-traditional math stereotypes favoring girls [35]. Thus, evidence from several studies in elementary and middle school suggests that Black adolescents may endorse less traditional gender stereotypes than White adolescents, though this has not been tested during high school. Most prior research has taken place among early adolescents when non-traditional or egalitarian views are expected. Research with late adolescents and longitudinal research across adolescence is needed to chart the developmental changes in Black adolescents' math gender stereotypes.

Moreover, few studies examine the prevalence of math ability gender stereotypes among Latinx and Asian adolescents. Much of the existing research focuses on citizens of Asian countries rather than Asian Americans; thus, it is unclear to what extent these findings generalize to Asian adolescents in the U.S. Similar to Asian families, Latinx families have been found to endorse more traditional gender roles [36]. However, this may not translate to more traditional math ability gender stereotypes for Latinx adolescents, given that two studies found Latinx mothers and parents had gender egalitarian views about individuals' math abilities [7,30], which were significantly less traditional than Asian and White parents' stereotypes [7]. Two additional studies among adult women found Latina women had significantly less traditional STEM-gender implicit stereotypes compared to White women [37,38]. Thus, past studies suggest that Asian adolescents might endorse more traditional math gender stereotypes similar to White adolescents, whereas Latinx adolescents may endorse less traditional stereotypes similar to Black adolescents, although it remains an open question among Asian and Latinx adolescents. It is important to explore math gender stereotypes among Asian and Latinx adolescents for multiple reasons. First, both racial/ethnic groups comprise sizable portions of the U.S. school age population and are some of the fastest growing racial/ethnic groups. Second, it is important to explore math gender stereotypes among Asian and Latinx adolescents because we cannot assume that the patterns for White families generalize to other groups, as emerging evidence among Black and White adolescents suggests stereotype processes vary by racial/ethnic group [16]. Finally, studying stereotypes of Asian and Latinx adolescents has critical theoretical implications. Both ethnic groups culturally endorse traditional gender roles, yet differ in their representation in STEM, which makes for an interesting theoretical test of social status theory [16].

1.3. Adolescents' Math Gender Stereotypes and Motivational Beliefs

According to situated expectancy-value theory, adolescents' gender stereotypes about math should shape their math motivational beliefs (such as expectancy-value beliefs) [14]. One study among high school students in the U.S. linked traditional math gender stereotype endorsement to lower math motivational beliefs for girls and higher beliefs for boys [10]. However, another study among U.S. fourth, sixth, and eighth graders (54% White, 30% Black) found that traditional gender stereotypes were positively related to boys' expectancy beliefs but not significantly related to girls' beliefs [39]. Much of past research has been conducted among children; however, it is also important to explore these relations replicated during late adolescence, when stereotypes are theorized to become more traditional and thus may be more likely to benefit boys while harming girls. Furthermore, it is important to explore the extent to which these processes replicate among Asian, Latinx, and Black adolescents in the U.S. or the extent to which the relations may differ across these groups. For example, because of the model minority stereotype, Asian adolescents may be stereotyped in school and society as excelling at math, regardless of gender [40]. Furthermore, studies have found that Asian adolescents do not demonstrate the gender differences in math motivational beliefs that White adolescents do [2,11,41]. However, stereotype threat research indicates that while highlighting Asian identity may improve performance among Asian American women, highlighting female identity may lower performance [42]. This suggests that gender-based stereotypes about math still affect Asian American women and girls. Relatedly, given that Latino boys are often stereotyped as not doing well academically by teachers [22], math gender stereotypes may not benefit them in the same way they do White boys. Thus, our study adds to this body of literature by testing the extent to which racially/ethnically diverse U.S. adolescents' gender stereotypes are related to their expectancy and value beliefs in two longitudinal datasets that span early and late adolescence and differ in terms of historical time.

1.4. Current Study

Based on situated expectancy-value theory [14], adolescents' gender stereotypes about math ability should vary by age, gender, or race/ethnicity and be differentially related to their expectancy-value beliefs. We utilized two large, longitudinal datasets that included early adolescents at the first time point (eighth and ninth grades) and late adolescents at the second time point (eleventh grade) from the mid-1990s to the 2010s. The two datasets were identified by consulting experts in the field and searching for datasets in databases of publicly available data (e.g., IES, OSF, and the Henry A. Murray Research Archive at Harvard) using terms such as "math gender stereotypes". These two datasets were included because they had data on math gender stereotypes and expectancy-value belief questions on adolescents in the U.S.; other datasets did not include both types of data for this age group. Using multiple datasets also afforded us the ability to test for conceptual replication across historical time, study design, and races/ethnicities, to see whether the results are similar across the datasets. Scholars argue that conceptual replication provides more compelling evidence than exact replication [12] because it addresses the robustness of the findings across different participant demographics, measures, and contexts [13]. As an open research question, we examined conceptual replication and historical differences in prevalence.

Based on the reviewed literature and situated expectancy-value theory [14], we examined the following hypotheses:

1.4.1. Hypothesis 1: Changes in the Prevalence of Math Gender Stereotypes over Time

Based on past research [10], we expected adolescents to move from same-gender biased or egalitarian gender stereotypes in early adolescence (eighth/ninth grades) to traditional stereotypes favoring boys in late adolescence (eleventh grade). Additionally, we expected this pattern to replicate across the two datasets, given that past research has found these developmental differences, and parents' stereotypes and youth's math self-concepts have not changed over historical time. Finally, we explored if these changes varied based on gender and race/ethnicity.

1.4.2. Hypothesis 2: Group Differences in the Prevalence of Math Gender Stereotypes

Based on past research [43], we expected that boys would hold significantly more traditional math gender stereotypes at all grade levels compared to girls. Furthermore, based on past research and theory [16], we expected that even though Asian and White adolescents would take the developmental path outlined above (favoring their own gender in early adolescence but boys in late adolescence), Black and Latinx adolescents would largely be gender egalitarian or favor their own gender across adolescence. Additionally, we expected that Asian and White adolescents would hold significantly more traditional math gender stereotypes at all grade levels compared to Black and Latinx. We expected these patterns to replicate across the two datasets, given a past study found that parent gender stereotypes (including racial/ethnic differences) were replicated across four datasets [7].

1.4.3. Hypothesis 3: Adolescents' Math Gender Stereotypes in Relation to Their Math Motivational Beliefs

We expected that more traditional gender stereotypes would be positively associated with math expectancy and value beliefs among boys but negatively associated with expectancy and value beliefs among girls. We additionally wished to explore these relations within each racial/ethnic group. We expected these patterns to replicate across both datasets, given past research has found these patterns, e.g., [10].

2. Method

2.1. Datasets

This present study used two longitudinal datasets. The first was a large, local dataset: the Maryland Adolescent Development in Context Study (MADICS). The second dataset was a large nationally representative dataset: the High School Longitudinal Study (HSLS). For more information on MADICS, including study questionnaires, see https://garp.education.uci.edu/madics.html (accessed on 21 August 2023). For more information about HSLS, see the National Center for Education Statistics website [44].

For this paper, we used all waves when the math gender stereotype data were collected in each dataset; MADICS and HSLS included the stereotype data at two waves—either the eighth or ninth grades and the eleventh grade. Adolescents were excluded from our sample if they were missing data for (a) their gender (MADICS n = 11; HSLS n = 0) or (b) their stereotype data at both time points (MADICS n = 285; HSLS n = 160). Detailed descriptions of the participants in the analytic sample for each study are in Table 1 and below.

Table 1. Participants by dataset.

Indicator	MADICS	HSLS
Design	Longitudinal, 1 cohort	Longitudinal, 1 cohort
Data included		
Years when collected	1993–1996	2009–2012
Youth waves (Year)	W3 (1993) and W4 (1996)	W1 (2009) and W2 (2012)
Youth's grades	8th and 11th Grades	9th and 11th Grades
Sample sizes		
Total N: Dataset	1482	23,500
Total N: Current study	1186	23,340
Demographic information		
% Girls (n)	49% (n = 585)	50% (n = 11,670)
% White (n)	30% (n = 350)	50% (n = 11,670)
% Black (n)	60% (n = 708)	13% (n = 3030)
% Latinx (n)	1% (n = 15)	22% (n = 5140)
% Asian (n)	2% (n = 20)	4% (n = 930)
% Other race/ethnicity	8% (n = 93)	11% (n = 2570)
% Parent college degree	42%	45%
Family income	\$30,000 or less: 23% \$30–60,000: 43% Over \$60,000: 34%	\$35,000 or less: 28% \$35–75,000: 32% Over \$75,000: 41%

W = Wave. HSLS analyses, including demographic information, were weighted to be nationally representative. HSLS SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), Base Year and First Year Follow-Up public dataset; n's rounded to the nearest tens place.

2.1.1. Maryland Adolescent Development in Context Study (MADICS): Eighth and Eleventh Grades

MADICS is a longitudinal study following one cohort of 1482 adolescents (49% girls) and their families. The sample includes primarily Black (60% of the sample) and White (30%) families (in addition to 15 Latinx families, 20 Asian families, and 93 families that were multiethnic or another race/ethnicity) from a range of socioeconomic statuses in Maryland and the D.C. area. The present study used data from adolescents in Waves 3 and 4 (1993 and 1996) when adolescents were in the eighth (ages 13–14; n = 1043) and eleventh grades (ages 16–17; n = 855).

2.1.2. High School Longitudinal Study (HSLS): Ninth and Eleventh Grades

HSLS is a longitudinal and nationally representative study of more than 20,000 adolescents (49% girls, 54% White, 17% Latinx, 11% Black, and 9% Asian) from 944 schools in the U.S. Eligible schools and adolescents were selected using a random, stratified sample design. The present study used data from adolescents in Wave 1 (2009) and Wave 2 (2012) when adolescents were in the ninth grade (ages 14–15; n = 20,720) and eleventh grade (ages 16–17; n = 20,010). To be representative of the U.S. population, analyses in this study were adjusted using the analytic weight W2STUDENT, strata, and clusters.

2.2. Measures

HSLS and MADICS included the same constructs with slight item-level variations, allowing us to examine conceptual replication [12]. For a full list of items by dataset, see Supplementary Table S1.

2.2.1. Adolescents' Math Ability Gender Stereotypes

Adolescents in HSLS and MADICS reported their gender math ability stereotypes using a similar item [45]. This item was measured in the ninth and eleventh grades in HSLS and the eighth and eleventh grades in MADICS; these were the only two time points the question was asked in both datasets. In the HSLS dataset, they were asked, "In general, how would you compare males and females in math?" (1 = females much better to 5 = males much better). In the MADICS dataset, adolescents were asked, "Who is better at math and science, girls or boys?" (1 = girls are a lot better, 2 = girls are somewhat better, 3 = girls and boys are the same, 4 = boys are somewhat better, and 5 = boys are a lot better). Adolescents' responses were recoded. The response "girls and boys are the same" was set to 0. Responses that endorsed the belief that girls are better at math were set to be below 0 (-1 = girls are somewhat better, -2 = girls are a lot better). Responses that reflected the traditional belief that boys are better at math were set to be above 0 (1 = boys are somewhat better, 2 = boys are a lot better).

Past research about math gender stereotypes has used this item [10,24,45]. The item has strong face validity as well as criterion validity because it directly asks about ability stereotypes, and prior research has found that the item significantly correlates with other scales in the expected direction [24]. Furthermore, prior research indicates that single items perform similarly to scales [46–48], and the use of a single item allows adolescents to endorse the belief that girls are better at math, that boys are better, or that both genders are the same at math [10].

2.2.2. Adolescents' Math Expectancy and Value Beliefs

Both HSLS and MADICS had math expectancy and value belief items based on Eccles' situated expectancy-value model that prior work has validated [49,50]. The expectancy belief scale had three items in MADICS (eighth grade: alpha = 0.82; eleventh grade: alpha = 0.85) and four items in HSLS (ninth grade: alpha = 0.89; eleventh grade: alpha = 0.90). An example expectancy belief item is "How good at math are you?" (1 = not at all good, 7 = very good). The value belief scale was three items in MADICS (eighth grade: alpha = 0.68; eleventh grade: alpha = 0.64) and six items in HSLS (ninth grade: alpha = 0.80; eleventh grade: alpha = 0.80). An example value belief item is "In general, how useful is what you learn in math?" (1 = not at all useful, 7 = very useful). HSLS used a response scale of 1–4, and MADICS used a 1–5 response scale for two value questions and a 1–7 scale for the remaining question; in order to combine the three items, the 1–5 response scale was converted to a 1–7 scale using a linear transformation. For more information regarding psychometric properties such as measurement invariance, see [41,51] (HSLS) and [52] (MADICS).

2.2.3. Background and Covariates

For the present analyses, we examined potential variations based on adolescent gender (girl or boy, reported by adolescents) and adolescent race/ethnicity (Asian, Black, Latinx, White, or other, reported by adolescents). Additionally, we controlled for household income (reported by the caregivers) and highest caregiver education (reported by the caregivers).

2.3. Plan of Analysis

Data were analyzed in SPSS version 26 for MADICS, and STATA version 14.1 or R version 4.0.1 and Rstudio were used for HSLS. The proportion of missing values across the two datasets ranged from 11–12% for adolescent gender stereotypes in early adolescence, 14–25% for adolescent gender stereotypes in late adolescence, 11–19% for motivational beliefs items, and 0–28% for family income and parent education level. Missing data were imputed if adolescents had stereotype data from at least one single time point. Within each dataset, we compared adolescents who were missing some data to adolescents who had complete data. In MADICS, participants who were missing some data reported significantly lower family incomes ($d = 0.22$) and parent education ($d = 0.13$), were more likely to be boys than girls ($\varphi = 0.09$) and were more likely to be Black than White ($\varphi = 0.04$) (ps < 0.003) compared to adolescents with complete data. There were no significant differences in adolescents' math motivational beliefs or math gender stereotypes (ds = −0.07–0.02; ps < 0.089). In HSLS, participants with some data missing reported a significantly lower parent education ($d = 0.29$), family income ($d = 0.28$), and math motivational beliefs (ds = 0.07–0.15); had more traditional math gender stereotypes (ds = 0.06–0.11); were more likely to be boys ($\varphi = 0.04$); and were more likely to be Asian ($\varphi = 0.05$), Black ($\varphi = 0.05$), Latinx ($\varphi = 0.05$), and less likely to be White ($\varphi = -0.12$) (all ps < 0.001) compared to adolescents with complete data. Thus, although there were significant differences between adolescents with missing data and those with complete data, all effects were small or less than small. Missing data in both datasets were handled by multiple imputations [53]. Thirty datasets were imputed in SPSS using auxiliary variables that included data from other time points, demographic data, and transcript data (e.g., math grade and SES).

2.3.1. Hypothesis 1: Changes in the Prevalence of Math Gender Stereotypes over Time

To understand the prevalence of adolescents' math gender stereotypes, it is important to describe how much these beliefs change over time and the extent to which those changes vary across groups. For Hypothesis 1, we expected adolescents would hold more traditional math gender stereotypes in late adolescence (i.e., eleventh grade) compared to early adolescence (i.e., eighth grade in MADICS and ninth grade in HSLS). We expected that this would replicate across the two datasets. We examined these changes through three steps: (a) repeated measures MANCOVAs in each dataset, (b) a series of one-sample t-tests in each dataset, and (c) testing the extent to which these effects replicate across the two datasets.

First, we examined the overall changes in adolescents' math gender stereotypes and the extent to which these changes varied across gender and race/ethnicity with repeated measures MANCOVA. We estimated one repeated measure MANCOVA in each dataset to investigate (a) the main effects of time, gender, and race/ethnicity, as well as (b) the interactions between time, gender, and race/ethnicity. We ran this MANCOVA both with and without using family income and parent education as the covariates; for the analyses with covariates, please see the Supplementary Materials. p-values and the effect sizes (eta squared) were used to determine meaningful significance. These analyses describe the extent to which there were changes in adolescents' math gender stereotypes and whether these changes varied across groups (e.g., boys and girls); however, they provide less information on what the changes look like for each group or if adolescents favor a particular gender or if their stereotypes significantly differ from gender egalitarian views. For instance, girls' stereotypes might shift from slightly favoring their own gender to slightly favoring boys over time, but these stereotypes may not differ from gender egalitarian beliefs at either time point. For that, we employed one-sample *t*-tests.

We examined the prevalence of these math gender stereotypes in early and late adolescence within each gender and racial/ethnic group using a series of one-sample *t*-tests, which examined the extent that each group's mean differed from zero. Because zero was coded to signify gender egalitarian beliefs (i.e., believing boys and girls are equally good at math), these *t*-tests examine the extent to which participants in each group hold egalitarian beliefs or favor one gender (and which gender they favor). For example, if early adolescents are more egalitarian as expected, boys' and girls' means should not significantly differ from zero, though they may still change over time. If late adolescents favor boys as expected, then boys' and girls' means should be different from zero with positive means, which indicates they favor boys, on average. In each dataset, we tested these one-sample *t*-tests for girls and for boys overall, which included all racial/ethnic groups, as well as for girls and for boys in each racial/ethnic group. For example, in MADICS, we estimated six one-sample *t*-tests: girls overall, boys overall, Black girls, White girls, Black boys, and White boys.

Finally, we examined the extent to which the analyses in the first two parts, namely the MANCOVA and one-sample *t*-tests, were replicated across the two datasets. After calculating the means, standard deviations, and effect sizes within each dataset separately, we used random-effects models to estimate the combined effect, or an average effect size across the datasets adjusted for sample sizes using Comprehensive Meta-Analysis, version 3.3 (CMA, [54]). In addition to the combined effect size, we obtained two statistics (Cochran's Q and I^2) that describe the heterogeneity of the effect sizes across the datasets. A significant Cochran's Q statistic indicates that there is significant heterogeneity across datasets, whereas the I^2 indicates the percentage of variance that differs between datasets [55,56].

2.3.2. Hypothesis 2: Group Differences in the Prevalence of Adolescents' Math Gender Stereotypes

A second important aspect in understanding the prevalence of adolescents' math gender stereotypes is describing the differences across groups. We expected boys would have more traditional stereotypes than girls. We also expected that Asian and White adolescents would have more traditional stereotypes than Black and Latinx adolescents. We tested these two hypotheses with the gender and race/ethnicity main effects and the gender by race/ethnicity interaction from the MANCOVA described under Hypothesis 1. A MANCOVA was estimated for each dataset separately. We examined the extent to which these findings replicated across datasets using the random-effects models described under Hypothesis 1.

2.3.3. Hypothesis 3: Adolescents' Math Gender Stereotypes in Relation to Their Math Motivational Beliefs

We hypothesized that adolescents' math gender stereotypes would significantly relate to their motivational beliefs (positively for boys, negatively for girls). Regressions controlling for parent education and income were run by grade and gender for adolescents' stereotypes and math motivational beliefs. Parent education and income were controlled for, given they are related to adolescents' motivational beliefs [57]. A regression was estimated separately at each grade level and for each gender in each dataset, and by race/ethnicity. We did not test if the regression coefficients replicated across the two datasets, as it is not appropriate to compute the relevant statistics (Cochran's Q and I^2) across regression coefficients [58]. Instead, to compare whether the effects were similar across datasets, gender, and race/ethnicity, we plotted the unstandardized B coefficients with 95% confidence intervals (using unstandardized coefficients is recommended when the predictor variable is easy to understand, which is the case with a single item predictor; [59]). This approach has been recommended to compare the effects [60].

3. Results

3.1. Hypothesis 1: Prevalence of Math Gender Stereotypes

3.1.1. Changes from Early to Late Adolescence

We anticipated that early adolescents would hold more gender egalitarian beliefs (i.e., endorse the belief that both genders are equal), and late adolescents would report more traditional stereotypes (i.e., the belief that boys are better than girls at math). To test this, we investigated age-related changes through repeated measures MANCOVAs (see Supplementary Materials for the results where we controlled for family income and parent education; the results were largely similar with one exception: in MADICS, the time by gender stereotypes main effect was no longer significant). As expected, boys' and girls' stereotypes significantly became more traditional from the eighth to eleventh grades in MADICS [$F(1) = 9.68$, $p = 0.002$, $\eta p^2 = 0.014$] and from the ninth to eleventh grades in HSLS [$F(1) = 510.77$, $p = < 0.001$, $\eta p^2 = 0.022$]. Overall, the longitudinal data supported our hypothesis that gender stereotypes would become more traditional by late adolescence, which was replicated in both datasets.

Next, we tested if these changes varied by gender and race/ethnicity. One of the six interactions was statistically significant. Specifically, the grade by gender interaction was statistically significant in HSLS [$F(1) = 54.40$, $p < 0.001$, $\eta p^2 = 0.002$], indicating that girls had larger increases towards the traditional stereotype over adolescence than boys in HSLS. This interaction was not significant in MADICS [$F(1) = 0.34$, $p = 0.562$]. The grade by race/ethnicity interaction was not significant in either MADICS [$F(3) = 0.095$, $p = 0.758$] or HSLS [$F(3) = 1.258$, $p = 0.287$], and neither was the three-way interaction between grade, gender, and race/ethnicity in MADICS [$F(1) = 0.861$, $p = 0.489$] nor HSLS [$F(1) = 2.612$, $p = 0.074$]. Thus, there were mixed findings concerning whether the changes in adolescents' stereotypes were larger for girls than boys and that there was no evidence that the changes varied by race/ethnicity.

3.1.2. Prevalence in Early Adolescence

We expected that early adolescents, on average, would endorse gender egalitarian beliefs. Unexpectedly, the *t*-tests suggested early adolescents largely favored their own gender as their ratings significantly differed from zero in favor of their own gender. This finding emerged for girls in both datasets and boys in HSLS; the one exception was that boys in MADICS, on average, held gender egalitarian beliefs as their mean did not significantly differ from zero (Table 2).

Additionally, favoring their own gender seemed more pronounced for girls than boys. Girls evidenced small effect sizes, ranging from −0.25 among eighth graders in MADICS to −0.15 among ninth graders in HSLS, with a combined effect size of −0.19 (Figure 1). The size of this effect significantly varied across the two datasets [$Q(1) = 4.537$, $p = 0.033$, $I^2 = 77.958$], with the effect of girls' favoring their own gender being more pronounced in MADICS than HSLS. The effect sizes for boys were negligible, ranging from 0.00 in the eighth grade in MADICS to 0.05 in the ninth grade in HSLS, with a combined effect size of 0.05; the size of the effect was similar across the two datasets [$Q(1) = 2.716$, $p = 0.099$, $I^2 = 63.180$]. Despite the differences in the *t*-tests, many early adolescents endorsed gender egalitarian beliefs. Seventy-nine percent (79%) of early adolescents in MADICS (eighth grade) and 58% of early adolescents in HSLS (ninth grade) endorsed the belief that females and males are equally good at math (see Figure 2a,b).

Table 2. Math gender stereotypes by race and gender: One-sample t-tests investigating whether the means for boys and girls significantly differ from egalitarianism (0).

Dataset	Boys				Girls			
	N	*M* (*SD*)	*t*	*d*	N	*M* (*SD*)	*t*	*d*
Overall								
Early adolescence								
MADICS	601	0.00 (0.71)	0.055	0.00	585	−0.16 (0.63)	−6.664 ***	−0.25
HSLS	11,750	0.05 (0.98)	5.175 ***	0.05	11,750	−0.13 (0.87)	−17.070 ***	−0.15
Combined				0.05				−0.19
Late adolescence								
MADICS	601	0.06 (0.81)	2.216 *	0.09	585	−0.07 (0.77)	−2.742 *	−0.09
HSLS	11,750	0.15 (0.98)	17.833 ***	0.16	11,750	0.08 (0.87)	10.617 ***	0.10
Combined				0.11				0.01
Black adolescents								
Early adolescence								
MADICS	378	−0.01 (0.74)	−0.409	−0.02	330	−0.21 (0.69)	−5.969 ***	−0.30
HSLS	1520	0.02 (1.07)	0.754	0.02	1520	−0.31 (1.02)	−10.412 ***	−0.30
Combined	--	--	--	0.01	--	--	--	−0.30
Late adolescence								
MADICS	378	0.03 (0.88)	0.911	.05	330	−0.10 (0.76)	−2.631 *	−0.15
HSLS	1520	0.06 (1.03)	2.190 *	0.06	1520	−0.07 (1.05)	−2.562 *	−0.08
Combined	--	--	--	0.06	--	--	--	−0.08
White adolescents								
Early adolescence								
MADICS	166	0.04 (0.68)	0.776	0.06	184	−0.06 (0.45)	−2.129 *	−0.16
HSLS	5840	0.05 (1.02)	4.455 ***	0.06	5840	−0.08 (0.85)	−7.952 ***	−0.10
Combined	--	--	--	0.06	--	--	--	−0.10
Late adolescence								
MADICS	166	0.14 (0.79)	2.564 *	0.19	184	0.00 (0.69)	0.099	0.01
HSLS	5840	0.18 (1.00)	15.104 ***	0.19	5840	0.13 (0.84)	11.960 ***	0.16
Combined	--	--	--	0.19	--	--	--	0.10
Asian adolescents								
Early adolescence								
HSLS	470	0.17 (0.94)	6.248 ***	0.20	470	−0.12 (0.83)	−4.656 ***	−0.15
Late adolescence								
HSLS	470	0.28 (0.91)	10.217 ***	0.33	470	0.09 (0.90)	3.321 **	0.11
Latinx adolescents								
Early adolescence								
HSLS	2570	0.03 (0.98)	2.291 *	0.04	2570	−0.19 (0.87)	−9.715 ***	−0.23
Late adolescence								
HSLS	2570	0.04 (1.41)	2.113 *	0.05	2570	0.02 (0.95)	1.112	0.03

0 = both genders equally good at math. 1 = Boys better, −1 = Girls better. All ηp^2 effects were small (0.01–0.13). Gender and race/ethnicity were significant among subject's factors. Pairwise comparisons indicated boys had significantly more traditional math gender stereotypes than girls across race/ethnicity. Additionally, when compared to Black and Latinx adolescents, White and Asian adolescents had significantly more traditional math gender stereotypes. Furthermore, grade level was significant within subject's factor among adolescents (eleventh grade higher than eighth or ninth grades). HSLS SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), Base Year and First Year Follow-Up. *n*'s rounded to the nearest tens place. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

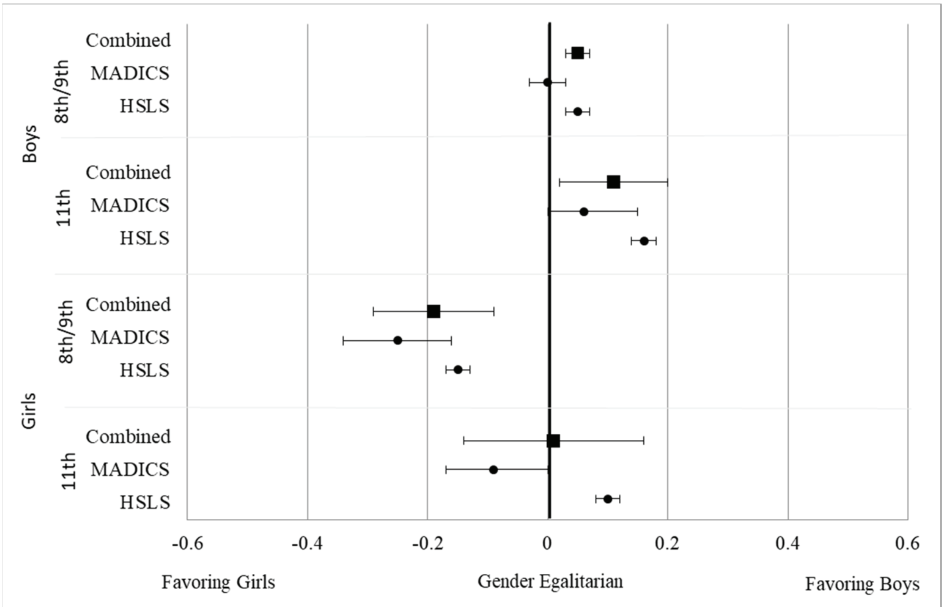
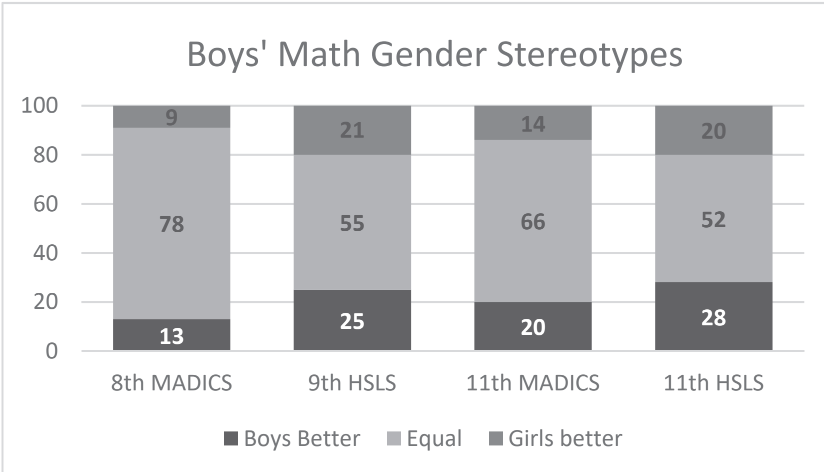


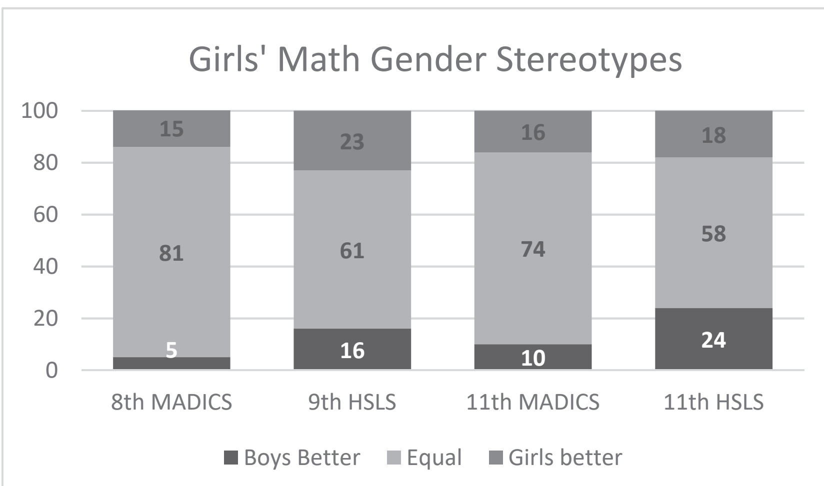
Figure 1. Adolescents' math gender stereotypes: effect sizes (Cohen's d) by grade and gender. Effect sizes to the left indicate bias towards girls/females, and effect sizes to the right indicate bias towards boys/males. HSLS SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), Base Year and First Year Follow-Up.

3.1.3. Prevalence in Late Adolescence

In late adolescence, we expected both genders to endorse the more traditional beliefs that boys are better at math compared to girls. We found evidence for this trend among boys and partial evidence among girls. In both MADICS and HSLS, boys in the eleventh grade were significantly more likely to endorse the traditional gender stereotype (Table 2), with 28% saying boys are better than girls (Figure S1a). The size of these effects for boys was a negligible to small bias towards boys in both datasets, ranging from 0.09 in MADICS and 0.16 in HSLS, with a combined effect size of 0.11 (Figure 1); this effect was significantly larger in HSLS than MADICS [$Q(1) = 6.903$, $p = 0.009$, $I^2 = 85.513$]. The pattern for girls in the eleventh grade varied by dataset. In MADICS, girls still favored their own gender. In HSLS, girls were significantly more likely to say that boys are better than girls at math, with 24% of girls saying boys are better than girls (Figure S1b). Among girls in late adolescence, the effect sizes were negligible, ranging from a slight bias towards girls ($d = -0.09$) in MADICS to a slight bias towards boys ($d = 0.10$) in HSLS, with a combined effect size of 0.01 (Figure 1); these effects significantly varied across the two datasets [$Q(1) = 20.874$, $p < 0.001$, $I^2 = 95.209$]. In sum, late adolescent boys in both datasets and girls in HSLS were significantly more likely to favor the traditional beliefs that boys are better at math than girls; however, the effect sizes were less than small and did not replicate across the two datasets.



(**a**)



(**b**)

Figure 2. (**a**) Boys' math gender stereotypes frequencies. (**b**) Girls' math gender stereotypes frequencies. Frequency of answers to the question "Who's better at math?" by gender and grade. HSLS SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), Base Year and First Year Follow-Up.

3.2. Hypothesis 2: Group Differences in Prevalence

3.2.1. Stereotype Prevalence by Gender

As expected, the MANCOVA gender main effect showed that boys had more traditional math gender stereotypes compared to girls in the two datasets [MADICS: $F(1) = 36.65$, $p < 0.001$, $\eta p^2 = 0.02$; HSLS $F(1) = 170.94$, $p < 0.001$, $\eta p^2 = 0.01$]. The means are shown in Table 2, and the effect sizes are shown in Figure 1, while the frequencies of gender stereotype endorsement by gender at each grade level are presented in Figure S1a,b. In sum, as expected, boys had more traditional math gender stereotypes than girls.

3.2.2. Stereotype Prevalence by Race/Ethnicity

We also tested for race/ethnic differences and the prevalence of gender stereotypes with each racial/ethnic group (Table 2). The MADICS dataset included Black and White adolescents, and HSLS included Asian, Black, Latinx, and White adolescents.

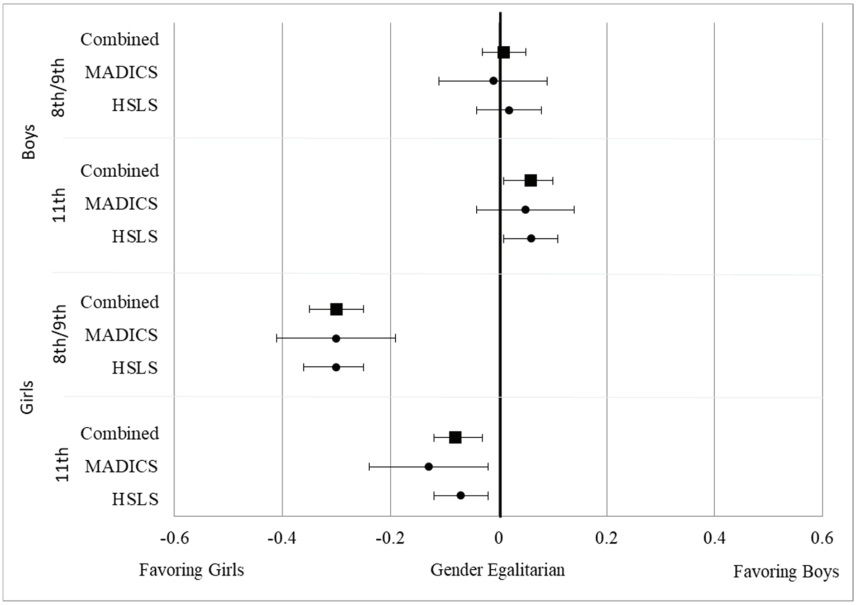
As expected, the racial/ethnic main effect in the MANCOVAs indicated significant differences. In MADICS, White adolescents had significantly more traditional math gender stereotypes than Black adolescents [$F(1) = 11.56$, $p = 0.001$, $\eta p^2 = 0.011$], which aligns with our hypothesis. This finding was replicated in the HSLS dataset, where White and Asian adolescents, who had similarly traditional stereotypes, were more traditional in their beliefs than Black and Latinx adolescents, who had similar non-traditional stereotypes [$F(1) = 46.52$, $p < 0.001$, $\eta p^2 = 0.007$].

Using the MANCOVA, we also examined whether these racial/ethnic differences emerged for boys and girls through the race/ethnicity by gender interaction term. The interaction term was not statistically significant in MADICS [$F(1) = 0.08$, $p = 0.774$], but was in HSLS [$F(3) = 10.53$, $p < 0.001$, $\eta p^2 = 0.002$]. Pairwise comparisons suggest that the racial/ethnic differences varied across boys and girls. Parallel to the racial/ethnic main effect described earlier, Asian boys had the most traditional stereotypes (ninth grade: *p*s < 0.001; eleventh grade: *p*s = 0.003 to < 0.001), followed by White boys (ninth grade: *p*s = 0.134–0.579; eleventh grade: *p*s < 0.001), followed by Black and Latino boys (who did not significantly differ from each other; ninth and eleventh grades: *p*s = 0.791–1.000). In contrast, White girls had more traditional stereotypes than Black and Latina girls (ninth and eleventh grades: *p*s < 0.001), and Asian and Latina girls had more traditional math gender stereotypes than Black girls (ninth grade: *p*s < 0.001; eleventh grade: *p*s = 0.019 to <0.001). Asian girls did not significantly differ from Latina and White girls at either grade level (*p*s = 0.182 to 0.796). The racial/ethnic differences among boys largely parallel the differences found overall, whereas the racial/ethnic differences among girls varied from the overall findings, with White girls having the most traditional beliefs and no differences among Asian and Latina girls.

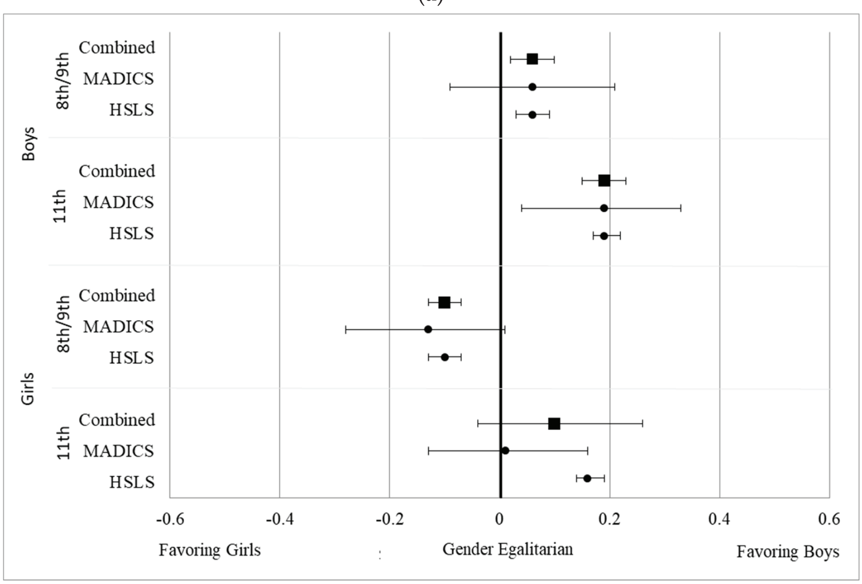
Next, we examined the prevalence of adolescent gender stereotypes within each racial/ethnic group (see Table 2). As expected, White and Asian boys had the most traditional gender stereotypes in early and late adolescence (Asian boys: $d = 0.20$ to 0.33; White boys: $\bar{d} = 0.06$ to 0.19). In contrast, early adolescent Black and Latina girls had the most non-traditional stereotypes, favoring girls (Black girls combined effect: $\bar{d} = -0.30$; Latina girls: $d = -0.23$). By late adolescence, Black girls still held non-traditional stereotypes, favoring girls, but with negligible effect sizes (combined effect: $\bar{d} = -0.08$), and Latina girls were gender egalitarian ($d = 0.03$). Thus, as expected, Asian and White boys were the most traditional, particularly in late adolescence, whereas Black and Latina girls were the least traditional, particularly in early adolescence. Both groups became more traditional across adolescence but did not switch from favoring one gender to the other.

In contrast, Asian and White girls and Black and Latino boys both switched from non-traditional or egalitarian beliefs to traditional beliefs during adolescence. Asian and White girls typically held non-traditional stereotypes in early adolescence (Asian girls: $d = -0.15$; White girls combined: $\bar{d} = -0.10$), but held traditional stereotypes favoring boys in late adolescence, with negligible to small effect sizes ($d = 0.11$ to 0.16). Black boys were gender egalitarian in early adolescence (combined: $\bar{d} = 0.01$), but held traditional stereotypes in late adolescence, with negligible effect sizes ($d = 0.04$ to 0.06). Thus, Asian and White girls and Black and Latino boys both switched to having traditional math gender stereotypes by late adolescence.

White and Black adolescents were included in both datasets, affording us the opportunity to test if the effect sizes replicated across the two datasets (Figure 3a,b). We found that the effect sizes replicated for seven of the eight comparisons, including Black adolescents and White boys in early and late adolescence, and White girls in early adolescence (*p*s = 0.354–0.998). The one exception was White girls in late adolescence [$Q(1) = 3.993$, $p = 0.046$, $I^2 = 74.955$]. In this case, White girls in HSLS held traditional stereotypes, with a small effect size (*d* = 0.16), whereas White girls in MADICS endorsed gender egalitarianism (*d* = 0.01). Thus, the results largely replicated across the datasets for Black and White adolescents.

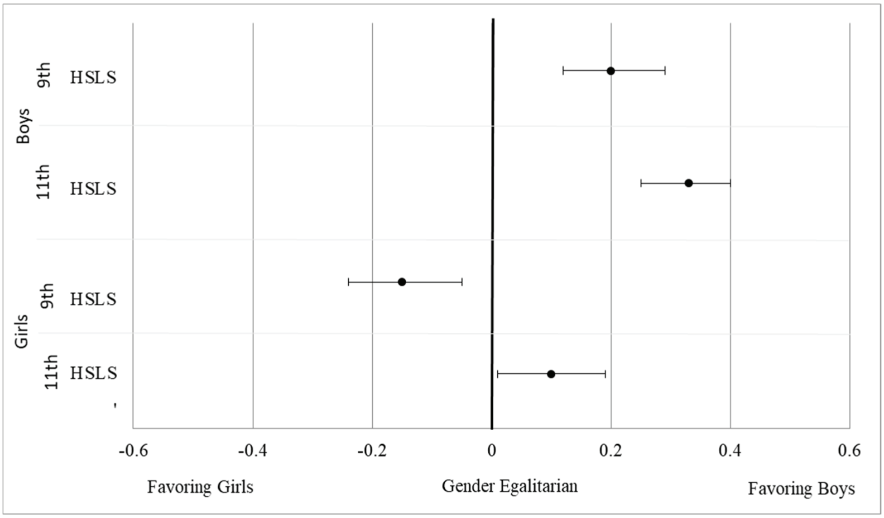


(**a**)

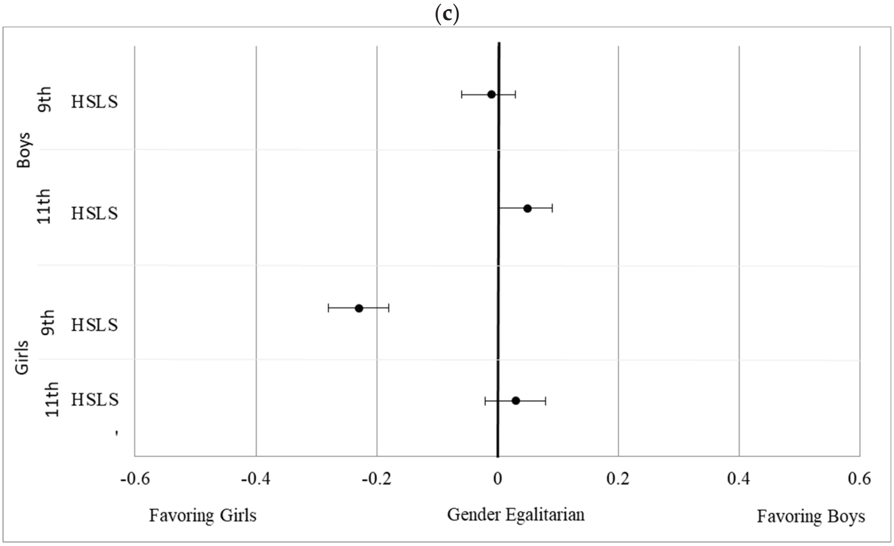


(**b**)

Figure 3. *Cont.*



(**c**)



(**d**)

Figure 3. (**a**) Black adolescents' math gender stereotypes: effect sizes by grade and gender. (**b**) White adolescents' math gender stereotypes: effect sizes by grade and gender. (**c**) Asian adolescents' math gender stereotypes: effect sizes by grade and gender. (**d**) Latinx adolescents' math gender stereotypes: effect sizes by grade and gender.

In conclusion, White adolescents had more traditional stereotypes than Black adolescents that were replicated across datasets. When additionally considering Asian and Latinx adolescents, Asian and White adolescents had more traditional stereotypes than Black and Latinx adolescents. When looking within race and gender groups, math stereotype effect sizes were replicated across the datasets in all cases except among White girls.

3.3. Hypothesis 3: Adolescents' Math Gender Stereotypes and Motivational Beliefs

Regressions were used to examine the associations between adolescents' stereotypes and motivational beliefs while controlling for parent income and education (see Table 3, see

Supplementary Table S2 for the full regression model). We first describe the patterns across all racial/ethnic groups and then separately for each racial/ethnic group. Across the full sample, adolescents' traditional gender math ability stereotypes were positively related to boys' expectancy beliefs (β = 0.23 to 0.27) and negatively related to girls' expectancy beliefs (β = −0.17 to −0.28) in early and late adolescence in both datasets with one exception—the relation was marginally significant for early adolescent girls in MADICS (β = −0.16). Regarding value beliefs, adolescents' stereotypes were positively related to boys' value beliefs (β = 0.10 & 0.20) and negatively related to girls' value beliefs in HSLS (β = −0.10 & −0.15) but were largely not related in MADICS (β = −0.23 to 0.13). Overall, our hypothesis that math gender stereotypes would significantly relate to expectancy and value beliefs was largely supported.

Table 3. Regression analysis coefficients: gender stereotypes and motivational beliefs, by grade and gender.

Predictor	Expectancy Beliefs Predicted by Stereotypes		Value Beliefs Predicted by Stereotypes	
	Boys *B* (SE)	Girls *B* (SE)	Boys *B* (SE)	Girls *B* (SE)
		All adolescents		
Early adolescence				
MADICS	0.23 (0.08) **	−0.16 (0.10) +	0.12 (0.08)	−0.06 (0.11)
HSLS	0.26 (0.02) ***	−0.17 (0.03) ***	0.10 (0.01) ***	−0.10 (0.02) ***
Late adolescence				
MADICS	0.25 (0.11) *	−0.28 (0.14) *	0.13 (0.11)	−0.23 (0.09) *
HSLS	0.27 (0.03) ***	−0.22 (0.03) ***	0.20 (0.02) ***	−0.15 (0.02) ***
		Black adolescents		
Early adolescence				
MADICS	0.18 (0.14)	−0.20 (0.16)	0.06 (0.14)	0.07 (0.20)
HSLS	0.04 (0.06)	−0.11 (0.07)	0.05 (0.05)	−0.06 (0.06)
Late adolescence				
MADICS	0.07 (0.12)	−0.21 (0.17)	0.00 (0.13)	−0.37 (0.17) *
HSLS	0.23 (0.07) ***	−0.29 (0.07) ***	0.16 (0.06) **	−0.01 (0.06)
		White adolescents		
Early adolescence				
MADICS	0.36 (0.16) *	−0.16 (0.20)	−0.12(0.13)	−0.18 (0.21)
HSLS	0.22 (0.03) ***	−0.18 (0.04) ***	0.14 (0.03) **	−0.10 (0.03) **
Late adolescence				
MADICS	0.35 (0.14) *	−0.12 (0.20)	0.21 (0.21)	−0.16 (0.31)
HSLS	0.33 (0.03) ***	−0.23 (0.04) ***	0.21 (0.03) ***	−0.16 (0.03) ***
		Asian adolescents		
Early adolescence				
HSLS	0.24 (0.11) *	−0.11 (0.09)	0.03 (0.09)	0.05 (0.08)
Late adolescence				
HSLS	0.48 (0.12) ***	−0.49 (0.08) ***	0.29 (0.09) ***	−0.24 (0.06) ***
		Latinx adolescents		
Early adolescence				
HSLS	0.23 (0.06) ***	−0.14 (0.06) *	−0.03 (0.05)	−0.10 (0.05) *
Late adolescence				
HSLS	0.14 (0.07) *	−0.04 (0.07)	0.18 (0.05) ***	−0.09 (0.05)

Adolescents' stereotypes and expectancy-value beliefs were all measured at the same time point. Numbers reported are pooled unstandardized beta coefficients (and standard errors) from separate regressions, controlling for parent education and income. HSLS SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School Longitudinal Study of 2009 (HSLS:09), Base Year and First Year Follow-Up. + $p < 0.10$. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Next, we examined these same relations within each racial/ethnic group (see Table 3). The positive relations found between boys' stereotypes and their expectancy beliefs overall replicated/emerged among Asian, Latinx, and White boys (β = 0.14 to 0.48) but were largely not significant for Black boys (β = 0.04 to 0.23). Similarly, the negative relations found between girls' stereotypes and their expectancy beliefs largely replicated across Asian, Black, Latinx, and White girls (β = −0.49 to −0.04). Thus, across races/ethnicities and time points, stereotypes were largely positively related to expectancy beliefs for boys but negatively related for girls. Regarding value beliefs among boys, the relations were largely not significant for boys in early adolescence (β = −0.12 to 0.14) but largely were positively and significantly related in late adolescence for Asian, Black, Latinx, and White boys (β = 0.00 to 0.29). Among girls, value beliefs were significantly and negatively related to stereotypes for White and Latinx girls in early adolescence (β = −0.18 to −0.16) and for Asian, Black, and White girls in late adolescence (β = −0.12 to −0.49). However, the relations during late adolescence for girls were largely not statistically significant for the value beliefs in MADICS. Thus, expectancy beliefs were consistently related to adolescents' math gender stereotypes across race/ethnicities and the datasets in early and late adolescence; however, value beliefs were only consistently related in late adolescence in the HSLS dataset.

We plotted the regression coefficients with 95% confidence intervals to examine whether the coefficients varied across the racial/ethnic groups and datasets (see Supplementary Figures S1a–d and S2a–d). The confidence intervals of the regression coefficients overlapped in all but two cases, both within the HSLS dataset. Among early adolescent boys, the positive relation between math gender stereotypes and value beliefs was larger for White boys than Latino boys. Additionally, among late adolescent girls, the negative relation between math gender stereotypes and expectancy beliefs was larger among Asian girls than White and Latinx girls. Asian and White boys had large, overlapping positive relations for expectancy beliefs, whereas girls' across races/ethnicities had negative relations of similar strength for expectancy beliefs in early adolescence. In sum, relations between gender stereotypes and motivational beliefs largely overlapped across Asian, Black, Latinx, and White adolescents.

4. Discussion

Gender gaps continue to persist in STEM, despite psychologists' attempts to study and minimize these gaps [1,2]. One potential reason for gender gaps is the traditional stereotype that boys are better at math than girls [1,4]. However, existing research has not systemically charted the changes in adolescents' stereotypes nor if these normative developmental processes vary across the four largest racial/ethnic groups in the U.S. Furthermore, much of the existing research focuses on a single dataset, and few studies examine replication or historical differences. Using two longitudinal datasets, we examined four primary research questions. First, how do individuals' gender stereotypes about math ability change across adolescence? Second, do these changes vary based on gender and race/ethnicity? Third, are adolescents' math ability gender stereotypes related to their math expectancy and value beliefs? Finally, do these findings replicate across the two datasets that spanned over 20 years? Findings from this study make several contributions in terms of extending the existing literature and implications for theory and applied work.

4.1. The Prevalence and Changes in Gender Stereotypes

We found that individuals' traditional stereotypes increased from early to late adolescence for all gender and racial/ethnic groups in both datasets. We know from prior research that children favor their own gender [24], whereas parents endorse traditional math gender stereotypes [7]. The current findings suggest that adolescence is a period characterized by changes in math ability stereotype beliefs and that individuals' more traditional beliefs may emerge by late adolescence and persist well into adulthood. Gender stereotypes about math, paired with cultural values that emphasize "following your passions", may result in adolescents relying on stereotypes to choose majors and careers, resulting in STEM gender gaps [61]. However, it is important to note that although math gender stereotypes moved toward being more traditional for all racial/ethnic groups, Black and Latina girls did not hold traditional stereotypes in late adolescence. Instead, they held gender egalitarian or non-traditional views. These findings help discern the prior mixed cross-sectional work on age differences. Thus, understanding when increases occur for each group defined by the intersection of race/ethnicity and gender is an important contribution, as it allows for a deeper understanding of math gender stereotype development and indicates when interventions might be most helpful and for whom.

Increases towards the traditional stereotype were theorized to occur in adolescence for multiple reasons. First, adolescents have the cognitive capacity to connect their social groups' stereotypes to themselves [20,62,63]. Second, because adolescence is a period of intense identity development, adolescents have the motivation to relate these social group stereotypes to their own identity and motivational beliefs [18,21]. Adolescents look to important people in their lives and society more broadly to help them understand their social groups' identities and, in turn, decide who they are [50]. Finally, many adolescents take advanced STEM coursework in late high school, and teacher practices in advanced math courses, such as allowing adolescents to shout out answers, may privilege boys while silencing girls [22]. Thus, although a significant portion of advanced math courses is comprised of girls (predominantly Asian and White girls), research suggests that girls are treated differently than boys in advanced math courses, which helps to explain why gender stereotypes about math and gender gaps in STEM persist, despite having a larger representation [22]. In sum, we found that traditional stereotypes increased from early to late adolescence in both datasets. These more traditional stereotypes have been found to persist through adulthood [7,10], helping us to understand the developmental time point when traditional math gender stereotypes begin to emerge.

Racial/Ethnic and Gender Comparisons

As hypothesized, Black and Latinx adolescents had less traditional stereotypes compared to Asian and White adolescents, and girls had less traditional stereotypes than boys. When examining the intersection of gender and race/ethnicity, the findings suggest that Black girls had the least traditional math gender stereotypes, whereas Asian and White boys had the most traditional math gender stereotypes, which aligned with our hypothesis. Other studies among Asian, Black, Latinx, and White U.S. high school students [10] and among Black U.S. families with adolescents and children [16,35] report similar findings. Social status theory [16] suggests that Black and Latinx adolescents and girls may be less likely to endorse traditional gender stereotypes, as disenfranchised groups have less to gain by upholding the status quo and may be more aware of social justice issues. Overall, our study indicates that Black and Latinx adolescents, particularly Black girls, were significantly less likely to endorse traditional math gender stereotypes when compared to White and Asian adolescents, particularly White and Asian boys.

The finding that Black adolescents demonstrated less traditional stereotypes compared to White adolescents replicated across the two datasets, which spanned nearly 20 years and used different methodologies (e.g., a nationally representative study versus a local study based in Maryland/D.C.). The one exception was that the prevalence of White girls' stereotypes in the eleventh grade did not replicate across the datasets. White girls held gender egalitarian beliefs in MADICS (collected in 1996), but traditional stereotypes favoring boys as better at math in HSLS (collected in 2012). Though one might expect girls' math ability stereotypes might become less traditional over historical time, this difference may have also resulted from methodological differences. For example, HSLS is a large national sample, whereas MADICS is a smaller, local sample from an urban area near Washington, D.C. Prior research finds that people in urban areas tend to have more progressive views than the rest of the country, which may have resulted in less traditional math gender stereotypes among White girls in the MADICS dataset [64]. However, it is unclear why this difference across datasets was only observed among White girls.

Our study is one of the first to examine math gender stereotypes among Asian and Latinx adolescents. Our findings suggest that a culture's gender role beliefs may not translate into gender stereotypes about specific academic subjects, such as math. Both Asian and Latinx cultures espouse more traditional gender role beliefs [36]; however, these two groups differ dramatically concerning their representation in STEM, with Asians having stronger representation and being stereotyped to do well in math with the model minority hypothesis [40]. Aligned with the patterns in STEM, Asian boys in HSLS endorsed the strongest traditional stereotypes, whereas Latinx adolescents were more gender egalitarian. This finding also supports social status theory [16]; Asian adolescents may be more likely to have something to gain by supporting stereotypes related to STEM, given their high representation in the STEM fields and advanced STEM classrooms, whereas Latinx adolescents may be less likely to buy into math-related stereotypes given their underrepresentation in STEM.

4.2. Adolescents' Math Gender Stereotypes and Their Expectancy and Value Beliefs

As hypothesized, boys' stereotypes positively related to their expectancy beliefs, whereas girls' traditional stereotypes negatively related to their expectancy beliefs. When examined by race/ethnicity, these relations were less prevalent among Black adolescents compared to other racial/ethnic groups; however, generally, the confidence intervals of the effect sizes overlapped between racial/ethnic groups within the same gender, suggesting similar effects. The relations with adolescents' math value beliefs followed a similar pattern but were less consistent. One explanation for why significant relations were less common for value beliefs than for expectancy beliefs is that the gender stereotypes focused on stereotypes about math ability, not value. Given that expectancy and value beliefs about math are related but separate constructs [14], it follows that math expectancy beliefs would be more strongly related to a stereotype specifically about math ability. Future research might also explore value-related stereotypes in conjunction with ability stereotypes to examine whether value beliefs are more strongly related to the former.

Moreover, the relations between math ability gender stereotypes and expectancy-value beliefs emerged in both datasets spanning nearly 20 years and evidenced effects that were similar in size. Believing that boys are better at math than girls was positively related to the boys' motivational beliefs and negatively related to the girls' motivational beliefs in each dataset. Thus, stereotypes about gender and math ability may contribute to the persistent gender gap in math expectancy beliefs [2]. These findings provide support for situated expectancy-value theory [14] as well as the potential value of using archived datasets, particularly when combined with newer nationally representative datasets.

4.3. Practical Implications

This research has several practical implications. Given that we found adolescence may be a period when gender stereotypes become more traditional, parents and high school teachers should try to counteract such trends, particularly among boys and Asian and White adolescents. Such interventions might counteract traditional stereotypes and highlight that boys and girls perform similarly in math. Furthermore, they might discuss the gender gaps in math self-confidence and expectancy beliefs, explaining that due to math gender stereotypes, boys might be inclined to believe they are doing better in math than they are, whereas girls might believe they are doing worse than they actually are. Such a strategy may work to reduce gender stereotypes and provide adolescents an accurate understanding of their math abilities. Youth in adolescence may be especially interested in fairness and social justice issues [65], which may make it an ideal time for teachers and parents to broach the topic of equity and inclusion in STEM.

It is also important to note that many adolescents, particularly Black and Latinx adolescents, endorsed the belief that boys and girls are equally good in math. In early adolescence, Black and Latina girls believed, on average, that girls were better than boys in math. Prior research has demonstrated that a strength of many Black families is their encouragement of competence and self-reliance among their daughters [17]. Our findings suggest that Black individuals carry this strength with them into STEM, where Black families and adolescents were more likely to endorse counter-traditional gender stereotypes about math ability and be especially supportive of their daughters in STEM. Researchers and practitioners should learn from the positive parenting practices among Black families to better support girls and other groups who are marginalized and help promote culturally responsive practices in classrooms as well as home-based learning among other racial/ethnic groups.

This study found that the pattern of late adolescents favoring boys in math has not significantly changed in 20 years, and prior studies have similarly found that parent stereotypes favoring boys and gender gaps in math motivational beliefs have not changed in decades [7]. Furthermore, recent research indicates that gender gaps in math self-concept have persisted for the past 23 years [2]. We need to do more as researchers, teachers, parents, and as a society. However, changing adolescents' math gender stereotypes at the level of the individual is only a small piece of creating more gender equality in STEM. For example, even though Black girls have the least traditional stereotypes, they are still underrepresented in STEM. There are many structural issues at the high school and college level, such as teacher practices that privilege boys and men and bias towards girls and women in STEM classrooms [22]. These structural issues must also be addressed to create greater gender equality in STEM.

4.4. Future Directions and Limitations

This study was unique in that it implemented two large datasets to investigate gender stereotypes. This approach has made replication possible; however, it also has limitations. The datasets differed demographically, and due to limited racial and ethnic diversity in the MADICS dataset, replication by race/ethnicity could only be examined among White and Black participants. Additionally, the people who identify as these pan-ethnic and racial categories are diverse, and the present study did not distinguish between important within-group ethnic variations. Furthermore, one dataset was nationally representative, whereas the other was a local dataset, making historical comparisons more difficult. Each dataset only had one item to measure gender stereotypes, which does not allow us to examine reliability or invariance over time. Furthermore, the MADICS gender stereotype question asked about both math and science, whereas the HSLS stereotype question focused on math. Despite this, we largely found that findings across the two datasets replicated, indicating that adolescents may hold similar gender stereotypes about math and science when it comes to ability. Relatedly, future studies might ask about STEM subjects beyond math. Relative to fields like computer science and engineering, math has greater gender parity [66]; however, individuals may hold stronger male bias regarding fields like computer science when compared to math [67]. However, we believe that assessing math gender stereotypes is still important, given that math is a gateway to these fields. We found that stereotypes became more traditional from early to late adolescence. Future studies might explore the mechanisms related to this increase, such as taking advanced math courses. Math gender stereotype increases may occur during late high school, and these beliefs may continue into adulthood. However, more longitudinal work would need to be conducted to better understand stereotype development.

This study investigated differences by race/ethnicity and gender, finding that underrepresented groups in STEM (girls, Black, and Latinx individuals) had significantly less traditional stereotypes. Future studies might also explore the differences related to SES and parent education; in the present study, we found some differences when we controlled for family income and parent education. Given that low-income and first-generation adolescents are also underrepresented in STEM, they may also be less likely to endorse math gender stereotypes. However, this might also differ at the intersections of race and gender (e.g., low-income White boys may still feel they have something to gain by upholding traditional stereotypes). Finally, both datasets used in this study are older (collected in the 1990s through to the 2010s); more recent data might further explore similar research questions to see if these effects hold up today. Given that gender gaps in STEM still remain [1], we believe our findings are relevant, despite the data being slightly older. Despite these limitations, the findings of this study contributed to the literature in several significant ways. The findings that replicated provide more confidence, as these patterns emerged consistently at the conceptual level and provided insight into areas of historical changes and stability. These findings also point to several new directions for further study, including racial/ethnic differences.

5. Conclusions

The present study makes several contributions to important areas. First, this study examines math ability gender stereotypes among adolescents across multiple datasets spanning the 1990s to the 2010s, allowing for replication and testing of potential historical differences. Second, this study allowed us to examine the developmental changes in individuals' math gender stereotypes and their motivational correlates from early to late adolescence, including differences based on race/ethnicity. Adolescents' math gender stereotypes became more traditional from early to late adolescence. This small increase towards believing boys are better in math than girls in late adolescence may be one reason why, and despite the many societal changes in regard to gender and similar math grades and test scores, math motivational beliefs have still not reached gender parity. However, many participants endorsed egalitarian beliefs or favored girls as better, especially among early adolescent girls and Black/Latinx adolescents. In particular, Black girls had non-traditional or gender egalitarian beliefs, highlighting a strength among Black girls and their families that future research might learn from.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/educsci13090866/s1.

Author Contributions: Conceptualization, C.R.S. and S.D.S.; methodology, C.R.S. and S.D.S.; formal analysis, C.R.S.; data curation, C.R.S., A.-L.D. and N.S.; writing—original draft preparation, C.R.S., Y.G., C.R. and G.L.; writing—review and editing, primarily C.R.S. and S.D.S. with contributions from all authors; supervision, S.D.S. and J.S.E.; project administration, S.D.S.; funding acquisition, S.D.S. and J.S.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation grant numbers [DRL-1760757 and DRL-2054956] awarded to Sandra Simpkins and Jacquelynne Eccles. The original data collection of MADICS was supported by funding from the MacArthur Research Network on Successful Adolescent Development in High-Risk Settings and funding from NICHD Grant R01 033437.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of University of California, Irvine (protocol HS#2018-4349)" for studies involving humans.

Informed Consent Statement: The present study used deidentified, secondary human participant data.

Data Availability Statement: Data and material is available at the following websites. HSLS: https://nces.ed.gov/surveys/hsls09/ (accessed on 21 August 2023); MADICS: https://garp.education.uci.edu/msalt.html (accessed on 21 August 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Breda, T.; Jouini, E.; Napp, C.; Thebault, G. Gender Stereotypes Can Explain the Gender-Equality Paradox. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 31063–31069. [CrossRef] [PubMed]
2. Rubach, C.; Lee, G.; Starr, C.R.; Gao, Y.; Safavian, N.; Dicke, A.-L.; Eccles, J.S.; Simpkins, S.D. Is There Any Evidence of Historical Changes in Gender Differences in American High School Students' Math Competence-Related Beliefs from the 1980s to the 2010s? *Int. J. Gend. Sci. Technol.* **2022**, *14*, 55–126.
3. Rickles, J.; Heppen, J.; Taylor, S.; Sorensen, N.; Walters, K.; Clements, P. *Course Progression for Students Who Fail Algebra I in Ninth Grade*; American Institutes for Research: Arlington County, VA, USA, 2017; Available online: https://www.air.org/sites/default/files/2022-03/Course-Progression-for-Students-Who-Fail-Algebra-I-in-Ninth-Grade-June-2017.pdf (accessed on 15 April 2023).
4. Regner, I.; Steele, J.R.; Ambady, N.; Thinus-Blanc, C.; Huguet, P. Our Future Scientists: A Review of Stereotype Threat in Girls from Early Elementary School to Middle School. *Rev. Int. De Psychol. Soc.* **2014**, *27*, 13–51.

5. Ceci, S.J.; Ginther, D.K.; Kahn, S.; Williams, W.M. Women in Academic Science: A Changing Landscape. *Psychol. Sci. Public. Interest.* **2014**, *15*, 75–141. [CrossRef]
6. Long, M.; Steinke, J.; Applegate, B.; Knight Lapinski, M.; Johnson, M.J.; Ghosh, S. Portrayals of Male and Female Scientists in Television Programs Popular Among Middle School-Age Children. *Sci. Commun.* **2010**, *32*, 356–382. [CrossRef]
7. Starr, C.R.; Ramos Carranza, P.; Simpkins, S.D. Stability and Changes in High School Students' STEM Career Expectations: Variability Based on STEM Support and Parent Education. *J. Adolesc.* **2022**, *94*, 906–919. [CrossRef]
8. World Economic Forum. Global Gender Gap Report. 2023. Available online: https://www.weforum.org/reports/global-gender-gap-report-2023/ (accessed on 16 August 2023).
9. McGuire, L.; Mulvey, K.L.; Goff, E.; Irvin, M.J.; Winterbottom, M.; Fields, G.E.; Hartstone-Rose, A.; Rutland, A. STEM Gender Stereotypes from Early Childhood through Adolescence at Informal Science Centers. *J. Appl. Dev. Psychol.* **2020**, *67*, 101109. [CrossRef]
10. Starr, C.R.; Simpkins, S.D. High School Students' Math and Science Gender Stereotypes: Relations with Their STEM Outcomes and Socializers' Stereotypes. *Soc. Psychol. Educ.* **2021**, *24*, 273–298. [CrossRef]
11. Seo, E.; Shen, Y.; Alfaro, E.C. Adolescents' Beliefs about Math Ability and Their Relations to STEM Career Attainment: Joint Consideration of Race/Ethnicity and Gender. *J. Youth Adolesc.* **2019**, *48*, 306–325. [CrossRef]
12. Maxwell, S.E.; Lau, M.Y.; Howard, G.S. Is Psychology Suffering from a Replication Crisis? What Does "Failure to Replicate" Really Mean? *Am. Psychol.* **2015**, *70*, 487–498. [CrossRef]
13. Plucker, J.A.; Makel, M.C. Replication Is Important for Educational Psychology: Recent Developments and Key Issues. *Educ. Psychol.* **2021**, *56*, 90–100. [CrossRef]
14. Eccles, J.S.; Wigfield, A. From Expectancy-Value Theory to Situated Expectancy-Value Theory: A Developmental, Social Cognitive, and Sociocultural Perspective on Motivation. *Contemp. Educ. Psychol.* **2020**, *61*, 101859. [CrossRef]
15. Plante, I.; O'Keefe, P.A.; Aronson, J.; Fréchette-Simard, C.; Goulet, M. The Interest Gap: How Gender Stereotype Endorsement about Abilities Predicts Differences in Academic Interests. *Soc. Psychol. Educ.* **2019**, *22*, 227–245. [CrossRef]
16. Rowley, S.J.; Kurtz-Costes, B.; Mistry, R.; Feagans, L. Social Status as a Predictor of Race and Gender Stereotypes in Late Childhood and Early Adolescence. *Soc. Social. Dev.* **2007**, *16*, 150–168. [CrossRef]
17. Black, A.R.; Peacock, N. Pleasing the Masses: Messages for Daily Life Management in African American Women's Popular Media Sources. *Am. J. Public. Health* **2011**, *101*, 144–150. [CrossRef] [PubMed]
18. Erikson, E.H. *Identity: Youth and Crisis*; Norton & Co.: Oxford, UK, 1968.
19. Patterson, M.M.; Bigler, R.S. Effects of Consistency between Self and In-Group on Children's Views of Self, Groups, and Abilities. *Soc. Social. Dev.* **2018**, *27*, 154–171. [CrossRef]
20. Umaña-Taylor, A.J.; Quintana, S.M.; Lee, R.M.; Cross, W.E., Jr.; Rivas-Drake, D.; Schwartz, S.J.; Syed, M.; Yip, T.; Seaton, E.; Ethnic and Racial Identity in the 21st Century Study Group. Ethnic and Racial Identity During Adolescence and Into Young Adulthood: An Integrated Conceptualization. *Child. Dev.* **2014**, *85*, 21–39. [CrossRef]
21. Eccles, J. Who Am I and What Am I Going to Do With My Life? Personal and Collective Identities as Motivators of Action. *Educ. Psychol.* **2009**, *44*, 78–89. [CrossRef]
22. Musto, M. Brilliant or Bad: The Gendered Social Construction of Exceptionalism in Early Adolescence. *Am. Sociol. Rev.* **2019**, *84*, 369–393. [CrossRef]
23. Steinke, J.; Lapinski, M.K.; Crocker, N.; Zietsman-Thomas, A.; Williams, Y.; Evergreen, S.H.; Kuchibhotla, S. Assessing Media Influences on Middle School–Aged Children's Perceptions of Women in Science Using the Draw-A-Scientist Test (DAST). *Sci. Commun.* **2007**, *29*, 35–64. [CrossRef]
24. Kurtz-Costes, B.; Copping, K.E.; Rowley, S.J.; Kinlaw, C.R. Gender and Age Differences in Awareness and Endorsement of Gender Stereotypes about Academic Abilities. *Eur. J. Psychol. Educ.* **2014**, *29*, 603–618. [CrossRef]
25. Cvencek, D.; Meltzoff, A.N.; Greenwald, A.G. Math–Gender Stereotypes in Elementary School Children. *Child. Dev.* **2011**, *82*, 766–779. [CrossRef] [PubMed]
26. Chatard, A.; Guimond, S.; Selimbegovic, L. "How Good Are You in Math?" The Effect of Gender Stereotypes on Students' Recollection of Their School Marks. *J. Exp. Soc. Social. Psychol.* **2007**, *43*, 1017–1024. [CrossRef]
27. Steffens, M.C.; Jelenec, P. Separating Implicit Gender Stereotypes Regarding Math and Language: Implicit Ability Stereotypes Are Self-Serving for Boys and Men, but Not for Girls and Women. *Sex. Roles* **2011**, *64*, 324–335. [CrossRef]
28. Hargreaves, M.; Homer, M.; Swinnerton, B. A Comparison of Performance and Attitudes in Mathematics amongst the 'Gifted'. Are Boys Better at Mathematics or Do They Just Think They Are? *Assess. Educ. Princ. Policy Pract.* **2008**, *15*, 19–38. [CrossRef]
29. Plante, I.; Théorêt, M.; Favreau, O.E. Student Gender Stereotypes: Contrasting the Perceived Maleness and Femaleness of Mathematics and Language. *Educ. Psychol.* **2009**, *29*, 385–405. [CrossRef]
30. Denner, J.; Laursen, B.; Dickson, D.; Hartl, A.C. Latino Children's Math Confidence: The Role of Mothers' Gender Stereotypes and Involvement Across the Transition to Middle School. *J. Early Adolesc.* **2018**, *38*, 513–529. [CrossRef]
31. Ward, L.M.; Grower, P. Media and the Development of Gender Role Stereotypes. *Annu. Rev. Dev. Psychol.* **2020**, *2*, 177–199. [CrossRef]
32. Bieg, M.; Goetz, T.; Wolter, I.; Hall, N. Gender Stereotype Endorsement Differentially Predicts Girls' and Boys' Trait-State Discrepancy in Math Anxiety. *Front. Psychol.* **2015**, *6*, 1404. [CrossRef]

33. Hyde, J.S.; Fennema, E.; Lamon, S.J. Gender Differences in Mathematics Performance: A Meta-Analysis. *Psychol. Bull.* **1990**, *107*, 139–155. [CrossRef]
34. Skinner, O.D.; Kurtz-Costes, B.; Vuletich, H.; Copping, K.; Rowley, S.J. Race Differences in Black and White Adolescents' Academic Gender Stereotypes across Middle and Late Adolescence. *Cult. Divers. Ethn. Minor. Psychol.* **2021**, *27*, 537–545. [CrossRef]
35. Evans, A.B.; Copping, K.E.; Rowley, S.J.; Kurtz-Costes, B. Academic Self-Concept in Black Adolescents: Do Race and Gender Stereotypes Matter? *Self Identity* **2011**, *10*, 263–277. [CrossRef] [PubMed]
36. Gutierrez, B.C.; Halim, M.L.D.; Leaper, C. Variations in Recalled Familial Messages about Gender in Relation to Emerging Adults' Gender, Ethnic Background, and Current Gender Attitudes. *J. Fam. Stud.* **2019**, *28*, 150–183. [CrossRef]
37. O'Brien, L.T.; Blodorn, A.; Adams, G.; Garcia, D.M.; Hammer, E. Ethnic Variation in Gender-STEM Stereotypes and STEM Participation: An Intersectional Approach. *Cult. Divers. Ethn. Minor. Psychol.* **2015**, *21*, 169–180. [CrossRef] [PubMed]
38. Starr, C.R. "I'm Not a Science Nerd!": STEM Stereotypes, Identity, and Motivation Among Undergraduate Women. *Psychol. Women Q* **2018**, *42*, 489–503. [CrossRef]
39. Kurtz-Costes, B.; Rowley, S.J.; Harris-Britt, A.; Woods, T.A. Gender Stereotypes about Mathematics and Science and Self-Perceptions of Ability in Late Childhood and Early Adolescence. *Merrill-Palmer Q* **2008**, *54*, 386–409. [CrossRef]
40. McGee, E. "Black Genius, Asian Fail": The Detriment of Stereotype Lift and Stereotype Threat in High-Achieving Asian and Black STEM Students. *AERA Open* **2018**, *4*, 16. [CrossRef]
41. Hsieh, T.; Simpkins, S.D.; Eccles, J.S. Gender by Racial/Ethnic Intersectionality in the Patterns of Adolescents' Math Motivation and Their Math Achievement and Engagement. *Contemp. Educ. Psychol.* **2021**, *66*, 101974. [CrossRef]
42. Gibson, C.E.; Losee, J.; Vitiello, C. Replication Attempt of Stereotype Susceptibility (Shih, Pittinsky, & Ambady, 1999): Identity Salience and Shifts in Quantitative Performance. *Soc. Psychol.* **2014**, *45*, 194–198. [CrossRef]
43. Passolunghi, M.C.; Rueda Ferreira, T.I.; Tomasetto, C. Math–Gender Stereotypes and Math-Related Beliefs in Childhood and Early Adolescence. *Learn. Individ. Differ.* **2014**, *34*, 70–76. [CrossRef]
44. National Center for Statistics (NCES). *High School Longitudinal Study of 2009 User Manuals*; Institute of Education Sciences: Washington, DC, USA, 2019. Available online: https://nces.ed.gov/surveys/hsls09/usermanuals.asp (accessed on 23 February 2023).
45. Eccles, J.S.; Jacobs, J.E.; Harold, R.D. Gender Role Stereotypes, Expectancy Effects, and Parents' Socialization of Gender Differences. *J. Soc. Social. Issues* **1990**, *46*, 183–201. [CrossRef]
46. Fisher, G.G.; Matthews, R.A.; Gibbons, A.M. Developing and Investigating the Use of Single-Item Measures in Organizational Research. *J. Occup. Health Psychol.* **2016**, *21*, 3–23. [CrossRef] [PubMed]
47. Hays, R.D.; Reise, S.; Calderón, J.L. How Much Is Lost in Using Single Items? *J. Gen. Intern. Med.* **2012**, *27*, 1402–1403. [CrossRef] [PubMed]
48. Nichols, A.L.; Webster, G.D. The Single-Item Need to Belong Scale. *Personal. Individ. Differ.* **2013**, *55*, 189–192. [CrossRef]
49. Eccles, J.S.; Wigfield, A. In the Mind of the Actor: The Structure of Adolescents' Achievement Task Values and Expectancy-Related Beliefs. *Pers. Soc. Psychol. Bull.* **1995**, *21*, 215–225. [CrossRef]
50. Lauermann, F.; Tsai, Y.-M.; Eccles, J.S. Math-Related Career Aspirations and Choices within Eccles et al.'s Expectancy–Value Theory of Achievement-Related Behaviors. *Dev. Psychol.* **2017**, *53*, 1540–1559. [CrossRef]
51. Jiang, S.; Simpkins, S.D.; Eccles, J.S. Individuals' Math and Science Motivation and Their Subsequent STEM Choices and Achievement in High School and College: A Longitudinal Study of Gender and College Generation Status Differences. *Dev. Psychol.* **2020**, *56*, 2137–2151. [CrossRef]
52. Diemer, M.A.; Marchand, A.D.; McKellar, S.E.; Malanchuk, O. Promotive and Corrosive Factors in African American Students' Math Beliefs and Achievement. *J. Youth Adolesc.* **2016**, *45*, 1208–1225. [CrossRef]
53. Enders, C.K. *Applied Missing Data Analysis*; Guilford Press: New York, NY, USA, 2010; Chapter xv; 377p.
54. Borenstein, M. *Comprehensive Meta-Analysis Software*; John Wiley & Sons Ltd: Hoboken, NJ, USA, 2022; pp. 535–548. [CrossRef]
55. Borenstein, M.; Hedges, L.; Higgins, J.; Rothstein, H. *Introduction to Meta-Analysis*; John Wiley & Sons Ltd: Hoboken, NJ, USA, 2021. [CrossRef]
56. Hedges, L.V.; Schauer, J.M. Statistical Analyses for Studying Replication: Meta-Analytic Perspectives. *Psychol. Methods* **2019**, *24*, 557–570. [CrossRef]
57. Schoon, I.; Eccles, J.S. *Gender Differences in Aspirations and Attainment: A Life Course Perspective*; Cambridge University Press: New York, NY, USA, 2014. [CrossRef]
58. Hunter, J.E.; Schmidt, F.L. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings/Frank L. Schmidt, University of Iowa, John E. Hunter, Michigan State University*, 3rd ed.; SAGE: Thousand Oaks, CA, USA, 2015.
59. Baguley, T. Standardized or Simple Effect Size: What Should Be Reported? *Br. J. Psychol.* **2009**, *100*, 603–617. [CrossRef]
60. Thompson, B. Effect Sizes, Confidence Intervals, and Confidence Intervals for Effect Sizes. *Psychol. Sch.* **2007**, *44*, 423–432. [CrossRef]
61. Siy, J.O.; Germano, A.L.; Vianna, L.; Azpeitia, J.; Yan, S.; Montoya, A.K.; Cheryan, S. Does the Follow-Your-Passions Ideology Cause Greater Academic and Occupational Gender Disparities than Other Cultural Ideologies? *J. Personal. Soc. Social. Psychol.* **2023**, *125*, 548–570. [CrossRef] [PubMed]
62. Marcia, J.E. The Empirical Study of Ego Identity. In *Identity and development: An Interdisciplinary Approach*; Bosma, H.A., Graafsma, T.L.G., Grotevant, H.D., de Levita, D.J., Eds.; Sage Publications, Inc.: Thousand Oaks, CA, USA, 1994; pp. 67–80, Chapter xii; 204p.

63. Bigler, R.S.; Patterson, M.M. Social Stereotyping and Prejudice in Children. In *The Wiley Handbook of Group Processes in Children and Adolescents*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2017; pp. 184–202. [CrossRef]
64. Lyons, J.; Utych, S.M. You're Not From Here!: The Consequences of Urban and Rural Identities. *Polit. Behav.* **2021**, *45*, 75–101. [CrossRef]
65. Brown, C.S.; Bigler, R.S. Children's Perceptions of Discrimination: A Developmental Model. *Child. Dev.* **2005**, *76*, 533–553. [CrossRef] [PubMed]
66. Miller, D.; Tanenbaum, C. *NSF Award Search: Award # 1920401—The Development of Gender Stereotypes About STEM Abilities: A Meta-Analysis*; National Science Foundation: Washington DC, USA, 2019. Available online: https://nsf.gov/awardsearch/showAward?AWD_ID=1920401 (accessed on 23 February 2023).
67. Master, A.; Meltzoff, A.N.; Cheryan, S. Gender Stereotypes about Interests Start Early and Cause Gender Disparities in Computer Science and Engineering. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2100030118. [CrossRef]

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Education Sciences Editorial Office
E-mail: education@mdpi.com
www.mdpi.com/journal/education

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-2448-9