ATLANTIS PRESS

Advances in Social Science, Education and Humanities Research, volume 615

Proceedings of the 2021 4th International Conference on Humanities Education and Social Sciences (ICHESS 2021)

A Systematic Review About Gender Difference in STEM Performance for Adolescents

Ke Sun[1], Xinwei Bian[2,*]

[1] *Moray House of Education and Sport, The University of Edinburgh, EH88JR, Edinburgh, United Kingdom*
[2] *Faculty of Social Science & Public Policy, King's College London, WC2R2LS, London, United Kingdom*
[*]*Corresponding author. Email: K1814423@kcl.ac.uk*

ABSTRACT

This study aims to synthesize the current qualitative studies associated with the reason why there is gender differences in STEM performance. In order to explore this question, the authors searched 4 databases: ERIC, EBSCO, Child Development & Adolescent Studies and PsycINFO, and obtained 13 studies to analyze. This essay selects a bio-psycho-social model to investigate the gender difference in STEM performance, which attempted to bring together the disparate findings from the perspective from cognition, psychology, and sociocultural environment. The results demonstrate that there were three factors related to gender difference in STEM performance, including cognitive skills, psychological factors and sociocultural effects. The three factors not only affect gender difference in STEM performance respectively, but also interact with each other and conduce in gender difference in STEM grades. The authors also provide several suggestions for future practice. All participants including parents, teachers and students themselves ought to pay attention to the phenomenon.

Keywords: *Gender differences, STEM education, Gender stereotypes, Self-efficacy*

1. INTRODUCTION

Gender has been a focus in science, technology, engineering and mathematics education (STEM) [1]. Some studies indicated that there was no difference between females and males in STEM performance [2-5], while numerous studies have found that males have outperformed females [6,7]. Researchers and policy makers considered that gender stereotype was one of the effective factors contributing to gender differences in STEM fields [8]. Existing literature shows that there are some factors relating to gender differences and STEM performance, such as problem solving, mathematical thinking abilities, mathematics anxiety, mathematical attitudes and self-efficacy [8-11].

Considering there has been numerous studies concentrating on gender differences in STEM performance, it is essential to review and summarize existing literature. In this study, the authors systematically review current studies, synthesize relevant literature about gender difference in STEM performance among adolescents. In addition, recommendations for future practice and research will also be mentioned.

1.1. Evidence for the Difference between Females and Males Outcome in STEM.

Researchers found that age and ability are the two crucial elements when exploring the gender difference in STEM. There was no significant gender difference in STEM skills with young children [12]. While in high school, researchers found there were small but important differences in STEM outcomes between females and males [12]. This finding indicates that gender differences in STEM performance changes over course development and students' ability level. Another study can verify this finding as well. Data from the National Assessment of Educational Progress (NAEP) showed that gender difference in eighth grade was 0.04, while twelfth grade was 0.10, which was slightly larger than it in eighth grade [13]. This result demonstrates gender difference in STEM may be positively correlated with age.

Several researchers illustrated that countries as a factor influenced the magnitude of gender difference as well. Stoet and Geary represented that gender difference was less significant in STEM lower-performing countries than in higher-performing countries [14].

1.2. The Relationship between STEM Performance and Gender Stereotype in Current Studies

Researchers have conducted several studies associated with STEM outcomes and gender stereotype. For instance, traditional math gender stereotypes, self-efficacy, math self-concepts and motivational beliefs were significantly relevant to math grades in middle school [15-17]. In high school, besides the factors in middle school, math GPA were related to self-reported math grades as well [18].

A large amount of studies demonstrated that the difference between females and males' grades attributed to cognitive variables involving problem solving and mathematical thinking abilities[19]. Geary and his colleague indicated that males performed better than females when they completed the same administered arithmetical computations and arithmetical reasoning tests. Although the IQ test showed no difference, the advantages of males in mathematical thinking abilities, including computational fluency and spatial cognition played crucial roles in the math test [20]. Another research conducted by Gallagher and his colleague found that the score of SAT-M and GRE-Q were impacted significantly by the strategy flexibility in mathematical problem solving in high school. Men outperformed women overall, especially the problems demanding shortcuts, or multiple solution paths [21]. In addition, several affective variables affect gaps in females and males STEM outcomes as well, such as mathematics anxiety, mathematical attitudes and self-efficacy [19]. In the study of Casey and Ganley, the conclusion that gender differences in mathematics attitudes triggered the gap in math outcomes [22]. Another study exhibited there was negative moderate correlation between mathematical thinking and mathematics anxiety, and mathematical thinking determined the mathematical performance. It can be seen that mathematics anxiety influenced mathematics scores [23]. Moreover, self-efficacy was considered in current studies as well. A research showed that women perceived themselves as academically weaker than men among STEM majors [24]. And self-efficacy had negative impacts on female STEM expectancy.

2. METHOD

This study aims to synthesize the current quantitative studies associated with the reason why there was gender differences in STEM performance. Thus, the systematic review is interested in exploring: What are the factors related to the gender differences in STEM performance? In order to analyze the influential elements, the author summarized and categorized all factors according to a bio-psycho-social model.

2.1 Inclusion Criteria for Systematic Review

The inclusion criteria used in this systematic review are quantitative study, adolescents as the target population, focusing on gender differences in STEM performance. Furthermore, published journals, empirical studies and published in English were also involved in the criteria.

2.2 Search Strategy

Gender differences in STEM performance has been researched from different perspectives and disciplines, and therefore databases relevant to education, social science and psychology were searched. Included databases were: ERIC, EBSCO, Child Development & Adolescent Studies and PsycINFO. The search terms contained four sections: (i) keywords related to gender difference, (ii) keywords related to STEM (iii) keywords related to adolescents (iv) keywords related to quantitative research. Search terms from each area were applied to each database and combined using AND.

The search process aimed to get any keyword from section 1, with a keyword from section 2 and a keyword from section 3. The word "OR" and wild card function—Asterisk were selected to cover variability spellings and derivatives. In addition, synonyms were used for literature searching. Table1 is used to display the search terms.

Table 1. Search Terms

Gender difference	Performance	Adolescents	STEM	Quantitative
Gender gaps Sex difference Sex gaps	Grades Scores	Teenagers Young adults Teens Youth Young persons	Science Technology Engineering Mathematics	Data analysis Statistic analysis

2.3 Screening Process and Selection of Studies

The database searches identified 329 articles. 220 articles were found in ERIC; 32 articles were found in PsycINFO; 23 articles were found in Child Development & Adolescent Studies; 54 were found in EBSCO. In addition, 9 studies identified through the reference's list of Gallagher and his colleagues. Therefore, 338 articles were included in the screening process that followed. All the articles were imported into Endnote and this led to the deletion of 20 articles as duplicates, and 318 articles remained. Screening refers to the reading of titles and/or abstracts of identified work and assessing them against the inclusion/exclusion criteria. The process of screening led to the exclusion of 274 articles and the inclusion of 44 articles for the first round of the full-text review. Articles were excluded for different reasons. Firstly, performance or grades in STEM subjects ought to be the first reason, or performance or grades in STEM subjects should make up a large majority of essays' contents. Some excluded studies paid more attention to the relationship between STEM performance and future decision making according to gender differences. Secondly, participants were inconsistent with the target group. Several researchers concentrated on children resulting from children in kindergarten or elementary schools have shown different interests and attitudes towards mathematics, which may affect future study and employment. Thirdly, research method was different. Although the word "quantitative" has been used in searching, the keyword was applied in both titles and abstracts. Searching results might conclude qualitative and mixed method research but mention quantitative in the abstract. Moreover, the published journal was the resource type to be included in this study. Finally, all the articles ought to be published in recent 15 years (2006-2021).

A total of 44 papers were included in the first round of full-text screening. 44 was still a relatively large number to read; therefore, these papers were read briefly, focusing on the introduction, method, and conclusion section. The papers were screened according to inclusion and exclusion criteria. There were 29 papers excluded in total. The first reason was irrelevant participants. Some studies chose children in primary school and adults in STEM industry as samples. Several articles were not eligible owing to the quantitative; and most articles were excluded since those studies paid attention to confirm that there are gender differences in some special elements. For instance, sense of belonging in course, career interests, and critical thinking skills. These elements may impact gender differences in STEM fields, but researchers did not mention their relationship with STEM performance in excluded studies. In addition, some studies were excluded because researchers aimed to confirm there were gender differences in STEM performances in a special area, instead of exploring the reasons of why the differences occurred. Furthermore, the studies getting inconsistent results with this essay were rejected as well.

The remaining 15 articles were included for a more detailed full-text screening. Finally, 2 articles were excluded, leaving 13 articles to be included in the review. The 2 articles were excluded due to the focus of gender differences in future decision and career development in STEM field, which means STEM performance was mentioned relatively less than other studies. The searching process was exhibited in Figure 1.

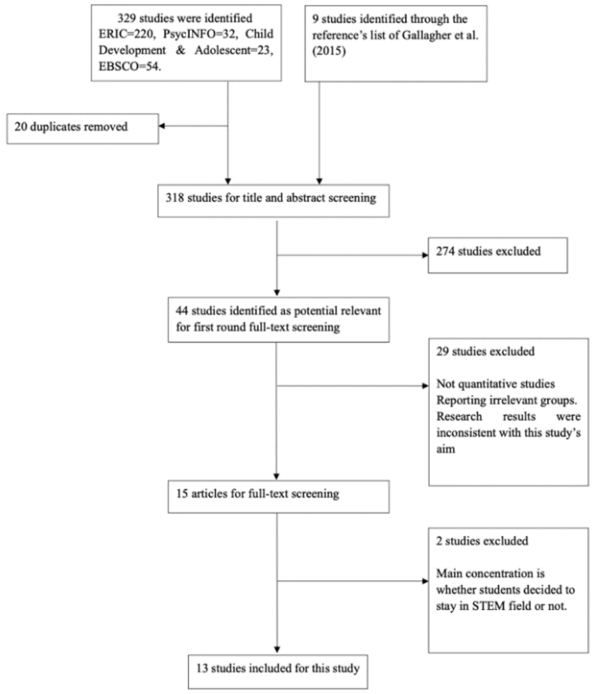

Figure 1 the PRISMA

2.4 Theoretical Framework

This essay selected a bio-psycho-social model to investigate the gender difference in STEM performance, which attempted to bring together the disparate findings from the perspective from cognition, emotion, and sociocultural environment. Initially, the origins and background of bio-psych-social models will be discussed.

Bio-psycho-social theories explore the reciprocal relationship among biological, psychological, and socio-environmental factors, in order to comprehend development deeper. This theory was proposed firstly when Sherman used biological/environmental correlations to illustrate gender differences, which was the foundation of a ''bent twig'' model of individual differences [25]. While Petersen presented bio-psycho-social theories to investigate gender differences in cognitive abilities. She found there were interconnections among socialization, biology, and psychology, which affected the development of gender difference. According to her theory, individuals were influenced by sociocultural elements, through broader society, peers, and family. In addition, psychological changes happened with development as well when biologically based hormonal varied in different time [26].

In this study, the authors referred to the model presented by Casey and Ganley [27], but modified in some details, which attributed gender differences in STEM performance to four major factors: 1) group gender differences in cognitive skills such as spatial and arithmetic skills, 2) group gender differences in STEM attitudes and anxieties, 3) group gender differences in sociocultural factors such as social economic status (SES) and school sex diversity. All three factors affected by biological, psychological, and socio-cultural influences.

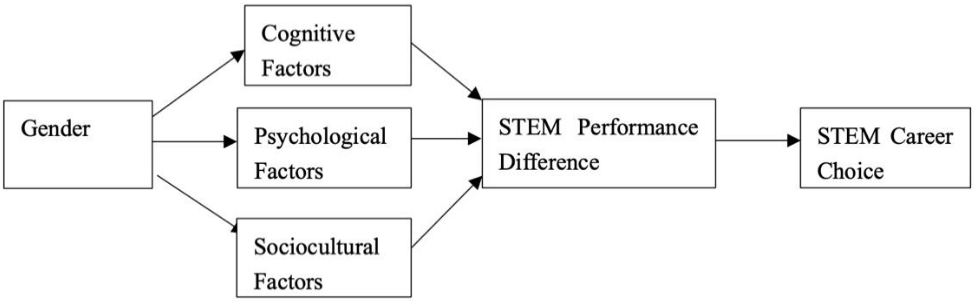

Figure 2 A bio-psycho-social model of gender difference in STEM

3. RESULTS

There are in total of 16 articles eligible for this review. In this section, studies' characteristics and quality of included studies will be discussed.

3.1 Studies' Characteristics

3.1.1 Methods Employed

10 included studies used first-hand data, and 3 studies used secondary data. For the 10 studies with used first-hand data, half of them took special tests to obtain performance data; while other 5 researches collected students' grades from past and future grades to analyse. In addition, 7 researches designed questionnaires to collect information about affective factors such as self-efficacy, anxiety and attitudes.

There were 3 studies used secondary data. The study conducted by Starr and Simpkins draw data from the High School Longitudinal Study (HSLS) dataset. The data was approved by the University of California, Irvine Institutional Review Board under the project name "Family Support of Math and Science: Examining an Untapped Source of Resilience for Diverse High School Students" [28]. Other 2 studies investigated the sociocultural factors related with STEM performance, employing data from the 2011 TIMSS assessment. The assessment is a large multinational study conducted by the International Association for the Evaluation of Educational Achievement (IEA) [29].

3.1.2 Participants Characteristics

The sample size of included studies ranged between 55 and 261738. The largest sample size 261738 was in the study of Reilly, Neumann and Andrews, since the study used secondary data from the 2011 Trends in Mathematics and Science Study (TIMSS) to investigate gender differences in STEM achievements and gender equality. The smallest sample size 55 was in a study from Gallagher and his colleague, which research strategy flexibility in mathematical problem solving from the perspective of gender. Except for the two secondary data studies, samples in 3 studies were students from junior and senior high schools, and other 7 studies selected undergraduate students as the sample. In addition, the age range was difficult to be identified since several studies did not provide detailed information about that. Only 4 studies reported the information related to ages, and age ranges from 12-23 years.

3.2 Quality of Included Studies

This study used the CASP list to assess the several elements reported by included studies, and the CASP list is shown in Table 2. There were nine criteria evaluated in the CASP list: aims and purpose; theoretical frameworks; suitability of qualitative design; a link between aim and study design; participant selection; data collection; ethical issues; data analysis and findings. In this list, "++" refers to clearly stated or discussed; "+" refers to partly addressed and "0" refers to no information provided.

The quality assessment results identified one influential study from Cherney and Campbell. The two researchers mentioned every factor in this study. Participant selection and ethical issues were the two factors that ignored the most frequently. Three articles provided no information about participant selection, since they used secondary data. Additionally, 10 studies neglected the information of ethical issues. Only three studies mentioned that and they provided full ethical information about the study to all respondents in order to obtain consents to engage in the research process.

Table 2. the CASP list

Study	Aims and purpose	Theoretical framework	Suitability of quantitive design	Link between aim and study design	Participant selection	Data collection	Ethical issues	Data analysis	Findings
Gallagher et al. (2015)	++	0	++	+	+	++	++	++	++
Lemos et al. (2013)	++	++	++	++	++	++	0	++	++
Cherney and Campbell (2013)	++	+	++	++	++	++	++	++	++
Rice et al. (2013)	++	+	+	++	+	+	0	++	++
Cotner et al. (2020)	++	++	++	++	+	++	0	++	++
Seyranian et al. (2018)	++	+	++	++	++	+	+	++	++
Jones, Ruff and Paretti (2013)	++	+	++	++	+	++	0	++	++
Marshman et al. (2018)	++	+	++	++	+	++	0	++	++
Marshman et al. (2020)	++	++	++	++	+	+	0	++	++
Smeding (2012)	++	+	++	++	+	+	0	+	++
Starr and Simpkins (2021)	++	+	+	++	0	+	0	++	++
Hamamura (2012)	++	++	+	+	0	0	0	+	+
Reilly, Neumann and Andrews (2019)	++	++	+	+	0	+	0	++	++

4. DISCUSSION

This study aims to synthesize the current qualitative studies associated with the reason why there was gender differences in STEM performance. In order to explore this question, a bio-psycho-social model was implemented in this section. According to this model, gender differences in STEM performance can attribute to three dimensions: 1) group gender differences in cognitive skills such as algorithmic and arithmetic skills, 2) group gender differences in psychological factors such as STEM anxieties and self-efficacy, 3) group gender differences in sociocultural factors such as social power distance and culture background.

4.1 Cognitive Skills

Two included studies demonstrated that cognitive skills triggered of gender differences in STEM performance. The research conducted by Gallagher and his colleague found that male students were more likely than female students to successfully match strategies to problem characteristics [30]. The strategies they used were categorized into algorithmic or insightful approach. Besides, intuitive strategy played a significant role as in this test as well. In Gallagher's study, no matter which type's test was employed, males' students always performed better than female students.

Lemos and his colleague investigated gender differences in STEM performance from the perspective of cognitive skill as well [31]. They completed five reasoning tests, including abstract [AR], numerical [NR], verbal [VR], mechanical [MR], and spatial [SR]. The results exhibited that males outperform females in all the subtests, and mechanical reasoning (MR) was the most significant one. The persistent advantage in mechanical reasoning (MR) for males can explain their higher presence in STEM field.

4.2 Psychological Factors

There were 9 included studies paying attention to psychological factors related to gender differences in STEM performance. During the process of searching literature, it can be found that majority of existing researches focused on the perspective of students' psychology. Psychological reasonings from the included studies were classified into 5 types: perfectionism, test anxiety, course identity and belonging, self-efficacy and stereotype, which will be discussed in the following.

4.2.1 Perfectionism

Research conducted by Rice and his colleague held that perfectionism impacted on students' score. Perfectionism was a facet of both conscientiousness and neuroticism [32]. Therefore, perfectionism can be defined as a personality trait consisted of extremely high, self-imposed performance expectations or standards; or unrealistically high personal standards and excessively critical self-evaluation [32]. Rice and his colleague combined perfectionism and gender difference, showed that maladaptively perfectionistic females were more likely to perform in STEM areas disappointedly, while adaptively perfectionistic females performed better in these courses. However, perfectionism was not substantially associated with males' STEM scores. This difference may explain the difficulties that women pursuing STEM careers.

4.2.2 Test Anxiety

Course anxiety was associated with students' performance as well. Cotner and his colleague's research illustrated that girls expressed more test anxiety in STEM test than boys, and the anxiety girls experienced negatively predicted their performance in class [33].

4.2.3 Course Identity and Belonging

Comparing to males, females expressed less course belonging and less course identification than males according to Seyranian and his colleagues [34]. In addition, students with higher course identification were more possible to perform better, and students who performed better reported increasing course identification at the end of the term. Under this circumstance, male students were more likely to obtain higher grades in STEM performance. Moreover, identification in STEM course was a crucial predictor of persistence in STEM field [35].

4.2.4 Self-efficacy

Self-efficacy was an important motivation in students' engagement, participation, and retention in STEM field [36]. 2 included studies concentrated on the relationship between self-efficacy and gender difference in STEM performance. The conclusion demonstrated that females had lower self-efficacy than male, and the low self-efficacy would last no matter what instructors and course formats. This result also indicates that females' low self-efficacy causes detrimental short-term and long-term impacts, including poorer performance in STEM performance.

4.2.5 Gender Stereotypes

There were 3 included studies related to stereotype threats in STEM areas, which indicated that negative gender stereotype in STEM subjects was likely to undermine self-perceptions of ability, performance and interest for females. The negative gender stereotype refers to women obtained relatively lower ability in mathematics and reasoning. In order to investigate the connection between gender stereotype and gender difference in STEM performance, Smeding found that female engineering students perceived weaker implicit gender- math and gender-reasoning stereotypes than female humanities, male engineering students [37]. While implicit stereotypes of humanity females were more negatively associated with STEM grades compared with engineering females. Another study conducted by Starr and Simpkins believed that students' gender stereotypes were related to parents' gender stereotypes, however, were unrelated to teacher stereotypes [38]. In addition, according to the study of Cherney and Campbell, females without gender stereotype threats performed better than those under the threats, while males were not [39].

4.3 Sociocultural Effects

In addition to the internal factors with students in gender difference, there were several external factors such as sociocultural effects were considered as well. Both the 2 included studies used the data of 2011 Trends in Mathematics and Science Study (TIMSS). Hamamura mentioned how low power distance societies impacted gender differences in STEM performance. He suggested that the societies' power distance predicted gender differences in STEM performance: The results that males performed better than females was more pronounced in low relative to high power distance societies [40]. In another study conducted by Reilly, Neumann and Andrews, they indicated that cross-cultural variability played a significant role in gender differences of STEM performance, no matter the influence was development or suppression [41].

4.4 The Relationship between Three Dimensions and Recommendations for Future Practice

Accounting for the bio-psycho-social model, there were three factors related to gender difference in STEM performance, including cognitive skills, psychological factors and sociocultural effects. How each element impact student respectively has already been discussed, but these elements may interact with each other, and influenced on STEM performance further.

Existing literature mentioned that there was gender difference in terms of cognitive skills, such as algorithmic and arithmetic skills, and the differences caused that males may outperform than females in STEM subject. However, the difference varied from different age. A study conducted by Logan and Lowrie found that the cognitive differences occurred at very young age, but varied with future development [42]. The difference in cognitive skills cannot predicted STEM performance, namely, and the cognitive skills were not positively correlated with STEM scores. During the period of junior school and senior high school, the environment paid more attention to “scores”. And students, parents and teachers in junior school and senior high school concentrated more on competition and comparison, since teachers selected pre-established benchmark levels and high-stakes standardized testing to rate students. In addition, current research indicated that girl students influenced more by anxiety and course identity and belonging. Depending on that, for girls, psychological factors may impede the influences of cognitive skills and turn to be a more prominent element in STEM performance. This result was consistent with the study of Wang and Degol [43]. In other words, psychological factors such as anxiety and peer pressure reacted on cognitive skills, and the interactive effects resulted in gender differences in STEM performance. Nevertheless, the interactive effects did not mean that females always performed worse than males. Some studies also indicated that there were no gender differences in STEM performance since girls have more positive perceptions towards STEM than boys, which was consistent with the research from Reilly and his colleagues. Additionally, mindset was malleable and can be shaped by social forces, and females benefited more than males from growth mindset training, which decreased the discernible gender gaps in STEM performance. Sociocultural effects reacted on cognitive and psychological factors as well. Students with high SES were likely to develop their cognitive skills and train mindset further. Moreover, teachers in schools and parents have higher possibility to express less gender stereotype and bias. The better environment and strengthened psychology may eliminate gender difference in STEM areas.

In order to address the main causes of the pervasive gender imbalance in STEM fields, the authors provide several suggestions for addressing the problem. Initially, teachers ought to be emphasized on both ability enhancement and interest enhancement. It has already been found that capability and interest are equally crucial to performance and career paths. Therefore, it is still significant to promote performance in STEM, but cultivating female’s interests in STEM subjects can produce more female scientists in the long run. Particularly, it is essential to cultivate interests for the women who have the talents to succeed in STEM but do not seem to have the interests. Secondly, teachers should break down their stereotype about females and STEM. Stereotype will affect people’s think, behave, and feel about their own abilities. Therefore, we are supposed to highlight the achievement of females in STEM field in order to break down negative stereotypes. Teachers and parents are the crucial role to practice this. In addition, teachers should emphasize that efforts and persistence play more significant roles instead of talents. Teachers ought to strengthen growth mindset for girls, and deepen the understanding that efforts and hard work result in good performance in STEM subjects instead of talents.

5. CONCLUSION

This study aims to synthesize the current qualitative studies associated with the reason why there was gender differences in STEM performance. In order to explore this question, the authors searched 4 databases: ERIC, EBSCO, Child Development & Adolescent Studies and PsycINFO, and obtained 13 studies to analyze. This essay selects a bio-psycho-social model to investigate the gender difference in STEM performance, which attempted to bring together the disparate findings from the perspective from cognition, psychology, and sociocultural environment. The results demonstrate that there were three factors related to gender difference in STEM performance, including cognitive skills, psychological factors and sociocultural effects. The three factors not only affect gender difference in STEM performance respectively, but also interact with each other and conduce in gender difference in STEM grades. The authors also provide several suggestions for future practice. All participants including parents, teachers and students themselves ought to pay attention to the phenomenon. In addition, there are some limitations in this study. Firstly, more databases should be searched since there are still some ignored elements related to gender and STEM. In addition, how the three factors interact is not clear enough, more empirical studies are needed to investigate deeply. Finally, quantitative studies are still dominant in this field. Therefore, researchers can select qualitative method to explore unanswered question about gender difference in STEM performance in the future.

REFERENCES

[1] Beede, D. N., Julian, T. A., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011, August 1). Women in STEM: A gender gap to innovation. Economics and Statistics Administration Issue Brief No. 04-11. Retrieved from https://ssrn.com/ abstract1⁄41964782

[2] Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. Science, 321(5888), 494-495.

[3] Kazu, H., & Ersözlü, Z. N. (2008). Öğretmen adaylarının problem çözme becerilerinin cinsiyet, bölüm ve öss puan türüne göre incelenmesi. Abant İzzet Bay

[4] Dede, Y., & Dursun, Ş. (2008). İlköğretim II. kademe öğrencilerinin matematik kaygı düzeylerinin incelenmesi. Eğitim Fakültesi Dergisi, 11(2), 295-312.

[5] Yücel, Z., & Koç, M. (2011). The relationship between the prediction level of elementary school students' math achievement by their math attitudes and gender. Elementary Education Online, 10(1), 133-143.

[6] Van de Gaer, E., Pustjens, H., Van Damme, J., & De Munter, A. (2008). Mathematics participation and mathematics achievement across secondary school: The role of gender. Sex Roles, 59(7-8), 568-585.

[7] Tate, W. F. (1997). Race-ethnicity, SES, gender, and language proficiency trends in mathematics achievement: An update. Journal for Research in Mathematics Education, 28(6), 652-679.

[8] Köğce, D., Yıldız, C., Aydın, M., & Altındağ, R. (2009). Examining elementary school students' attitudes towards mathematics in terms of some variables. Procedia-Social and Behavioral Sciences, 1(1), 291- 295.

[9] Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics—A "Hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. European Journal of Psychology of Education, 22(4), 497-514.

[10] Kargar, M., Tarmizi, R. A., & Bayat, S. (2010). Relationship between mathematical thinking, mathematics anxiety and mathematics attitudes among university students. Procedia-Social and Behavioral Sciences, 8(2010), 537-542.

[11] Çakıroğlu, E., & Işıksal, M. (2009). Pre-service elementary teachers' attitudes and self-efficacy beliefs toward mathematics. Education and Science, 34(151), 132- 139.

[12] Hutchison, J. E., Lyons, I. M., & Ansari, D. (2019). More similar than different: Gender differences in children's basic numerical skills are the exception not the rule. Child Development, 90(1), e66–e79.

[13] Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of Educational Progress assessments. Journal of Educational Psychology, 107(3), 645–662.

[14] Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within-and across-nation assessment of 10 years of PISA data. PloS one, 8(3), Article e57988.

[15] Casad, B. J., Hale, P., & Wachs, F. L. (2017). Stereotype threat among girls: Differences by gender iden- tity and math education context. Psychology of Women Quarterly, 41(4), 513–529. https://doi. org/10.1177/0361684317711412.

[16] Passolunghi, M. C., Rueda Ferreira, T. I., & Tomasetto, C. (2014). Math–gender stereotypes and math- related beliefs in childhood and early adolescence. Learning and Individual Differences, 34, 70–76. https://doi.org/10.1016/j.lindif.2014.05.005.

[17] Plante, I., de la Sablonnière, R., Aronson, J. M., & Théorêt, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. Contemporary Educational Psychology, 38(3), 225–235. https://doi.org/10.1016/j.cedpsych.2013.03.004.

[18] Chatard, A., Guimond, S., & Selimbegovic, L. (2007). "How good are you in math?" The effect of gender stereotypes on students' recollection of their school marks. Journal of Experimental Social Psychol- ogy, 43(6), 1017–1024. https://doi.org/10.1016/j.jesp.2006.10.024.

[19] Nurlu, &. (2017). Developing a teachers' gender stereotype scale toward mathematics. International

Electronic Journal of Elementary Education, 10(2), 287-299.

[20] Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. Journal of Experimental Child Psychology, 77(4), 337-353.

[21] Gallagher, A., De Lisi, R., Holst, P., McGillicuddy-De Lisi, A., Morely, M., & Cahalan, C. (2000). Gender Differences in Advanced Mathematical Problem Solving. Journal of Experimental Child Psychology, 75(3), 165-190.

[22] Casey, B., & Ganley, C. (2021). An examination of gender differences in spatial skills and math attitudes in relation to mathematics success: A bio-psycho-social model. Developmental Review, 60, 100963.

[23] Kargar, M., Tarmizi, R., & Bayat, S. (2010). Relationship between Mathematical Thinking, Mathematics Anxiety and Mathematics Attitudes among University Students. Procedia, Social and Behavioral Sciences, 8, 537-542.

[24] MacPhee, D., Farro, S., & Canetto, S. (2013). Academic Self-Efficacy and Performance of Underrepresented STEM Majors: Gender, Ethnic, and Social Class Patterns. Analyses of Social Issues and Public Policy, 13(1), 347-369.

[25] Sherman, J. (1978). Sex-related cognitive differences. Springfield, IL: Charles C. Thomas.

[26] Petersen, A. C. (1980). Biopsychosocial processes in the development of sex-related differences. The Psychobiology of Sex Differences and Sex Roles, 31–55.

[27] Casey, B., & Ganley, C. (2021). An examination of gender differences in spatial skills and math attitudes in relation to mathematics success: A bio-psycho-social model. Developmental Review, 60, 100963.

[28] Starr, C., & Simpkins, S. (2021). High school students' math and science gender stereotypes: Relations with their STEM outcomes and socializers' stereotypes. Social Psychology of Education, 24(1), 273-298.

[29] Reilly, D., Neumann, D., & Andrews, G. (2019). Investigating Gender Differences in Mathematics and Science: Results from the 2011 Trends in Mathematics and Science Survey. Research in Science Education, 49(1), 25-50.

[30] Gallagher, M., Burwell, R., & Burchinal, M. (2015). Severity of spatial learning impairment in aging: Development of a learning index for performance in the Morris water maze. Behavioral Neuroscience, 129(4), 540–548.

[31] Lemos, G. C., Abad, F. J., Almeida, L. S., & Colom, R. (2013). Sex differences on g and non-g intellectual performance reveal potential sources of STEM discrepancies. Intelligence, 41(1), 11–18.

[32] Cook, K. S., Cheshire, C., Rice, E. R. W., & Nakagawa, S. (2013). Social Exchange Theory. Handbooks of Sociology and Social Research, 61–88.

[33] Aguillon, S. M., Siegmund, G.-F., Petipas, R. H., Drake, A. G., Cotner, S., & Ballen, C. J. (2020). Gender Differences in Student Participation in an Active-Learning Classroom. CBE—Life Sciences Education, 19(2), ar12.

[34] Seyranian, V., Madva, A., Duong, N., Abramzon, N., Tibbetts, Y., & Harackiewicz, J. M. (2018). The longitudinal effects of STEM identity and gender on flourishing and achievement in college physics. International Journal of STEM Education, 5(1).

[35] Jones, B. D., Ruff, C., & Paretti, M. C. (2013). The impact of engineering identification and stereotypes on undergraduate women's achievement and persistence in engineering. Social Psychology of Education, 16(3), 471–493.

[36] Marshman, E., Kalender, Z., Nokes-Malach, T., Schunn, C. and Singh, C. (2018) Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for alarm?. Physical Review Physics Education Research, 14(2).

[37] Smeding, A. (2012). Women in Science, Technology, Engineering, and Mathematics (STEM): An Investigation of Their Implicit Gender Stereotypes and Stereotypes' Connectedness to Math Performance. Sex Roles, 67(11-12), 617–629.

[38] Starr, C. R., & Simpkins, S. D. (2021). High school students' math and science gender stereotypes: relations with their STEM outcomes and socializers' stereotypes. Social Psychology of Education, 24(1), 273–298.

[39] Cherney, I. D., & Campbell, K. L. (2011). A League of Their Own: Do Single-Sex Schools Increase Girls' Participation in the Physical Sciences? Sex Roles, 65(9-10), 712–724.

[40] Hamamura, T. (2012). Social Class Predicts Generalized Trust But Only in Wealthy Societies.

Journal of Cross-Cultural Psychology, 43(3), 498–509.

[41] Reilly, D., Neumann, D. L., & Andrews, G. (2019). Investigating Gender Differences in Mathematics and Science: Results from the 2011 Trends in Mathematics and Science Survey. Research in Science Education.

[42] Logan, T., & Lowrie, T. (2017). Gender perspectives on spatial tasks in a national assessment: a secondary data analysis. Research in Mathematics Education, 19(2), 199–216.

[43] Wang, M.-T., & Degol, J. L. (2016). Gender Gap in Science, Technology, Engineering, and Mathematics (STEM): Current Knowledge, Implications for Practice, Policy, and Future Directions. Educational Psychology Review, 29(1), 119–140.