Bridging the Mathematics Gender Divide: A Mixed Methods Study of Teacher Biases and Their Impact on Female Students

Abstract

This paper investigates how British secondary mathematics teachers' gendered ability-based expectations influence female students' motivational beliefs, mathematical academic self-efficacy, and achievement outcomes. The research question emerged from persistent gender disparities in mathematics education. Grounded in Social Cognitive Theory, Expectancy-Value Theory, Stereotype Threat Theory, and Implicit Theories of Intelligence, the proposed mixed-methods study employs a critical realist paradigm. Quantitative data from surveys and academic records will be analysed using descriptive and inferential statistics. Qualitative data from interviews and observations will undergo thematic analysis. Findings will inform recommendations for reducing gender bias in mathematics education and promoting equitable outcomes for all students.

Keywords: mathematics education, gender bias, stereotype threat, expectations

Acknowledgement

To my Norm – without which I would not have finished this. I love you.

Chapter 1 – Introduction

Mathematics education has a problem: the gender differences in academic outcomes are significant and widening (Department for Education, 2024). Persistent stereotypes hinder gender equality in mathematics outcomes and representation (Fennema and Sherman, 1976; Becker and Hall, 2024), and this paper addresses the question of how teachers' gender ability-based expectations influence female students' motivational beliefs, self-efficacy and achievement outcomes in mathematics. The rationale behind this study will be introduced in this chapter alongside the theoretical framework, literature review framing, research question development and context.

Despite significant progress in gender equality around the world in the last fifty years, mathematics stereotypes persist, affecting self-concepts and perceived self-efficacy from early education to career aspirations (Barth et al., 2018; Wang et al., 2020). As education is formed from the relationship between students and teachers, teachers' implicit biases must be addressed to make progress, with teacher-held stereotypes having a direct effect on students (Steele and Aronson, 1995) and teacher expectations affecting students' self-concepts (Rosenthal and Jacobson, 1968). It is, therefore, vital that this area is studied extensively. Furthermore, addressing these biases can potentially improve female students' mathematics outcomes and increase their representation in STEM fields, contributing to a more diverse and equitable workforce in these crucial areas.

This study links to the module themes from E822 in several ways. Firstly, the introduction of research paradigms within this course has led to the confrontation of what paradigm fits into this area of study best (Bhaskar, 1998; Grix, 2002; Corbetta, 2011; Fletcher, 2017), and the foci on quantitative and qualitative data collection has influenced a mixed methods approach (Creswell and Creswell, 2018). Additionally, this study contributes to the ongoing debate about research paradigms in educational research, reflecting the module's emphasis on critical engagement with methodological approaches. The EP pathway has enabled the research to build throughout the year, further deepening a conceptual understanding of gender disparities in mathematics education. This module has developed both my

researcher identity and also how my own background and beliefs can influence the research process and interpretation of results (Cohen, Manion and Morrison, 2018).

The first half of this paper outlines an extended literature review of quantitative and qualitative studies on gender differences in mathematics. To frame this literature review, a comprehensive search was conducted using key educational databases, including ERIC, JSTOR and British Educational Index. Search terms included gender, mathematics achievement, girls, STEM, stereotypes, and expectations. Peer-review journals were prioritised, as were journals from the last twenty years. The main focus areas are teacher expectations, student self-efficacy, mathematics anxiety, stereotype threat, and achievement outcomes. This systematic review approach ensures a comprehensive understanding of the current state of research in this field, providing a solid foundation for the proposed study. The literature review highlights gaps in the current knowledge of this area and guides the development of research questions and methodology. The position of gender as an influence on mathematics outcomes is repeatedly framed within an intersectionality perspective, highlighting the importance of a nuanced approach to this area of study (Casanova, Vukovic and Kieffer, 2021; Battey et al., 2022).

A conceptual framework is developed from the pillars of Social Cognitive Theory (SCT) (Bandura, 1986), Expectancy-Value Theory (EVT) (Wigfield and Eccles, 2000), Stereotype Threat Theory (STT) (Steele and Aronson, 1995) and Implicit Theories of Intelligence (ITI) (Dweck, 1990). This framework provides a robust theoretical foundation for exploring the complex interplay between teacher expectations, student beliefs, and academic outcomes in mathematics education. The proposed research study will adopt a mixed methods design grounded in a critical realist paradigm (Bhaskar, 1998; Fletcher, 2017). A critical realist stance acknowledges both objective reality and socially constructed understanding (Maxwell and Mittapalli, 2015), and the stratified ontology of critical realism drives research searching for underlying mechanisms and understanding (Bhaskar, 1998).

This year of study has guided me through the process of honing a set of research questions, with feedback from TMA01 and TMA02 making me consider the scope and actionability of the research question more deeply. Alongside personal feedback, the literature review has identified important gaps to investigate, further

influencing the research question. This iterative refinement process demonstrates the development of critical thinking and research skills throughout the module.

The research question addresses key aspects of the problem identified in the literature, focusing on the crucial role of teacher expectations in shaping student outcomes.

“How do British secondary mathematics teachers’ gendered ability-based expectations influence female students’ motivational beliefs, mathematical academic self-efficacy, and achievement outcomes?”

With sub-questions:

1. “To what extent do mathematics teachers’ assumptions about innate skills exhibit gender biases reflecting societal stereotypes about mathematical abilities?”
2. “How do teachers’ gendered expectations manifest in classroom practices and environments, and how do these influence female students’ mathematics anxiety and engagement?”

The research proposed will be undertaken in British secondary schools in the West Midlands, specifically within their mathematics classrooms. Secondary education is a critical period for developing mathematical self-concept and career aspirations and therefore should be investigated. These results strive to give a window into the underlying mechanisms at play in shaping the gender divide in mathematics education in the West Midlands and, therefore, the national landscape. As a mathematics teacher, this area is very important to me, and this personal connection to the research topic enhances the study’s relevance and potential impact on real-world educational practices.

In conclusion, the importance of addressing the gender disparities in mathematics education cannot be understated. In order for society to become more numerate, this must be addressed. Focusing on teacher expectations and their impact on female students gives us a lens through which we can identify areas for improvement. The literature on this topic will be extensively analysed in the following chapters, a conceptual framework will be presented, and the research proposal will be outlined. This study can potentially influence educational policy and practice in promoting

gender equity in mathematics education, and by connecting the importance of this issue to improving overall numeracy in society, this research underscores its broader societal impact and the urgent need for evidence-based interventions in mathematics education.

Chapter 2 – Literature Review – The Topic

Introduction

Society has progressed considerably in terms of gender equality over the past half-century, but stereotypes about girls' abilities in mathematics persist. Despite being known to hinder gender equality in mathematics outcomes and representation from the 1970s (Fennema and Sherman, 1976), stereotypes are still being seen worldwide today (Becker and Hall, 2024). This may be detrimental due to the impact stereotypes have on self-concepts and perceived self-efficacy from early education to career aspirations (Barth et al., 2018; Wang et al., 2020; Becker and Hall, 2024). This chapter will explore the historical and contemporary contexts of gender stereotypes in mathematics education, the impact of teacher beliefs and expectations, the motivational beliefs and self-efficacy of female students, and the resulting achievement outcomes.

Gender Stereotypes in Mathematics Education

Gender stereotypes and bias in mathematics education have been known for a long time; historical and contemporary sources both agree that this is prevalent and addressing the problem is vital for our society. Stereotypes have persisted despite societal shifts, and evidence shows that this is both not improving and not from a biological source. Gender stereotypes affect students' performance, their self-concept and career aspirations and must be addressed.

Almost half a century ago, Fennema and Sherman's (1977) published their seminal paper investigating the extent to which mathematics was perceived as male-dominated. Their findings identified key differences in areas of strength between male and female participants and found a tendency to assume boys are better at mathematics by parents and teachers. The nine Fennema-Sherman scales (Fennema and Sherman, 1976) are Likert-style (Likert, 1932) scales that proved highly instrumental in probing mathematical beliefs around gender and societal influences. Further research indicated that these harmful stereotypes were also prevalent in students and adults (Hyde, Fennema and Lamon, 1990; Lindberg et al., 2010) despite evidence showing little significant gender difference in mathematics

performance. While societal shifts have occurred, it is important to show that these stereotypes are not new. Nurlu's (2017) development of a Teacher's Gender Stereotype Scale towards Mathematics indicates a more contemporary attempt to codify the extent to which teachers hold gender stereotypes. Significant differences were indicated despite a focus on Turkey, which must be further investigated in various cultural contexts. More recent research highlights that persistent gender stereotypes in mathematics exist and significantly impact both attainment and education as a whole (Becker and Hall, 2024).

It is evident that not only do teachers hold biased views based on a student's gender, but that these expectations can significantly influence students' performance and self-concept. Gentrup and Rjosk (2018) built on the work of Rosenthal and Jacobson (1968) by identifying that in Germany, teachers held higher expectations for boys in mathematics, which led to self-fulfilling prophecies for both boys and girls, resulting in lower academic outcomes for girls; however, it is unclear if these expectations result from implicit biases or explicit cultural norms. The perception of success also differs, assuming that boys do better due to their innate mathematics ability, whereas girls do better due to high levels of effort (Tiedemann, 2000). This finding is reflected in Heyder, Steinmayr and Kessels' (2019) study on the self-concept of mathematical talent among students. Both of these studies required self-reported statistics similar to Fennema and Sherman's (1977) and must be viewed critically. Flore and Wicherts' (2015) meta-analysis warns against the validity of self-reported data in research investigating stereotype threat, especially concerning recall bias and social desirability bias. So, while these are significant findings, they must be strengthened with more objective data.

It has been found that there is also a gender bias in other aspects of classroom teaching. Firstly, biased mathematical assessment grading significantly affects outcomes and career interest, creating a call to action to challenge implicit biases and explicit gender stereotypes in mathematics education (Barth et al., 2018; Doornkamp et al., 2022). Keller's (2001) study even indicates that this inherent bias may cause girls to self-handicap themselves to try to protect their self-esteem in the face of blatant stereotype threat; this shows that a teacher's gender bias can become internalised in their students. Secondly, Gjøvik, Kaspersen and Farsani (2023) show that stereotypical images of male and female mathematics teachers

impact gender stereotypes around who can be successful in mathematics, creating a cycle of less representation of role models for girls. However, Jaremus (2021) suggests that the issue with mathematics could be our monoglossic view of gender, and the adoption of a heteroglossic view of all people having masculine and feminine traits would lead to higher female involvement in mathematics as a whole.

The topic of gender stereotypes in mathematics education has substantial gaps in the literature in terms of form, content, and geography. The literature lacks long-term knowledge of how gender stereotypes impact education over time through longitudinal studies (Heyder, Steinmayr and Kessels, 2019). The interventions that effectively reduce gender bias and stereotypes remain unclear, suggesting a further investigation into strategies is needed (Bergold et al., 2017; Lee, Lee and Bong, 2022). Furthermore, reviewing the efficacy of growth mindsets in cultural contexts would lead to a more robust consensus of thought about a way forward (Heyder, Weidinger and Steinmayr, 2021; Lee, Lee and Bong, 2022). Becker and Hall's (2024) narrative review of the current research in this area suggests that taking gender in isolation does not address the broader picture of the influences and that research on intersectionality in nuanced studies would shed light on this area. They go on to suggest that peer-to-peer influences also need to be addressed in order to understand gender biases fully (ibid.). Also, while gender bias in mathematics is a truly international research portfolio, research in a UK setting is minimal, and further research in this area would be welcome to find effective local interventions.

Mathematics education has a significant international problem with gender stereotypes and bias; this has led to differences in expectations, teaching decisions, outcomes and career aspirations. More research is needed to gain objective and longitudinal data from various cultural contexts to address this problem and develop effective interventions to promote gender equity in mathematics.

Teachers' Beliefs and Expectations

Many influences form teachers' beliefs and expectations of students' abilities. Van Eycken, Demanet and Van Houtte (2023) indicate three main influences:

- Personal experience: what the teacher has experienced in their teaching careers
- Educational background: the teacher's own formal education and training programs
- Social contexts: societal and cultural norms and norms for the school context in which they work.

Further studies connect teachers' self-efficacy, which is their self-belief in their ability to affect student outcomes, to their beliefs about their students. Bandura (1997) indicated that this is generally due to a firm belief that their ability impacts the learning culture in their settings, thus leading to a positive learning environment for their students. While the framework Bandura (1997) posited is robust, it may require further empirical evidence in a modern UK setting. Wolters and Daugherty (2007) expand on this work by suggesting that teachers' self-efficacy is linked to mastery goal structures that focus on learning and understanding rather than grades and performance. Li (1999) says that the primary influence of a strongly gendered perception of students' performance is created mainly from social and cultural norms, which has been replicated more recently and shown to affect student outcomes significantly (Nürnberger et al., 2016). Despite the age of these three papers, they present strong evidence for the formation of teacher beliefs, but methodologically, they rely on self-reporting (Wolters and Daugherty, 2007), which can be prone to biases (Flore and Wicherts, 2015) or rely on secondary data, which brings the validity into question (Li, 1999).

Teachers' expectations are a significant influence on student performance. Firstly, Rosenthal and Jacobson (1968) demonstrated the Pygmalion effect, which showed that higher teacher expectations can lead to an improved student outcome. Despite methodological criticism of this initial study, Jussim and Harber (2005) confirmed that there is a real effect from teacher expectations, but potentially not as strong as previously stated. Other studies indicate that differential treatment based on teacher expectations can affect student performance, but the effect size was varied for different contexts (de Boer, Bosker and van der Werf, 2010; Rubie-Davies, 2014); this is especially prevalent in differing expectations due to gender and ethnic groups substantially affecting student achievement (Rubie-Davies, Hattie and Hamilton, 2006).

While teacher expectations have a high impact on students' self-concept, their perception of their own competence and ability both generally and within a mathematical domain (Wigfield and Eccles, 2000; Timmermans and Rubie-Davies, 2023), it is unclear if it has an impact on students' perception of the utility of mathematics (Timmermans and Rubie-Davies, 2023). Higher self-concept does lead to higher motivation levels (Wigfield and Eccles, 2000), but there has been a call for studies that control for initial achievement and self-concept levels to understand this relationship better (Timmermans, de Boer and van der Werf, 2016). Sorhagen (2013) indicates that teacher outcomes can have long-term, lasting effects, especially with low-income students. However, there is a lack of consensus, with some studies finding a range of effect sizes and a recommendation for more longitudinal studies to understand better how and why expectation effects vary (de Boer, Bosker and van der Werf, 2010; Johnston, Wildy and Shand, 2019). Clear links between teacher expectations based on gender and academic achievement have been shown (Gentrup and Rjosk, 2018), but the differences in teacher expectations can also be shown to have an effect based on minority backgrounds as well (McKown and Weinstein, 2008; van den Bergh et al., 2010; Peterson et al., 2016), with teacher ethnicity potentially causing differing expectations (Gershenson, Holt and Papageorge, 2016).

Eccles et al. (1983) proposed an expectancy-value theory incorporating, amongst other factors, differing beliefs based on gender. Expanding the concept that teacher expectations are important for a student's achievement suggests that a student's perception of the teacher's beliefs is even more crucial. Lazarides and Watt (2015) found that teachers often hold different beliefs about students' mathematical abilities based on gender, which leads to different outcomes. As this is a longitudinal study, it helps to bolster the claim that these differing expectations are harmful in the long term, but as this was based on the teacher's expectations as perceived by their students, it adds nuance to Eccles et al.'s (1983) claims.

Within mathematics education, it is clear that teachers hold beliefs based on the gender of their students, this leads to different levels of attainment (Nürnberger et al., 2016; Heyder, Steinmayr and Kessels, 2019; Doornkamp et al., 2022) and this is a trend that can be shown around the world; boys generally outperform in mathematics, girls generally outperform in reading (Marks, 2008). However, the

cross-sectional design from Marks (2008) does indicate a lack of causal evidence, but when coupled with longitudinal studies (Lindberg et al., 2010; Upadyaya and Eccles, 2014; Lazarides and Watt, 2015), it helps to triangulate findings. Upadyaya and Eccles (2014) suggest that teacher beliefs shape students' academic mathematical careers and affect their intrinsic motivation and interest in mathematics.

Real or perceived teacher expectations are complex and require a multi-faceted perspective to understand. Firstly, the concept of high expectations underscores the relationship, with teachers having lower expectations of students who do not exhibit culturally expected behaviours around attentiveness and participation (Good and Lavigne, 2017) and positive interactions leading to perceived higher expectations (Rubie-Davies, 2014; Good and Lavigne, 2017). Forgasz and Leder (2017) indicate that societal norms and expectations continue to be a limiting factor for gender equity in mathematics despite interventions aimed at reducing the gender gap, and it has been shown that more gender-neutral families lead to higher outcomes in mathematics (Rodríguez-Planas and Nollenberger, 2018). Glock and Krolak-Schwerdt (2014) warn against the concept of implicit biases based on protected characteristics; while their study mainly focuses on ethnicity and socioeconomic status, their findings indicate that more education is needed to help minimise the effects of implicit biases.

Teachers generally hold beliefs about their students, which leads to expectations, and these expectations lead to different outcomes. While this is mainly implicit and cultural, teachers must address these beliefs and understand their effects on children's lives.

Female Students' Motivational Beliefs and Self-Efficacy

The complex web of motivational effects and beliefs surrounding self-efficacy must be approached from a gender perspective, as fundamental differences have been observed within the literature. Bandura's (1997) defined self-efficacy as a student's belief in their abilities, which is formed from many sources. Lazarides and Watt (2015) analysed the perceptions of mathematics students and found that female students perceive mathematics teachers as having higher prestige beliefs about their

subject when compared to other subjects. However, paradoxically, they have lower expectations of female students than male students; this would affect the ‘verbal persuasion’ and ‘vicarious learning’ sources of self-efficacy (Bandura, 1997). The cumulative effect of socialisation cannot be understated. Girls have been shown to receive more negative input concerning their mathematical ability and experience more gender stereotyping; this has led to a lower self-concept and self-efficacy for girls compared to their male peers when they got older (Becker and Hall, 2024). Bandura (1997) model explains that self-efficacy affects motivation and actions; higher mathematics self-efficacy has been shown to lead to more engagement when analysed using multidimensional profiles (Miller, Perera and Maghsoudlou, 2021). The influence of teachers is substantial, with the idea that teachers’ outward displays of enthusiasm and interest in mathematics positively influence students’ motivation and enjoyment (Frenzel et al., 2009). The expectations teachers have and the feedback given to students can impact their self-efficacy, emotions and academic achievement (Frenzel, Pekrun and Goetz, 2007). We must be careful when highlighting the gender differences in expectations; drawing attention to this area can cause more issues with the societal culture of mathematics and girls (Forgasz and Leder, 2010).

While self-efficacy is multidimensional, the societal and educational contexts must be considered when discussing female students’ motivational beliefs. Adopting a growth mindset in the classroom can help foster a sense of achievability for the subject content and reduce attainment gaps and negative stereotypes (Dweck, 2006, 2015). The growth mindset is contested by Sisk et al. (2018), who found that growth mindsets have a limited impact on academic achievement, especially in mathematics. Growth mindsets align with a mastery-orientated classroom goal, which shifts the focus of the classroom to gaining knowledge and mastering skills rather than on exam results, and challenges are viewed as opportunities for growth (Ames, 1992; Eccles and Roeser, 2009) and personal improvement is valued over competition and comparison (Murayama and Elliot, 2009). Lazarides and Watt (2015) discovered that mastery-orientated classrooms led to higher student motivation in mathematics, subsequently positively influencing mathematical career intentions. Equally, the self-efficacy and motivation of students are affected by sources outside of the control of a classroom teacher; Suárez-Álvarez, Fernández-

Alonso and Muñiz (2014) showed that one of the significant predictors of mathematical academic performance is the self-concept and motivation of students, but this was moderated by socioeconomic level.

Mathematics anxiety is a significant area of concern for the self-efficacy of mathematics students. For over thirty years, it has been known that mathematics performance throughout secondary teaching has correlated negatively with levels of mathematics anxiety (Hembree, 1990), but recent studies have shown that this is more significant in girls than boys (Devine et al., 2012; Dowker, Sarkar and Looi, 2016). Also, the gendered difference in mathematics anxiety seems to increase over time, with higher levels for girls being exhibited in primary school and these levels increasing through their educational journey (Hill et al., 2016). We must view this critically, with Hembree's (1990) study having a weaker positive correlation and many studies muddying the concepts of general test anxiety with mathematics anxiety; despite these methodological issues, it remains a strong foundation of understanding. Becker and Hall (2024) demonstrated that male-domain stereotypes affect not only performance but levels of anxiety for female students, and Beilock et al. (2010) indicate that parent and teacher expectations lead to higher levels of anxiety for girls. Despite issues surrounding social desirability bias in self-reported levels of anxiety, Frenzel, Pekrun and Goetz (2007) show that in a large number of students, even when outcomes are similar, girls still report higher levels of anxiety in mathematics. A growth mindset is a common intervention to reduce mathematics anxiety levels, and there have been good outcomes, but it remains unclear whether this is effective (Lee et al., 2021), but it is evident that early interventions and self-regulation are paramount for combatting mathematics anxiety (Hill et al., 2016; Lee et al., 2021). Recent studies indicate that the gender disparity in mathematics anxiety remains a problem, with many studies showing that boys have higher self-concepts and self-efficacy than girls (Becker and Hall, 2024). Interestingly in a longitudinal study by Wang et al. (2020), we see a negative reciprocal relationship between maths anxiety and self-perceived ability in boys, so high anxiety leads to lower a self-perception of your abilities, but this is not present for girls and indicates there is another underlying mechanism that drives mathematics anxiety.

Research gaps exist in female mathematics students' motivational beliefs and self-efficacy. Firstly, while the negative impact of gender stereotypes on girls' self-efficacy

is clear, deep investigations are needed to understand the nuances of this link in specific contexts (Heyder, Steinmayr and Kessels, 2019; Mejía-Rodríguez, Luyten and Meelissen, 2021). Longitudinal studies also remain a crucial area for improvement, especially concerning investigating the connection between self-efficacy beliefs in girls and STEM careers (Suárez-Álvarez, Fernández-Alonso and Muñiz, 2014). Also, proposed interventions for tackling this problem are effective, but it is evident that this needs to be expanded to show long-term effects (Lee, Lee and Bong, 2022; Becker and Hall, 2024).

Achievement Outcomes in Mathematics

In national and international surveys, there appears to be a difference in achievement based on gender. Cascella, Williams and Pampaka (2022) indicate that male students typically score higher in mathematics in many countries, with a few exceptions, but the gap varies significantly between countries; this is reflected in our national results as girls achieve nearly a quarter of a grade lower in GCSE mathematics than boys (Department for Education, 2024). Bergold et al. (2017) showed a difference in mean mathematics test scores, but the effect size was small; they did show an overrepresentation of male students at either end of the competency spectrum, but at the top end, they seemed more consistent internationally.

It is indicated that gender differences were more pronounced in gender-unequal societies; specifically, countries with higher gender inequality socially showed higher differences between male and female students in mathematics, indicating that societal push and pull effects have a significant role in mathematics education (Else-Quest, Hyde and Linn, 2010; Cascella, Williams and Pampaka, 2022). Forgasz and Leder (2017) indicate that societal expectations and stereotypes about gender roles significantly contribute to the observed differences in mathematics achievement between boys and girls. Cascella, Williams and Pampaka (2022) recommend analysing sub-national data as there were considerable differences in mathematical achievement based on socioeconomic and cultural differences in different regions. Else-Quest, Hyde and Linn's (2010) meta-analysis reinforces this perspective; the gender stratification hypothesis raises significant implications for policy and practice,

and the power of societal and cultural factors must be considered, but as this does not take into account all aspects of the achievement gap, it is important to use this in conjunction with other theories.

North and Ryan (2018) suggest that self-regulation may be one aspect of differences in self-efficacy. Girls were more prone to sensitivity from feedback, seeing it as a diagnostic of their ability, which could lead to higher levels of anxiety. However, girls used more self-regulation and self-monitoring strategies to aid their improvement, leading to better academic outcomes and more anxiety about pleasing others (ibid.). Self-regulation and perceived self-efficacy have been shown to correlate positively with self-regulation forming higher levels of motivation (Zimmerman and Martinez-Pons, 1990; Shores and Shannon, 2007; Hodis, 2018). A suggested approach is to foster motivation through constructive feedback inside a supportive classroom environment with a growth mindset approach (Hodis, 2018; Miller, Perera and Maghsoudlou, 2021). Moreover, integrating technology in classrooms to provide personalised feedback may enhance students' self-regulation skills, leading to improved outcomes (Miller, Perera and Maghsoudlou, 2021).

Teachers' beliefs and expectations influence student achievement in mathematics. Bandura (1997) demonstrates the power of self-efficacy to change how students think, feel and act, and teachers significantly influence this positively or negatively (Lazarides and Watt, 2015). Students with higher teacher expectations are likelier to engage in mathematics and perform well (Miller, Perera and Maghsoudlou, 2021). However, with boys receiving more positive reinforcement than girls (Heyder, Steinmayr and Kessels, 2019), it is clear that there is considerable inequity in teacher attention and expectations. Peterson et al. (2016) connect teacher expectations to students' mathematical outcomes, showing that explicit and implicit biases contribute considerably to the attainment gap. While gender is directly relevant to this topic, it is also clear that teachers' expectations differ by other characteristics, such as socioeconomic backgrounds and ethnicity; both of these also show a significant difference in outcomes from teachers' expectations (Rubie-Davies, Hattie and Hamilton, 2006; Sorhagen, 2013; Van Eycken, Demanet and Van Houtte, 2023). Recent studies that consider the history of research in this area are in consensus about how teacher expectations affect students' self-efficacy, which leads directly to differing outcomes in mathematics and, therefore, less representation in

STEM career trajectories (Sheu et al., 2018; Johnston, Wildy and Shand, 2019). To boil down student outcomes to just one factor is simplistic, and it is clear that we need to approach motivation as a multifaceted quantity that requires a student-centred perspective (Hodis, 2018).

This evidence must be scrutinised for validity and generalisability in the broader landscape of the literature. While it is clear that there is a difference in mathematics outcomes based on gender, it is debated whether gender has a significant effect size on standardised test scores (Suárez-Álvarez, Fernández-Alonso and Muñiz, 2014) or to what extent gender is the deciding factor in determining outcomes, especially when the intersectionality of many privileges makes it difficult to parse out (Cascella, Williams and Pampaka, 2022). The validity of a gender-neutral mathematics assessment is also questioned, especially when many studies we discuss involve standardised assessments and assessment structures (Leder and Forgasz, 2018). In conclusion, a holistic approach that combines multiple theories and perspectives is essential to fully understand and address the complexities of gender differences in mathematics achievement.

Intersectionality and Contextual Factors

As mentioned throughout this chapter, all gender inequities must be viewed with the intersectionality of the contributing factors in mind; seldom can gender be considered without other intersecting identities such as race, socioeconomic status and sexuality. Firstly, despite Becker and Hall's (2024) focus on the reconceptualisation of gender as a fluid, non-binary construct, they highlight the importance of gender intersecting with race and call for further empirical data to reinforce their theoretical propositions. Hsieh, Simpkins and Eccles (2021) indicate that the combination of gender alongside other identities, such as race, can significantly affect motivational beliefs. Also, the need for underrepresented ethnic groups in research to be included in future studies around gender is widely suggested (Casanova, Vukovic and Kieffer, 2021; Battey et al., 2022). Equally, the intersection with socio-economic status can be seen in Cascella, Williams and Pampaka (2022) study of local and international trends in mathematics attainment; while the nuance of why certain countries scored lower than others may be missing, the socio-economic variability within countries

proved explanatory. Intersectionality is also critical for literature validity due to the potential lack of investigation (Lindberg et al., 2010; Jaremus et al., 2020; Forgasz, 2021; Rubel and Bay-Williams, 2022).

Conclusion

Addressing gender stereotypes in mathematics education is paramount to cultivating an inclusive environment where all students can learn and tackling the underrepresentation of women in STEM careers. Both historically and contemporarily, gender stereotypes are present in mathematics classrooms, and these can affect teachers' beliefs, expectations and actions. Gendered expectations can lead to radically different self-efficacy and motivational beliefs in students, affecting mathematics outcomes. In conclusion, eliminating gender stereotypes in mathematics education can improve academic performance and greater self-confidence in female students. In order to truly consider our society equal, this must be addressed by all shareholders in the education of our students.

Chapter 3 – Literature Review – The Conceptual Framework

Introduction

This chapter will outline the theoretical and philosophical underpinnings of the proposed research into gender differences in mathematics education and detail the methodological approach to be adopted. Gender differences in mathematics education is a complex topic requiring a deep understanding of the interplay between theory, methodology, and researcher identity. This chapter will focus on the areas of Social Cognitive Theory (SCT) (Bandura, 1986), Expectancy-Value Theory (EVT) (Wigfield and Eccles, 2000), Stereotype Threat Theory (STT) (Steele and Aronson, 1995) and Implicit Theories of Intelligence (ITI) (Dweck, 1990) to form a conceptual framework from the synergy of these theories. An appropriate philosophical paradigm will be discussed, and the researcher's inherent ontological and epistemological stances will be described.

Theoretical Underpinnings

Social Cognitive Theory (SCT) was first described by Bandura (1986, 1997) as an answer to his perceived limitations of behaviourism and refocus on cognition, observational learning and self-efficacy beliefs; this last point is essential to this research. Self-efficacy is a person's perception of their ability to achieve; it has four primary sources of development:

- Mastery experiences – their own experiences of success and failure
- Vicarious experiences – observing others succeeding and how they got there
- Social persuasion – encouragement from teachers, parents and peers
- Physiological and emotional states – their mood and stress levels (Pajares, 1996)

These influences demonstrate that personal, behavioural and environmental factors can modify self-efficacy. Self-efficacy has been shown to affect academic outcomes significantly (Zimmerman, 2000), and Sheu et al.'s (2018) meta-analysis highlights the importance of mastery and social factors for self-efficacy in STEM subjects. This theory is critical to this research as it describes one mechanism for differing outcomes in mathematics education. Additional research highlights the intertwining

of self-efficacy with motivation and anxiety; Shores and Shannon (2007) show that motivation and anxiety considerably influenced mathematics achievement. Suárez-Álvarez, Fernández-Alonso and Muñiz (2014) demonstrate that a student's self-concept in mathematics, which includes their self-efficacy and motivation, is highly predictive of academic success. Sheu et al. (2018) added to our understanding of self-efficacy in STEM subjects, suggesting that mastery experiences and social persuasion are of utmost importance in mathematics; this highlights the crucial role of positive learning experiences and encouragement in shaping students' beliefs about their mathematical abilities. These findings underscore the multifaceted nature of self-efficacy and its critical role in shaping academic trajectories.

Expectancy-Value Theory (EVT) of education then builds on SCT (Wigfield and Eccles, 2000); this describes a student's motivation and achievement in a task based on their domain-specific, subjective expectancy, their self-efficacy and how much effort they are willing to use, and the perceived value of the task. The value consists of:

- Intrinsic value – their enjoyment or interest in the subject.
- Utility value – the task's relevance to future goals or needs.
- Attainment value – the personal significance of doing well.
- Cost – how much engaging in one activity limits access to other activities, effort required or emotional cost (Eccles and Wigfield, 2002; Wigfield and Cambria, 2010).

It is suggested that teachers, acting as socialisers, can influence children's expectancies and values by using their expectations and attitudes, and positive teacher evaluations can lead to more favourable self-concepts in students (Timmermans and Rubie-Davies, 2023). There is a gendered divide in beliefs and values that conform to gender stereotypes in students as young as six years old (Wigfield and Eccles, 2000). However, recent research by Timmermans and Rubie-Davies (2023) indicates that in mathematics, teacher expectations were higher for girls than for boys and that teacher expectations at the beginning of the year were predictive of their students' end-of-year achievement and self-concept; this supports the idea of socialisers' beliefs influencing students' self-concept. Becker and Hall (2024) imply the complex relationship between gender, motivation and achievement,

suggesting that cultural and societal factors are essential to consider and that intersectionality cannot be overlooked when assessing student motivation.

Almost thirty years ago, Steele and Aronson (1995) framed stereotype threat theory (STT) as the idea of an individual risking confirming negative stereotypes about their social group. Steele (1997) refined this with the notion that extended exposure to stereotype threat can lead to disidentification, where a person detaches their self-esteem from a stereotyped domain as a self-protective measure, reducing motivation and persistence. While these original studies focused on ethnicity, it has been shown that women are also affected by stereotype threat; when given a mathematics assessment, it was shown that if a gender divide expectation was verbalised at the start, a gender divide was observed in the results (Spencer, Steele and Quinn, 1999). Good, Aronson and Harder (2008) demonstrate that stereotype threat affects top-performing students, highlight the effect of teachers as exacerbating or mitigating influences and warn against teachers accidentally creating threatening environments. The extent to which this has a significant effect on students is questioned due to a large proportion of studies involving children having small effect sizes (Flore and Wicherts, 2015), but a significant body of the literature indicates that the adoption of growth mindsets and identity-safe classrooms would help alleviate the effects of stereotypes (Good, Aronson and Harder, 2008; Boucher et al., 2012). Pennington et al.'s (2016) review of psychological mediators in stereotype threat indicates that while the effects of stereotype threat are real, they may be more nuanced and context-dependent than initially thought.

Implicit Theories of Intelligence (ITI) are the beliefs about whether intelligence is fixed – entity theories – or can be developed – malleable or incremental theories (Dweck, 1990). Within the malleable theories lies the concept of growth mindset – a way to view intelligence as changeable and learning as improvable, which is suggested to positively impact academic outcomes and resilience (Dweck, 2006, 2015; Yeager and Dweck, 2012). Embedded incremental theories of intelligence held by students helped students improve their academic performance (Blackwell, Trzesniewski and Dweck, 2007), with Burnette et al. (2013) suggesting that students' theories of intelligence significantly predict achievement goals and intrinsic motivation in mathematics. Sisk et al.'s (2018) meta-analyses indicate that while growth mindsets may be effective, it is still unclear whether they are effective within

mathematics. The positive impact of incremental beliefs on resilience is undoubtable (Blackwell, Trzesniewski and Dweck, 2007; Yeager and Dweck, 2012), and their intrinsic motivation in mathematics is also noticeable (Burnette et al., 2013). This is important not just for the students, but teachers' ITI also can be shown to influence students' resilience and motivation, but teachers' beliefs can also be contradictory to their actions, leading to unforeseen outcomes (Copur-Gencturk, Thacker and Quinn, 2021). While the impact of a growth mindset is unclear, especially in different cultures (Costa and Faria, 2018), it remains evident that teachers should seek ways to foster incremental theories of intelligence (Yeager and Dweck, 2012).

Conceptual Framework

Using the four theories stated, we can create a conceptual framework to act as a lens through which we can observe the complex interplay of factors affecting gender differences in mathematics education. All theories address cognitive and motivational factors affecting academic achievement from different perspectives: SCT focuses on self-efficacy, EVT discusses motivational factors, STT highlights the power of negative stereotypes, and ITI explains the power of incremental growth mindsets. Evidently, there is a connection between all four theories that will be discussed individually. Firstly, the concepts of self-efficacy from SCT (Zimmerman, 2000) link closely with EVT's expectancy component, with teachers' gendered ability-based expectations influencing female students' self-efficacy and expectancy beliefs in mathematics (Lazarides and Watt, 2015). Furthermore, Jussim and Harber (2005) indicate that teacher expectations create self-fulfilling prophecies, which hints at an overlap between self-efficacy and stereotype threat theory, and stereotypes have been shown to impact female students' performance, motivation and intellectual and academic identity (Steele, 1997; Spencer, Steele and Quinn, 1999).

As mentioned previously, motivation is a significant factor across many theories. Whilst motivation is a pillar of EVT (Eccles and Wigfield, 2002), it remains an important element of Implicit Theories of Intelligence and self-efficacy (Dweck, 2006; Heyder, Steinmayr and Kessels, 2019). Costa and Faria's (2018) meta-analysis connects incremental theories to academic achievement, thus linking ITI and performance outcomes, especially in mathematics. Moreover, Yeager and Dweck

(2012) discuss how mindsets promote resilience and suggest that teaching students about the malleability of intelligence can improve academic resilience and performance, connecting ITI with SCT and EVT. Also, societal stereotypes and teacher expectations have been shown to affect academic performance (Copur-Gencturk, Thacker and Quinn, 2021). However, Lindberg et al.'s (2010) meta-analysis indicates that gender differences in mathematical performance are not biological but societal, thus linking STT to mathematics outcomes.

The four theories complement each other in many ways; both SCT and EVT explain how self-efficacy and expectancies influence motivation and performance, and STT and ITI show that stereotype threat potentially reinforces fixed mindsets about mathematical ability. Integrating these theories is vital in gaining a more comprehensive understanding of the factors influencing gender differences in mathematics education. SCT and EVT are valuable but do not address the societal influences that STT captures; equally, ITI alone will not give enough detail to explain the persistence of gender gaps. Together, these theories can form a foundation of understanding that will help us investigate this area further by understanding the relationship between teachers' gendered ability-based expectations and female students' motivational beliefs, self-efficacy and achievement outcomes in mathematics.

Philosophy of the Research

When considering the philosophical standpoint this research should adopt, it is important to consider the extent to which each paradigm would be appropriate. Firstly, a positivist paradigm has many advantages; positivism emphasises the importance of objective measurements (Cohen, Manion and Morrison, 2018), which is vital when analysing test scores or quantifying gender gaps in mathematics achievement (Cascella, Williams and Pampaka, 2022). Replicability is a central tenet of positivism, strengthening the generalisability of the findings on gender stereotypes in mathematics education (Grix, 2002). These points also link to quantifiable data that can identify trends and patterns in large-scale studies on gender differences in mathematics performance (Bergold et al., 2017). However, the oversimplification of positivism (Scotland, 2012) leads it to be inappropriate in this case, as its neglect of

context and limited exploration of subjective experiences mean it may not adequately capture the nuanced experiences and perceptions of students and teachers within a cultural and social context (Grix, 2002; Else-Quest, Hyde and Linn, 2010).

Interpretivism would also be advantageous if it were not for considerable limitations. Firstly, interpretivism focuses on rich, contextual understanding that allows for deep exploration of the social constructs associated with gender stereotypes (Scotland, 2012). The focus on subjective experiences could provide valuable insights into how students and teachers perceive and interpret gender-related expectations (Grix, 2002). Also, interpretivism's adaptable nature would allow a researcher to explore emerging themes and unexpected findings related to gender differences in mathematics (Scotland, 2012). The limitations of this paradigm outweigh these advantages: interpretivism's focus on specific context and individual experiences makes it challenging to form generalisable findings and highlight causal relationships within the data (Grix, 2002), and the investigative nature of the researcher can lead to researcher bias (Scotland, 2012), which has been highlighted within the current literature in this area (Frenzel, Pekrun and Goetz, 2007; Flore and Wicherts, 2015).

A more suitable approach is to adopt a paradigm between the two; thus, postpositivism or critical realism seems appropriate. Critical realism's strength is in the integration of objective and subjective realities: it acknowledges the existence of an objective reality, for instance, there is a measurable gender difference in mathematics achievement, but recognises that our understanding of this reality is socially constructed and fallible (Bhaskar, 1998; Fletcher, 2017). The generalisability of knowledge is counterbalanced with the importance of context in understanding social phenomena, which aligns well with the need to consider cultural and societal factors in gender and mathematics education research (Archer *et al.*, 1998). Also, the emancipatory potential of critical realism allows for identifying false beliefs and their causes, which could help address gender inequities in mathematics education (Bhaskar, 1998).

The stratified ontology of critical realism, as described by Bhaskar (1998), is a key feature which posits three distinct domains of reality:

- The empirical – observable experiences and events, in our context this could be test scores, classroom behaviours or survey responses

- The actual – events and experiences that happen, whether or not they are observed, for example, unobserved classroom interactions or unrecognised biases
- The real – underlying mechanisms, structures, and tendencies that generate the events in the actual domain, such as societal gender norms, cognitive processes, and institutional structures that influence gender differences in mathematics education.

These three levels help researchers to investigate the deeper causal mechanisms rather than just following surface-level observations (ibid.) and make steps towards a more comprehensive understanding of the complex interplay between individual experiences, social structures and underlying generative mechanisms (Archer *et al.*, 1998). The epistemology stance of critical realism is based on relativism; there is an objective reality, but our knowledge of that reality is always conceptually mediated, socially situated and fallible, and our understanding is always partial and subject to revision (Maxwell and Mittapalli, 2015; Fletcher, 2017). The fundamental mode of inference in critical realism is retroduction, which involves identifying mechanisms that could explain phenomena (Fletcher, 2017); in mathematics education, this could be theorising about the underlying social or psychological mechanisms that could explain observed gender differences in mathematics performance.

While critical realism seems to be the most appropriate fit for a research paradigm into gender disparities in mathematics education, it is prudent to have a critical approach to the limitations and how they should be addressed going forward. The flexibility of this paradigm is mirrored in its complexity to execute effectively; it is suggested that this be addressed with clear explanations of methodological approaches (Fletcher, 2017). Furthermore, there can be a tendency to overreach and make overly broad claims without sufficient evidence, but reflexivity and openness about the limitations of findings are paramount for addressing this (Archer *et al.*, 1998). Also, the limit to which a researcher can empirically verify deep structures and mechanisms is a criticism that calls for robust evidence and openness to alternative explanations (Cruickshank, 2004). Despite these limitations, it is the most appropriate paradigm for investigating gender differences in mathematics education.

Methodology

The overall methodological approach within a critical realist framework is centred around a mixed methods approach, as it allows for the integration of quantitative and qualitative data to provide a more comprehensive understanding of the factors involved in gender differences in mathematics education (Mcevoy and Richards, 2006; Creswell and Creswell, 2018). This combination enhances the robustness of the research by enabling a more nuanced exploration of both observable patterns and underlying causal mechanisms (ibid.). Danermark, Ekström and Karlsson (2019) indicate that a mixed methods approach within a critical realist position is helpful because we can use different methods to investigate the different aspects of reality, thereby providing a richer and more accurate understanding of complex factors. Using mixed methods can help uncover the patterns of gender differences in mathematics education with quantitative methods and the underlying mechanisms and contexts that produce these patterns through qualitative methods (Zachariadis, Scott and Barrett, 2013).

The extensive (quantitive) data in this research could be student assessment results or the numbers of each gender who are joining higher mathematics courses, and the intensive (qualitative) data could be interviews or observations of classroom dynamics (Fletcher, 2017). Surveys can help identify patterns in teachers' expectations and students' self-efficacy, while interviews and observations can provide insight into the contextual factors and mechanisms influencing these patterns (Smith and Elger, 2012). Fletcher (2017) describes this process as looking for demi-regularities, trends or patterns in data and then using a system of abduction to theorise about potential mechanisms, such as stereotype threat or gendered self-efficacy beliefs, and retroduction to investigate the contextual conditions in which these mechanisms operate.

This range of data can prove to be a limitation itself; integrating quantitative and qualitative data can be challenging and may cause tensions between different assumptions and theories (Maxwell and Mittapalli, 2015), but a critical realism framework reconciles these conflicts: Scotland (2012) suggests that the researcher must critically analyse the participants' interpretations but also consider broader structural influences; thus, a flexible approach to coding is required to ensure that

both individual and societal factors are considered carefully and that any theorised mechanism is tested against collected empirical data (Fletcher, 2017).

This methodology can be used to investigate the observable differences in mathematics achievement between genders and the underlying social, cultural and psychological factors that contribute to these differences, providing a robust framework for capturing both empirical evidence and theoretical insights to address the complexities of gender disparities in education.

Researcher Identity

The journey towards an appropriate paradigm reflects my evolving identity as a researcher committed to exploring complex educational phenomena through a lens that integrates theory and practice. At first, being a mathematician, I naturally saw things through a positivist lens, but as this only produced the “what”, I realised that limitations were considerable for this paradigm. Through discussions with other students during a group assignment, I appreciated the effect of nuance and interpretivist outcomes. Exposure to research paradigms and methodologies has led to a better approach to research, enabling a selection of a critical realist paradigm for gender differences in mathematics (Kempster and Parry, 2011; Hoddy, 2019). Castelló et al. (2021) indicate that researcher identity can be considered multidimensional, highlighting four distinct dimensions, of which I consider myself to have developed a more dynamic view of my researcher identity, which shifts over time. Also, while I have always been aware of the multiplicity of my researcher identity, it has become more integrated as I have progressed through this year of study. However, I consider myself to be working towards the action aspect of research rather than the thinking aspect when considering identity trajectories over time (McAlpine, 2012). This research was chosen out of my professional experience, leading me to an insider-researcher perspective (Bridges, 2009), but through analysing and synthesising literature, it has become a much more complex viewpoint. My understanding of reflexivity and positionality has deepened from initial exposure to current comprehension, leading me to confront my assumptions and biases (Mason-Bish, 2019).

Conclusion

This area of study requires a conceptual framework built from Social Cognitive Theory (Bandura, 1986), Expectancy-Value Theory (Wigfield and Eccles, 2000), Stereotype Threat Theory (Steele and Aronson, 1995) and Implicit Theories of Intelligence (Dweck, 1990) in order to interrogate gender differences in mathematics education effectively. This leads to the objective ontology and subjective epistemology of a critical realism paradigm (Bhaskar, 1998), which leads the researcher to adopt a mixed methods methodological approach that can question the foundational mechanisms at play (Creswell and Creswell, 2018). These choices require a robust methodology to enable the research to be effectively undertaken and address all limitations of this paradigm. Future chapters will go into the specific methodological and ethical considerations that must be considered.

Chapter 4 – The Research Proposal

This dissertation is titled “Bridging the Mathematics Gender Divide: A Mixed Methods Study of Teacher Biases and Their Impact on Female Students”, which reflects the focus on teacher expectations, student self-efficacy, and stereotype threat. The area of research was of initial interest due to my professional position as a secondary mathematics teacher who is required to analyse trends in school data but also confront national data, which shows a historical difference by gender in mathematics education at this level (Department for Education, 2024). Through engaging with the literature to determine why this may be the case, my review revealed persistent gender stereotypes in mathematics education (Forgasz and Leder, 2017), which led to a deeper understanding of the mechanisms at play. This engagement enabled me to refine research ideas by critically analysing current research to determine gaps, contradictions and areas that need further exploration (White, 2009).

From the extensive exploration, four compatible theories emerged to form a conceptual framework for viewing this phenomenon: Social Cognitive Theory (Bandura, 1986), Expectancy-Value Theory (Wigfield and Eccles, 2000), Stereotype Threat Theory (Steele and Aronson, 1995), and Implicit Theories of Intelligence (Dweck, 1990), this was integral for framing my study (Punch, 2014) and contributed to a greater understanding of gender differences in mathematics education. These theories then formed the basis of the choice of paradigm to adopt; Grix (2002) describes the strength of critical realism as how the paradigm allows for objective and subjective perspectives in examining this topic, with the ontological standpoint being objectively real, while our epistemology can be subjective in how we view it. Throughout this process, it was important for my own curiosity and openness to surprise to help drive the research process and increase my understanding of reflexivity (White, 2009). Bryman (2007) discussion of particularistic and universalistic approaches to research enabled an element of reflection around the methodological choices made and helped refine a model of literature driving research questions and research questions driving methodological approaches – the researcher’s preferences are not important to this process. I must acknowledge my subjective standpoint on the journey to objectivity (Berger, 2015). A mixed methods approach seems logical for this investigation as it will provide a more comprehensive

understanding of the area than a single method would (DeCuir-Gunby and Schutz, 2018).

The research question has evolved throughout this process, with an initial emphasis on “what”, leading to a “how” or “why”. This echoes my journey from the identity of a mathematics educator to more of my developing researcher identity. The iterative process of writing a research question has been guided by an exploration of the literature, which, while international in scope, struggles to be specific enough to my national viewpoint. This research question will address the issue of UK-based mathematical education and investigate why within our cultural context. I do not see this as derivative of other studies, as the factors I am investigating do not appear in this combination in most of the literature, never mind in the UK context and a mixed method approach

The final research question is:

“How do British secondary mathematics teachers’ gendered ability-based expectations influence female students’ motivational beliefs, mathematical academic self-efficacy, and achievement outcomes?”

With sub-questions:

1. “To what extent do mathematics teachers’ assumptions about innate skills exhibit gender biases reflecting societal stereotypes about mathematical abilities?”
2. “How do teachers’ gendered expectations manifest in classroom practices and environments, and how do these influence female students’ mathematics anxiety and engagement?”

Chapter 5: Research Design, Research Methods and Methods of Analysis

Overall Design Frame

This mixed methods study will adopt an explanatory sequential design, allowing a comprehensive understanding of the complex relationships in gender differences in mathematics education (Creswell and Plano Clark, 2018). This design allows for the collection and analysis of quantitative data in the first phase, followed by a second phase for qualitative data for explanation and elaboration (Ivankova, Creswell and Stick, 2006). This mixed methods design will help overcome the limitations of single method designs, offering more robust inferences through triangulation, quantitative data's generalisability and in-depth qualitative analysis while minimising the weaknesses of each individually (Johnson and Onwuegbuzie, 2007; DeCuir-Gunby and Schutz, 2018). This two-phase explanatory sequential study will have an initial phase of quantitative data collection, which will provide a general understanding of the research problem and aid in selecting participants for qualitative follow-up (Ivankova, Creswell and Stick, 2006). The qualitative second phase refines and explains the statistical results (ibid.) in a convergent parallel multiple-case study approach, collecting and analysing interview and observational data simultaneously (Creswell and Plano Clark, 2018). This study is grounded in a critical realist paradigm, supporting the integration of objective and subjective approaches (Bhaskar, 1998) while providing flexibility in choosing methods that best address this problem while acknowledging the existence of an objective reality (Maxwell and Mittapalli, 2015). The main emphasis of this paradigm is uncovering and understanding the underlying mechanisms and structures in mathematics education (Archer *et al.*, 1998).

While data integration will be the most significant challenge with this approach, it can be completed by connecting, building and merging data from both phases of the study to gain a more nuanced understanding of the mechanisms in play (Fetters, Curry and Creswell, 2013). Data will be integrated at multiple points throughout the study: data collection, analysis and interpretation and a fully integrated data set enhances the value of mixed methods research (Bryman, 2007; Creswell and Plano Clark, 2018). The quantitative and qualitative data are utilised to investigate different aspects of gender differences in mathematics education, but their insights will be

complementary (DeCuir-Gunby and Schutz, 2018). This is shown clearly in the connecting point between the phases by selecting participants and developing interview protocols based on the quantitative data (Ivankova, Creswell and Stick, 2006). Therefore, a single-method approach would be insufficient for an investigation into the complex interplay of factors in mathematics education; a mixed-methods design provides a more comprehensive understanding of multifaceted relationships (Greene, Caracelli and Graham, 1989) and allows for triangulation of complementarity data that allow for the expansion of findings (Creswell and Plano Clark, 2018). A mixed methods design also allows for simultaneous investigation of confirmatory and exploratory research questions (Teddlie and Tashakkori, 2009). This approach enables examining both patterns and processes in gender differences in mathematics education, providing a more comprehensive understanding (DeCuir-Gunby and Schutz, 2018). A convergent parallel approach in the qualitative phase allows for the capture of a holistic snapshot of teacher perspectives and classroom practices (Creswell and Plano Clark, 2018).

Quantitative Phase

The explanatory sequential design indicates that this study will fall into a quantitative phase followed by a qualitative phase, which will be used to explain the quantitative data (Creswell and Plano Clark, 2018). This is firmly grounded in a critical realism paradigm's realist ontology and constructivist epistemology (Bhaskar, 1998; Fletcher, 2017), which indicates that gender differences have real causes and effects, but participants subjectively interpret these. Multiple quantitative data collection streams will be created during the first phase, followed by the first data analysis point.

The first phase will have two sources of data. Firstly, in the setting of this study, students' current academic data will be collected to provide a baseline and valuable comparison when combined with enrolment data and other data sources (Fletcher, 2017). Understanding the ethical implications of using secondary data, such as ensuring data protection and student privacy (Cohen, Manion and Morrison, 2018), is vital. However, as this data is an initial point of contrast, it is clear that as long as the data is internally valid, it will provide valuable comparative data on students' objective attainment (ibid.).

Additionally, two web-based cross-sectional surveys will be undertaken to investigate both students' and teachers' gender-based views and biases. Compared to a paper-based survey, a web-based survey is cheaper, easier to distribute, and easier to collect data (Cohen, Manion and Morrison, 2018). However, the main advantage is that a web-based survey can help reduce social desirability bias through its format (van den Bergh *et al.*, 2010; Koivula, Räsänen and Sarpila, 2019), but a clear plan is required to ensure a high uptake. Both surveys will require certain demographic information to help form samples and themes for the second phase of this study. Also, a 5-point Likert (Likert, 1932) scale enables participants to indicate their agreement along a scale, although an argument can be made for a 7-point scale to add more nuance to the findings.

The questionnaire given to the students (Appendix A) draws from many different scales and instruments. Firstly, the questionnaire investigates mathematics self-efficacy and anxiety by adapting the Mathematics Self-Efficacy and Anxiety Questionnaire (MSEAQ) (May, 2009) alongside the Mathematics Anxiety Rating Scale-Short (MARS-Short) (Richardson and Suinn, 1972) to question how students feel when the encounter mathematics in their lives, this aligns with the importance of Bandura's (1986) self-efficacy within SCT. Next, the Attitudes Toward Mathematics Inventory (ATMI) (Tapia and Marsh, 2004) questions students regarding four aspects of their feelings towards mathematics: self-confidence, value, enjoyment and motivation, all of which help to ground the investigation using the lens of SCT (Bandura, 1986) and EVT (Wigfield and Eccles, 2000). The Fennema-Sherman Mathematics Attitudes Scale-Short Form (FSMAS-SF) (Fennema and Sherman, 1976) investigates the innate stereotypes that are prevalent in the student population, which reflects the STT (Steele and Aronson, 1995) aspect of our concept framework, enabling the researcher to probe into cultural values. Finally, the questionnaires have a series of questions that reflect the aspirations and expectations of students within the classroom and education in general, and this leads into an open-ended section that helps clarify some of the other data points and aids in selecting participants for the second phase.

The teacher questionnaire (Appendix B) is designed similarly to the student version, but the elements are more general rather than specific to an individual class. Firstly, Nurlu's (2017) Teacher's Gender Stereotype Scale toward Mathematics is adapted to

investigate gender stereotypes of mathematics in four areas of focus: environment, career, competence and attribution. These four areas will help generate a specificity to the gender perceptions prevalent in the teacher population. Rubie-Davies, Hattie and Hamilton (2006) inspire the second section, with expectations on what teachers expect of their students based on gender, specifically in success vs longevity of study. The third section is a reflective section based on Copur-Gencturk, Thacker and Quinn (2021) paper on teachers' awareness of gender diversity in mathematics education and specifically questions teachers on their practice in the classroom.

Qualitative Phase

Following an analysis of the data collected in the first phase, the second phase will take the form of a multiple-case study approach, and this will enable the in-depth exploration of quantitative results from the first phase and a focus on the varying levels of gender disparities (Yin, 2013) and purposive sampling (Denzin *et al.*, 2023) will be adopted to identify lines of inquiry relevant to the data collected. The primary data sources during this phase will be semi-structured interviews with teachers and students and lesson observations.

The semi-structured interviews will be taken in-person and in-private within the school setting to provide more of an opportunity to create rapport with participants and make it easier to investigate (Krämer *et al.*, 2016; Cohen, Manion and Morrison, 2018). A semi-structured interview gives flexibility for investigation (Cohen, Manion and Morrison, 2018) but will require an interview guide (Appendices C and D). While the nature of the questions is highly dependent on the quantitative data collected in the first phase, the appendices show a range of questions that could be included from the literature that influenced the surveys. The interviews may also investigate intersectionality with more clarity than a questionnaire could, as suggested by Becker and Hall (2024). For student interviews, the argument can be made for focus groups based on findings from the initial quantitative phase.

Observations will also be undertaken to help triangulate findings by introducing a window into the classroom dynamic. An observational protocol (Appendix E) will be used to ensure consistent data collection (Cohen, Manion and Morrison, 2018).

These observations will remain anonymous, but through multiple observations, a data set will be developed for analysis (ibid.)

While these instruments have been produced (Appendices A to E), they should be developed and adapted based on expert consultation and a proposed pilot to suit the context in which this study will be undertaken (DeCuir-Gunby and Schutz, 2018).

Research Participants and Context

This study focuses on secondary school mathematics teachers and students aged between fourteen and sixteen; this encompasses students during key stage four up to their GCSE exams. This age range represents a critical period in mathematics education and career choices (Cohen, Manion and Morrison, 2018), and attitudes at this point can significantly impact future academic and career trajectories. The target sample size for the quantitative phase is approximately 300 students and 50 teachers across three to five schools in the West Midlands region of the United Kingdom, but these numbers will be confirmed using a power analysis (Cohen, 1988). Within these schools, a stratified sample will be taken to ensure representation across different demographics and geographic regions; this will enhance the generalisability of the findings and address potential sample biases (Cohen, Manion and Morrison, 2018) while also ensuring the scale of the study is manageable within the resources available. The second phase will take a purposive sample based on the findings of the first phase to triangulate and form an explanation of the results (Creswell and Plano Clark, 2018). Depending on uptake, a statistical weighting could be applied to reduce the effects of under- and over-sampling of underrepresented groups, which would form a sampling bias; however, this would only be applied if necessary and would be carefully implemented transparently (Marks, 2008; Cohen, Manion and Morrison, 2018).

Data will be collected over an academic year, and careful consideration will be given to the timing of the data collection to avoid disruption during exam periods. Teaching practices, resources, and the overall academic climate are important factors to focus on while in different schools to address the potential impact of school culture (Else-Quest, Hyde and Linn, 2010). Special consideration should also be given to the characteristics that may impact data quality; this could be due to the reading age of

participants or other factors that may influence their ability to provide accurate self-reports (Cohen, Manion and Morrison, 2018). To address this, all data collection methods will be age-appropriate, and questions will be framed in a way easily understood by the target age group (Cohen, Manion and Morrison, 2018; Tourangeau, 2020). Another consideration is the impact of the Hawthorne Effect on participants' actions, which indicates that when a participant knows they are being observed, they will act differently (McCambridge, Witton and Elbourne, 2014). This would be challenging to mitigate entirely, but data collection will be designed to be as unobtrusive as possible and occur in a naturalistic setting (ibid.); the participants will be treated with respect, and all ethical guidelines will be followed. The aims and outline of the study will be shared with participants (BERA, 2024), but the hypotheses of the study may be withheld to enable more typical responses, but this will be only undertaken following extensive consultation with experts (Cohen, Manion and Morrison, 2018).

Proposed Methods of Data Analysis

The mixed methods nature of this study requires two periods of data analysis, one between the phases that will drive the direction of the second phase and a second period after the second phase to be used as an explanatory data set alongside the quantitative data.

Quantitative data will be collected in the first phase of this explanatory sequential mixed methods study. Firstly, descriptive statistics to summarise demographic data and key variables will provide an overall picture of the sample and distribution, using such measures as central tendency and dispersion (Cohen, Manion and Morrison, 2018). Next, inferential statistics will be generated using t-tests to examine group differences based on gender and other factors to understand whether there are statistically significant differences in mathematics performance or attitudes (Creswell and Plano Clark, 2018). Other statistical tools could be utilised to investigate further, such as discriminant function analysis (Ivankova, Creswell and Stick, 2006), which would enable the identification of which factors best predict high or low achievers or a structural equation model to allow for an analysis of interrelationships between multiple variables (Klein, 2016).

Qualitative data will be in the form of interview transcripts and observation notes. A thematic analysis of the coded interview transcripts will identify recurring themes and patterns in participants' experiences and perspectives (Creswell, 2014). Content analysis will be carried out on observation notes to systematically categorise and interpret classroom dynamics and teaching practices (Cohen, Manion and Morrison, 2018). In order to develop a grounded theory, an iterative process will compare new data with previously collected data and emerging categories to develop a theoretical framework grounded in the data (ibid.).

The integration of findings across the two phases of this study will be undertaken in a convergent design to compare and contrast quantitative and qualitative results (Creswell and Plano Clark, 2018). A narrative integration will then explain how qualitative findings elaborate or explain quantitative results, especially focusing on unexpected outcomes (Fetters, Curry and Creswell, 2013). One way of forming this may be using a joint display to bring both sides of the data together for consideration (Guetterman, Fetters and Creswell, 2015). Data integration will fall into three strategies:

- Connecting – selecting participants from the first phase to be focused on in the second phase enables a more profound exploration (Ivankova, Creswell and Stick, 2006).
- Building – the qualitative findings are used to explain or elaborate on quantitative findings (Creswell and Plano Clark, 2018).
- Merging – Combining the two data sets to create a comprehensive picture of the complex factors influencing gender differences in mathematics education.

Credibility, Trustworthiness, and Confirmability

Ensuring credibility, trustworthiness, and confirmability of research is crucial in educational studies, particularly when investigating gender differences in mathematics education. Firstly, all results will be formed by careful triangulation of findings across the phases of this explanatory sequential study to corroborate findings and enhance the overall research quality (Creswell and Plano Clark, 2018); the range of data sources will provide a comprehensive understanding of this area of study and employing both quantitative and qualitative methods aids in creating a

holistic view of gender differences in mathematics education. The fundamental principle will be based on the researcher's reflexivity; through first-person language, transparent reporting and acknowledgement of the researcher's position and potential influences will lead to continual reflection (Berger, 2015; Cohen, Manion and Morrison, 2018).

This study will adopt various other techniques to ensure credibility, trustworthiness, and confirmability. The validity of the mixed-methods approach will be ensured by a careful approach to sample integration and techniques that reduce researcher bias and strive to achieve inside-outside legitimation (DeCuir-Gunby and Schutz, 2018). Firstly, peer researchers will be consulted throughout the study to help identify potential biases, assumptions or oversights in the researcher's analysis (Berger, 2015; Creswell and Plano Clark, 2018), the use of a three-part log to document participant statements, interpretations, and researcher reflections (Berger, 2015) as part of a detailed audit trail documenting decisions, processes, field notes, and coding schemes throughout the study (Creswell and Plano Clark, 2018). The surveys will be refined to be short and simple, using clear language and avoiding technical language or jargon (Cohen, Manion and Morrison, 2018). Techniques for minimising non-response and data cleaning will be used to ensure an optimum data set (ibid.), but the main challenge will be addressing biases in self-reported data to enhance the data quality. To tackle this, a series of techniques could be adopted:

- A counterbalancing technique to vary the order of questions across participants (Dillman, Smyth and Christian, 2016).
- Nurlu (2017) suggests using expert panels to review and refine item wording and instructions.
- Include questions to address implicit biases, stereotype prevalence and social desirability scales (Crowne and Marlowe, 1960; Greenwald, McGhee and Schwartz, 1998; Inglis and O'Hagan, 2022).

Interviews will adopt a series of measures to ensure the validity of the results. Firstly, member checking will be used to clarify if the participants' comments align with the researcher's interpretation (Creswell and Plano Clark, 2018) and Berger's (2015) suggestion of repeated review of interviews after a lapse in time. A conscientious awareness is also needed concerning the power dynamics of interviews with

children, which will require strategies to mitigate power imbalances, such as clear communication about the voluntary nature of participation, assurances of confidentiality, and situating interviews in a comfortable and non-threatening environment (Kivunja and Kuyini, 2017).

Ethical Considerations

This study will adhere to the ethical guidelines set forth by the British Educational Research Association (BERA, 2024) and the Economic and Social Research Council (ESRC, 2015 enabling research to be carried out with respect. Appendix F is this study's ethical appraisal form.

Firstly, obtaining informed consent is a crucial ethical requirement as an ongoing process (Cohen, Manion and Morrison, 2018). This requires all information about the research to be age-appropriate, transparent and available to all potential participants, including students, teachers, and parents and guardians (BERA, 2024). The understanding that participation in the study is voluntary and free from coercion is required (ESRC, 2015). Written consent is required for all adult participants, parental consent is needed for student participants (Cohen, Manion and Morrison, 2018), and ongoing assent from student participants throughout the research process (Graham, Powell and Taylor, 2015).

Secondly, protecting the participants' privacy is essential, and all participants will have their confidentiality and anonymity protected throughout the study (Cohen, Manion and Morrison, 2018). All data will be stored securely, with access restricted to authorised research team members (ESRC, 2015), and all GDPR regulations will be followed (UK Government, 2018). Pseudonyms or codes will be used to anonymise all participant data (British Psychological Society, 2021), and no individual or school will be identifiable in any published results (Cohen, Manion and Morrison, 2018).

A duty of care is required where harm is minimised to participants while maximising the benefits of the study. The impact of the research activities will be considered in the planning to ensure that participants' time and workload are not affected (BERA, 2024). Also, as this study investigates gender stereotypes, special thought will be

taken to be sensitive to the potential for reinforcing gender stereotypes through research questions and practices (Hammersley and Traianou, 2015). Appropriate support or referrals will be provided if sensitive issues arise during the research (British Psychological Society, 2021).

To enhance the ethical framework of this study, ongoing ethical reflection will be embedded into the organisation of the study. Regular team meetings that reflect on ethical concerns (BERA, 2024), consultation with institutional ethics committees for more complex ethical dilemmas (ESRC, 2015) and using reflective journals to document ethical decisions and rationales will be adopted (Cohen, Manion and Morrison, 2018). Also, Stutchbury and Fox's (2009) ethical analysis framework will be utilised to address ethical issues from multiple philosophical perspectives. Also, all research findings will be made available to participants; this may be through email or a study website (BERA, 2024).

By adhering to these ethical principles and practices, this study aims to conduct rigorous and valuable research while respecting the rights and dignity of all participants involved in exploring gender differences in mathematics education.

Postscript – Narrative Critical Reflection

Category	Feedback received, targets achieved and areas of development worked on	How did this shape my dissertation?
Knowledge and understanding: Targets, reflections or feedback relating to knowledge of current debate and issues in your specific area of focus; drawing out concepts and themes; choosing a focus area for your dissertation; identifying and overcoming ethical issues.	TMA 02 feedback: "…both your topic and the design issues are thoroughly examined and critiqued although there are some omissions, for example, race/ethnicity and class don't feature as such, you refer to 'Europe' without specificity and you confuse England with the UK despite the fact that the four nations have their own systems, histories, cultures and school structures."	I began using a table to identify the locations and themes in each paper I read. I used this feedback to create a diagram of the specifics of my research proposal. Through this process, it became clear exactly what areas of the world and what topics were prevalent, and the scope of my research proposal was too broad and unclear. I then spent time highlighting where the gaps in the research were with specificity.
Critical analysis and evaluation: Targets, reflections or feedback relating to justifying or challenging your personal perspective; interpreting and critically analysing evidence and methodologies from your own and others' research; analysing and evaluating themes and issues; sourcing and critically reviewing a wide range of publications; creating an academic argument using synthesis; comparing and connecting practice and theory.	TMA 01 feedback: "With regard to criticality, I think you demonstrate this in your treatment of the selected papers and also with regard to your possible choice of hypothetical research design, though the latter will need a bit more thought as I think you are interested in connecting with practice and so methodology will be important in what you propose." TMA 02 feedback: "Your literature review and design both show that you can analyse, synthesise and develop your own 'voice' although occasionally I thought that voice had a tinge of passionate conviction	I had initially set myself a PDP target to develop criticality in my writing. This has been a journey for me through this year of study as indicated by the two pieces of feedback. My criticality initially was potentially too negative, as indicated by "passionate conviction" rather than the outsider perspective I was hoping for. I read many critical analyses to understand the style I was attempting to develop. I adopted a discussion style that indicates fair consideration of all papers read.

	rather than the clinical and forensic."	
Links to professional practice: Targets, reflections or feedback relating to: designing and/or applying research methods; developing ideas from previous research and frameworks; reflecting and making adaptations during the research and writing process; addressing problems in research design; identifying implications for practice and professional debate; challenging your own assumptions; managing workload and personal motivation.	TMA 01 feedback: "Application to context is fine but a little more detail on the teaching context needed as I wasn't sure about factors such as class and race intruding into the neat category of 'girls'." EMA feedback (See Appendix G): "Some comment on what the research findings or debates might imply for practice would be welcome but may not be right for this particular extract."	My researcher identity has been a little at odds with my professional identity as a mathematics teacher for a significant period of time during this year of study. In order to try and develop a more robust sense of researcher identity, I have attempted to distance myself from my profession. From the feedback here, it is clear that this approach led to a lack of application of theory to practice. From this feedback, I have taken a careful approach to consider how theories can be contextualised.
Structure, communication and presentation: Targets, reflections or feedback relating to using academic style and referencing; presenting, managing and sharing information in different modes; communicating concepts, findings and ideas for different audiences.	My previous module indicated that referencing could be clearer by using a reference manager. At the beginning of the dissertation module, I set myself targets to synthesise ideas coherently rather than quote directly from papers.	I have utilised a reference manager to ensure that no references were missed from my reference list. Also, I began collecting ideas as direct quotes and then grouped themes to be coherent rather than seeming like a list of thoughts. This has been a process I have developed over the dissertation module, but it has become easier to manage for this dissertation.

References:

Ames, C. (1992) ‘Classrooms: Goals, Structures, and Student Motivation’, *Journal of Educational Psychology*, 84(3). Available at: https://doi.org/10.1037/0022-0663.84.3.261.

Archer, M. *et al.* (1998) *Critical Realism: Essential Readings (Critial Realism: Interventions)*, *Centre For Critical Realism*.

Bandura, A. (1986) ‘Social foundations of thought and action : a social cognitive theory / Albert Bandura.’, *New Jersey: Prentice-Hall, 1986*, 16(1).

Bandura, A. (1997) ‘Self-efficacy: The exercise of control. New York: W.H. Freeman and Company’, *American Psychological Association*, 23.

Barth, J.M. *et al.* (2018) ‘Matching Abilities to Careers for Others and Self: Do Gender Stereotypes Matter to Students in Advanced Math and Science Classes?’, *Sex Roles*, 79(1–2). Available at: https://doi.org/10.1007/s11199-017-0857-5.

Battey, D. *et al.* (2022) ‘Racialized and Gendered Labor in Students’ Responses to Precalculus and Calculus Instruction’, *Journal for Research in Mathematics Education*, 53(2). Available at: https://doi.org/10.5951/jresematheduc-2020-0170.

Becker, J.R. and Hall, J. (2024) ‘Research on gender and mathematics: exploring new and future directions’, *ZDM - Mathematics Education*, 56(1). Available at: https://doi.org/10.1007/s11858-023-01510-6.

Beilock, S.L. *et al.* (2010) ‘Female teachers’ math anxiety affects girls’ math achievement’, *Proceedings of the National Academy of Sciences of the United States of America*, 107(5). Available at: https://doi.org/10.1073/pnas.0910967107.

BERA (2024) *Ethical Guidelines for Educational Research, fifth edition (2024)*, *https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-fifth-edition-2024*.

Berger, R. (2015) ‘Now I see it, now I don’t: researcher’s position and reflexivity in qualitative research’, *Qualitative Research*, 15(2). Available at: https://doi.org/10.1177/1468794112468475.

van den Bergh, L. *et al.* (2010) ‘The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap’, *American Educational Research Journal*, 47(2). Available at: https://doi.org/10.3102/0002831209353594.

Bergold, S. *et al.* (2017) ‘Academic competencies: Their interrelatedness and gender differences at their high end’, *Journal of Educational Psychology*, 109(3). Available at: https://doi.org/10.1037/edu0000140.

Bhaskar, R. (1998) ‘Philosophy and Scientific Realism. Critical Realism: essential readings’, *Critical Realism: Essential Readings* [Preprint].

Blackwell, L.S., Trzesniewski, K.H. and Dweck, C.S. (2007) ‘Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study

and an intervention', *Child Development*, 78(1). Available at: https://doi.org/10.1111/j.1467-8624.2007.00995.x.

de Boer, H., Bosker, R.J. and van der Werf, M.P.C. (2010) 'Sustainability of Teacher Expectation Bias Effects on Long-Term Student Performance', *Journal of Educational Psychology*, 102(1). Available at: https://doi.org/10.1037/a0017289.

Boucher, K.L. *et al.* (2012) 'Reducing stereotype threat in order to facilitate learning', *European Journal of Social Psychology*, 42(2). Available at: https://doi.org/10.1002/ejsp.871.

Bridges, D. (2009) 'Education and the possibility of outsider understanding', *Ethics and Education*, 4(2). Available at: https://doi.org/10.1080/17449640903326714.

British Psychological Society (2021) *BPS Code of Human Research Ethics [2nd-edition-2014]*, *British Psychological Society*.

Bryman, A. (2007) 'The research question in social research: What is its role?', *International Journal of Social Research Methodology*. Available at: https://doi.org/10.1080/13645570600655282.

Burnette, J.L. *et al.* (2013) 'Mind-sets matter: A meta-analytic review of implicit theories and self-regulation', *Psychological Bulletin*, 139(3). Available at: https://doi.org/10.1037/a0029531.

Casanova, S., Vukovic, R.K. and Kieffer, M.J. (2021) 'Do girls pay an unequal price? Black and Latina girls' math attitudes, math anxiety, and mathematics achievement', *Journal of Applied Developmental Psychology*, 73. Available at: https://doi.org/10.1016/j.appdev.2021.101256.

Cascella, C., Williams, J.S. and Pampaka, M. (2022) 'Gender differences in mathematics outcomes at different levels of locality to inform policy and practice', *European Educational Research Journal*, 21(5). Available at: https://doi.org/10.1177/1474904121997211.

Castelló, M. *et al.* (2021) 'What perspectives underlie "researcher identity"? A review of two decades of empirical studies', *Higher Education*. Available at: https://doi.org/10.1007/s10734-020-00557-8.

Cohen, J. (1988) 'Statistical power analysis for the behavioural sciences. Hillside', *NJ: Lawrence Earlbaum Associates* [Preprint].

Cohen, L., Manion, L. and Morrison, K. (2018) 'Research Methods in Education. Eighth Edition', *Research Methods in Education* [Preprint].

Copur-Gencturk, Y., Thacker, I. and Quinn, D. (2021) 'K-8 Teachers' Overall and Gender-Specific Beliefs About Mathematical Aptitude', *International Journal of Science and Mathematics Education*, 19(6). Available at: https://doi.org/10.1007/s10763-020-10104-7.

Corbetta, P. (2011) *Social Research: Theory, Methods and Techniques*, *Social Research: Theory, Methods and Techniques*. Available at: https://doi.org/10.4135/9781849209922.

Costa, A. and Faria, L. (2018) 'Implicit theories of intelligence and academic achievement: A meta-analytic review', *Frontiers in Psychology*. Available at: https://doi.org/10.3389/fpsyg.2018.00829.

Creswell, J. (2014) 'Second Edition Qualitative Inquiry & Research Design Choosing Among Five Approaches', *Public Administration*.

Creswell, J.W. and David Creswell, J. (2018) 'Research Design: Qualitative, Quantitative, and Mixed Methods Approaches', *Journal of Chemical Information and Modeling*, 53.

Creswell, J.W. and Plano Clark, V.L. (2018) *Designing and Conducting Mixed Methods Research*, *Organizational Research Methods*.

Crowne, D.P. and Marlowe, D. (1960) 'A new scale of social desirability independent of psychopathology', *Journal of Consulting Psychology*, 24(4). Available at: https://doi.org/10.1037/h0047358.

Cruickshank, J. (2004) 'A tale of two ontologies: An immanent critique of critical realism', *Sociological Review*. Available at: https://doi.org/10.1111/j.1467-954X.2004.00496.x.

Danermark, B., Ekström, M. and Karlsson, J.C. (2019) *Explaining society: Critical realism in the social sciences*, *Explaining Society: Critical Realism in the Social Sciences*. Available at: https://doi.org/10.4324/9781351017831.

DeCuir-Gunby, J.T. and Schutz, P.A. (2018) *Developing a Mixed Methods Proposal: A Practical Guide for Beginning Researchers*, *Developing a Mixed Methods Proposal: A Practical Guide for Beginning Researchers*. Available at: https://doi.org/10.4135/9781483399980.

Denzin, N.K. *et al.* (2023) *The SAGE handbook of qualitative research*. Sixth edition. Los Angeles: SAGE.

Department for Education (2024) *'National data' from 'Key stage 4 performance'*, *https://explore-education-statistics.service.gov.uk/data-tables/permalink/635bdbf7-66c6-4038-bab9-08dcafcfd634*.

Devine, A. *et al.* (2012) 'Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety', *Behavioral and Brain Functions*, 8. Available at: https://doi.org/10.1186/1744-9081-8-33.

Dillman, D.A., Smyth, J.D. and Christian, L.M. (2016) *Crítica de libros Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method*, *Reis. Rev.Esp.Investig.Sociol. ISSN*.

Doornkamp, L. *et al.* (2022) ‘Understanding gender bias in teachers’ grading: The role of gender stereotypical beliefs’, *Teaching and Teacher Education*, 118. Available at: https://doi.org/10.1016/j.tate.2022.103826.

Dowker, A., Sarkar, A. and Looi, C.Y. (2016) ‘Mathematics anxiety: What have we learned in 60 years?’, *Frontiers in Psychology*. Available at: https://doi.org/10.3389/fpsyg.2016.00508.

Dweck, C.S. (1990) ‘Self-theories and goals: their role in motivation, personality, and development.’, *Nebraska Symposium on Motivation. Nebraska Symposium on Motivation*, 38.

Dweck, C.S. (2006) ‘Mindset: the new psychology of success’, *Choice Reviews Online*, 44(04). Available at: https://doi.org/10.5860/choice.44-2397.

Dweck, C.S. (2015) ‘Carol Dweck Revisits the “Growth Mindset”’, *Education Week*, 35(05).

Eccles, J.S. *et al.* (1983) ‘Expectancies, Values, and Academic Behaviors’, *Achievement and achievement motivation* [Preprint].

Eccles, J.S. and Roeser, R.W. (2009) ‘Schools, Academic Motivation, and Stage-Environment Fit’, in *Handbook of Adolescent Psychology*. Available at: https://doi.org/10.1002/9780470479193.adlpsy001013.

Eccles, J.S. and Wigfield, A. (2002) ‘Motivational beliefs, values, and goals’, *Annual Review of Psychology*, 53. Available at: https://doi.org/10.1146/annurev.psych.53.100901.135153.

Else-Quest, N.M., Hyde, J.S. and Linn, M.C. (2010) ‘Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis’, *Psychological Bulletin*, 136(1). Available at: https://doi.org/10.1037/a0018053.

ESRC (2015) ‘ESRC Framework for Research Ethics’, *Economics and Social Research Council* [Preprint], (January).

Van Eycken, L., Demanet, J. and Van Houtte, M. (2023) ‘Teachers’ efficacy, trust, and students’ features: Internal and external forces affecting teachers’ teachability perceptions’, *Social Psychology of Education* [Preprint]. Available at: https://doi.org/10.1007/s11218-023-09865-0.

Fennema, E. and Sherman, J. (1977) ‘Sex-Related Differences in Mathematics Achievement, Spatial Visualization and Affective Factors’, *American Educational Research Journal*, 14(1). Available at: https://doi.org/10.3102/00028312014001051.

Fennema, E. and Sherman, J.A. (1976) ‘Fennema-Sherman Mathematics Attitudes Scales: Instruments Designed to Measure Attitudes toward the Learning of Mathematics by Females and Males’, *Journal for Research in Mathematics Education*, 7(5). Available at: https://doi.org/10.2307/748467.

Fetters, M.D., Curry, L.A. and Creswell, J.W. (2013) ‘Achieving integration in mixed methods designs - Principles and practices’, *Health Services Research*, 48(6 PART2). Available at: https://doi.org/10.1111/1475-6773.12117.

Fletcher, A.J. (2017) ‘Applying critical realism in qualitative research: methodology meets method’, *International Journal of Social Research Methodology*, 20(2). Available at: https://doi.org/10.1080/13645579.2016.1144401.

Flore, P.C. and Wicherts, J.M. (2015) ‘Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis’, *Journal of School Psychology*, 53(1). Available at: https://doi.org/10.1016/j.jsp.2014.10.002.

Forgasz, H. (2021) ‘Gender: A dilemma for large-scale studies in mathematics education’, *Mathematics Education Research Journal*, 33(4). Available at: https://doi.org/10.1007/s13394-020-00353-8.

Forgasz, H.J. and Leder, G.C. (2010) ‘Equity and Quality of Mathematics Education: Research and Media Portrayals’, in *Mapping Equity and Quality in Mathematics Education*. Available at: https://doi.org/10.1007/978-90-481-9803-0_15.

Forgasz, H.J. and Leder, G.C. (2017) ‘Persistent gender inequities in mathematics achievement and expectations in Australia, Canada and the UK’, *Mathematics Education Research Journal*, 29(3). Available at: https://doi.org/10.1007/s13394-017-0190-x.

Frenzel, A.C. *et al.* (2009) ‘Emotional Transmission in the Classroom: Exploring the Relationship Between Teacher and Student Enjoyment’, *Journal of Educational Psychology*, 101(3). Available at: https://doi.org/10.1037/a0014695.

Frenzel, A.C., Pekrun, R. and 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). Available at: https://doi.org/10.1007/BF03173468.

Gentrup, S. and Rjosk, C. (2018) ‘Pygmalion and the gender gap: do teacher expectations contribute to differences in achievement between boys and girls at the beginning of schooling?’, *Educational Research and Evaluation*, 24(3–5). Available at: https://doi.org/10.1080/13803611.2018.1550840.

Gershenson, S., Holt, S.B. and Papageorge, N.W. (2016) ‘Who believes in me? The effect of student–teacher demographic match on teacher expectations’, *Economics of Education Review*, 52. Available at: https://doi.org/10.1016/j.econedurev.2016.03.002.

Gjøvik, Ø., Kaspersen, E. and Farsani, D. (2023) ‘Stereotypical images of male and female mathematics teachers’, *Research in Mathematics Education*, 25(2). Available at: https://doi.org/10.1080/14794802.2022.2041471.

Glock, S. and Krolak-Schwerdt, S. (2014) ‘Stereotype activation versus application: how teachers process and judge information about students from ethnic minorities and with low socioeconomic background’, *Social Psychology of Education*, 17(4). Available at: https://doi.org/10.1007/s11218-014-9266-6.

Good, C., Aronson, J. and Harder, J.A. (2008) ‘Problems in the pipeline: Stereotype threat and women’s achievement in high-level math courses’, *Journal of Applied Developmental Psychology*, 29(1). Available at: https://doi.org/10.1016/j.appdev.2007.10.004.

Good, T.L. and Lavigne, A.L. (2017) *Looking in Classrooms*, *Looking in Classrooms*. Available at: https://doi.org/10.4324/9781315627519.

Graham, A., Powell, M.A. and Taylor, N. (2015) ‘Ethical Research Involving Children: Encouraging Reflexive Engagement in Research with Children and Young People’, *Children and Society*, 29(5). Available at: https://doi.org/10.1111/chso.12089.

Greene, J.C., Caracelli, V.J. and Graham, W.F. (1989) ‘Toward a Conceptual Framework for Mixed-Method Evaluation Designs’, *Educational Evaluation and Policy Analysis*, 11(3). Available at: https://doi.org/10.3102/01623737011003255.

Greenwald, A.G., McGhee, D.E. and Schwartz, J.L.K. (1998) ‘Measuring individual differences in implicit cognition: The implicit association test’, *Journal of Personality and Social Psychology*, 74(6). Available at: https://doi.org/10.1037/0022-3514.74.6.1464.

Grix, J. (2002) ‘Introducing Students to the Generic Terminology of Social Research’, *Politics*, 22(3). Available at: https://doi.org/10.1111/1467-9256.00173.

Guetterman, T.C., Fetters, M.D. and Creswell, J.W. (2015) ‘Integrating quantitative and qualitative results in health science mixed methods research through joint displays’, *Annals of Family Medicine*, 13(6). Available at: https://doi.org/10.1370/afm.1865.

Hammersley, M. and Traianou, A. (2015) *Ethics in Qualitative Research: Controversies and Contexts*, *Ethics in Qualitative Research: Controversies and Contexts*. Available at: https://doi.org/10.4135/9781473957619.

Hembree, R. (1990) ‘The Nature, Effects, and Relief of Mathematics Anxiety’, *Journal for Research in Mathematics Education*, 21(1). Available at: https://doi.org/10.2307/749455.

Heyder, A., Steinmayr, R. and Kessels, U. (2019) ‘Do Teachers’ Beliefs About Math Aptitude and Brilliance Explain Gender Differences in Children’s Math Ability Self-Concept?’, *Frontiers in Education*, 4. Available at: https://doi.org/10.3389/feduc.2019.00034.

Heyder, A., Weidinger, A.F. and Steinmayr, R. (2021) ‘Only a Burden for Females in Math? Gender and Domain Differences in the Relation Between Adolescents’ Fixed Mindsets and Motivation’, *Journal of Youth and Adolescence*, 50(1). Available at: https://doi.org/10.1007/s10964-020-01345-4.

Hill, F. *et al.* (2016) ‘Maths anxiety in primary and secondary school students: Gender differences, developmental changes and anxiety specificity’, *Learning and Individual Differences*, 48. Available at: https://doi.org/10.1016/j.lindif.2016.02.006.

Hoddy, E.T. (2019) ‘Critical realism in empirical research: employing techniques from grounded theory methodology’, *International Journal of Social Research Methodology*, 22(1). Available at: https://doi.org/10.1080/13645579.2018.1503400.

Hodis, F.A. (2018) ‘Underpinnings of expectancies of success in mathematics: An analysis of general, school-related, and domain-specific motivation antecedents’, *Journal of Educational Psychology*, 110(3). Available at: https://doi.org/10.1037/edu0000218.

Hsieh, T. yang, Simpkins, S.D. and Eccles, J.S. (2021) ‘Gender by racial/ethnic intersectionality in the patterns of Adolescents’ math motivation and their math achievement and engagement’, *Contemporary Educational Psychology*, 66. Available at: https://doi.org/10.1016/j.cedpsych.2021.101974.

Hyde, J.S., Fennema, E. and Lamon, S.J. (1990) ‘Gender differences in mathematics performance: A meta-analysis.’, *Psychological Bulletin*, 107(2). Available at: https://doi.org/10.1037/0033-2909.107.2.139.

Inglis, M. and O’Hagan, S. (2022) ‘Stereotype threat, gender and mathematics attainment: A conceptual replication of Stricker & Ward’, *PLoS ONE*, 17(5 May). Available at: https://doi.org/10.1371/journal.pone.0267699.

Ivankova, N. V., Creswell, J.W. and Stick, S.L. (2006) ‘Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice’, *Field Methods*, 18(1). Available at: https://doi.org/10.1177/1525822X05282260.

Jaremus, F. *et al.* (2020) ‘Girls are still being “counted out”: teacher expectations of high-level mathematics students’, *Educational Studies in Mathematics*, 105(2). Available at: https://doi.org/10.1007/s10649-020-09986-9.

Jaremus, F. (2021) ‘When girls do masculinity like boys do: establishing gender heteroglossia in school mathematics participation’, *Mathematics Education Research Journal*, 33(4). Available at: https://doi.org/10.1007/s13394-020-00355-6.

Johnson, R.B. and Onwuegbuzie, A.J. (2007) ‘Toward a Definition of Mixed Methods Research’, *Journal of Mixed Methods Research*, 1(2). Available at: https://doi.org/10.1177/1558689806298224.

Johnston, O., Wildy, H. and Shand, J. (2019) ‘A decade of teacher expectations research 2008–2018: Historical foundations, new developments, and future pathways’, *Australian Journal of Education*, 63(1). Available at: https://doi.org/10.1177/0004944118824420.

Jussim, L. and Harber, K.D. (2005) ‘Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies’, *Personality and Social Psychology Review*. Available at: https://doi.org/10.1207/s15327957pspr0902_3.

Keller, C. (2001) ‘Effect of teachers’ stereotyping on students’ stereotyping of mathematics as a male domain’, *Journal of Social Psychology*, 141(2). Available at: https://doi.org/10.1080/00224540109600544.

Kempster, S. and Parry, K.W. (2011) ‘Grounded theory and leadership research: A critical realist perspective’, *Leadership Quarterly*, 22(1). Available at: https://doi.org/10.1016/j.leaqua.2010.12.010.

Kivunja, C. and Kuyini, A.B. (2017) ‘Understanding and Applying Research Paradigms in Educational Contexts’, *International Journal of Higher Education*, 6(5). Available at: https://doi.org/10.5430/ijhe.v6n5p26.

Klein, R.B. (2016) ‘Principles and Practice of Structural Equation Modeling, Fourth Edition’, *The Guilford Press*, 8(5).

Koivula, A., Räsänen, P. and Sarpila, O. (2019) ‘Examining Social Desirability Bias in Online and Offline Surveys’, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Available at: https://doi.org/10.1007/978-3-030-22646-6_11.

Krämer, N.C. *et al.* (2016) ‘Closing the gender gap in STEM with friendly male instructors? on the effects of rapport behavior and gender of a virtual agent in an instructional interaction’, *Computers and Education*, 99. Available at: https://doi.org/10.1016/j.compedu.2016.04.002.

Lazarides, R. and Watt, H.M.G. (2015) ‘Girls’ and boys’ perceived mathematics teacher beliefs, classroom learning environments and mathematical career intentions’, *Contemporary Educational Psychology*, 41. Available at: https://doi.org/10.1016/j.cedpsych.2014.11.005.

Leder, G.C. and Forgasz, H.J. (2018) ‘Measuring who counts: gender and mathematics assessment’, *ZDM - Mathematics Education*, 50(4). Available at: https://doi.org/10.1007/s11858-018-0939-z.

Lee, J. *et al.* (2021) ‘Enhancing children’s math motivation with a joint intervention on mindset and gender stereotypes’, *Learning and Instruction*, 73. Available at: https://doi.org/10.1016/j.learninstruc.2020.101416.

Lee, J., Lee, H.J. and Bong, M. (2022) ‘Boosting children’s math self-efficacy by enriching their growth mindsets and gender-fair beliefs’, *Theory into Practice*, 61(1). Available at: https://doi.org/10.1080/00405841.2021.1932156.

Li, Q. (1999) ‘Teachers’ beliefs and gender differences in mathematics: A review’, *Educational Research*. Available at: https://doi.org/10.1080/0013188990410106.

Likert, R. (1932) ‘A technique for the measurement of attitudes’, *Archives of Psychology*, 140.

Lindberg, S.M. *et al.* (2010) ‘New Trends in Gender and Mathematics Performance: A Meta-Analysis’, *Psychological Bulletin*, 136(6). Available at: https://doi.org/10.1037/a0021276.

Marks, G.N. (2008) ‘Accounting for the gender gaps in student performance in reading and mathematics: Evidence from 31 countries’, *Oxford Review of Education*, 34(1). Available at: https://doi.org/10.1080/03054980701565279.

Mason-Bish, H. (2019) ‘The elite delusion: reflexivity, identity and positionality in qualitative research’, *Qualitative Research*, 19(3). Available at: https://doi.org/10.1177/1468794118770078.

Maxwell, J.A. and Mittapalli, K. (2015) ‘Realism as a Stance for Mixed Methods Research’, in *SAGE Handbook of Mixed Methods in Social & Behavioral Research*. Available at: https://doi.org/10.4135/9781506335193.n6.

May, D.K. (2009) ‘Mathematics self-efficacy and anxiety questionnaire’, *Doctoral Dissertation, University of Georgia Georgia* [Preprint].

McAlpine, L. (2012) ‘Identity-Trajectories: Doctoral Journeys from Past to Present to Future’, *Australian Universities’ Review*, 54(1).

McCambridge, J., Witton, J. and Elbourne, D.R. (2014) ‘Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects’, *Journal of Clinical Epidemiology*. Available at: https://doi.org/10.1016/j.jclinepi.2013.08.015.

Mcevoy, P. and Richards, D. (2006) ‘A critical realist rationale for using a combination of quantitative and qualitative methods’, *Journal of Research in Nursing*, 11(1). Available at: https://doi.org/10.1177/1744987106060192.

McKown, C. and Weinstein, R.S. (2008) ‘Teacher expectations, classroom context, and the achievement gap’, *Journal of School Psychology*, 46(3). Available at: https://doi.org/10.1016/j.jsp.2007.05.001.

Mejía-Rodríguez, A.M., Luyten, H. and Meelissen, M.R.M. (2021) ‘Gender Differences in Mathematics Self-concept Across the World: an Exploration of Student and Parent Data of TIMSS 2015’, *International Journal of Science and Mathematics Education*, 19(6). Available at: https://doi.org/10.1007/s10763-020-10100-x.

Miller, C.J., Perera, H.N. and Maghsoudlou, A. (2021) ‘Students’ multidimensional profiles of math engagement: Predictors and outcomes from a self-system motivational perspective’, *British Journal of Educational Psychology*, 91(1). Available at: https://doi.org/10.1111/bjep.12358.

Murayama, K. and Elliot, A.J. (2009) ‘The Joint Influence of Personal Achievement Goals and Classroom Goal Structures on Achievement-Relevant Outcomes’, *Journal of Educational Psychology*, 101(2). Available at: https://doi.org/10.1037/a0014221.

North, E.A. and Ryan, A.M. (2018) ‘The Association of Peer Academic Reputations in Math and Science With Achievement Beliefs and Behaviors During Early Adolescence’, *Journal of Early Adolescence*, 38(6). Available at: https://doi.org/10.1177/0272431617692441.

Nurlu, Ö. (2017) ‘Developing a teachers’ gender stereotype scale toward mathematics’, *International Electronic Journal of Elementary Education*, 10(2). Available at: https://doi.org/10.26822/iejee.2017236124.

Nürnberger, M. *et al.* (2016) ‘Implicit Gender Stereotypes and Essentialist Beliefs Predict Preservice Teachers Tracking Recommendations’, *Journal of Experimental Education*, 84(1). Available at: https://doi.org/10.1080/00220973.2015.1027807.

Pajares, F. (1996) ‘Self-efficacy beliefs in academic settings’, *Review of Educational Research*, 66(4). Available at: https://doi.org/10.3102/00346543066004543.

Pennington, C.R. *et al.* (2016) ‘Twenty years of stereotype threat research: A review of psychological mediators’, *PLoS ONE*, 11(1). Available at: https://doi.org/10.1371/journal.pone.0146487.

Peterson, E.R. *et al.* (2016) ‘Teachers’ explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap’, *Learning and Instruction*, 42. Available at: https://doi.org/10.1016/j.learninstruc.2016.01.010.

Punch, K. (2014) ‘Introduction to Social Research: Quantitative & Qualitative Approaches, 3rd Edition’, *ProtoView*, p. NA. Available at: https://link.gale.com/apps/doc/A734450668/AONE?u=tou&sid=bookmark-AONE&xid=5e2ef69f.

Richardson, F.C. and Suinn, R.M. (1972) ‘The Mathematics Anxiety Rating Scale: Psychometric data’, *Journal of Counseling Psychology*, 19(6). Available at: https://doi.org/10.1037/h0033456.

Rodríguez-Planas, N. and Nollenberger, N. (2018) ‘Let the girls learn! It is not only about math … it’s about gender social norms’, *Economics of Education Review*, 62. Available at: https://doi.org/10.1016/j.econedurev.2017.11.006.

Rosenthal, R. and Jacobson, L. (1968) ‘Pygmalion in the classroom’, *The Urban Review*, 3(1). Available at: https://doi.org/10.1007/BF02322211.

Rubel, L.H. and Bay-Williams, I. by: J.M. (2022) ‘“Speaking Up and Speaking Out about Gender Identity”’, *Mathematics Teacher: Learning and Teaching PK-12*, 115(7). Available at: https://doi.org/10.5951/mtlt.2022.0115.

Rubie-Davies, C. (2014) *Becoming a high expectation teacher: Raising the bar*, *Becoming a High Expectation Teacher: Raising the bar*. Available at: https://doi.org/10.4324/9781315761251.

Rubie-Davies, C., Hattie, J. and Hamilton, R. (2006) ‘Expecting the best for students: Teacher expectations and academic outcomes’, *British Journal of Educational Psychology*, 76(3). Available at: https://doi.org/10.1348/000709905X53589.

Scotland, J. (2012) ‘Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms’, *English Language Teaching*, 5(9). Available at: https://doi.org/10.5539/elt.v5n9p9.

Sheu, H. Bin *et al.* (2018) ‘Sources of self-efficacy and outcome expectations in science, technology, engineering, and mathematics domains: A meta-analysis’,

Journal of Vocational Behavior, 109. Available at: https://doi.org/10.1016/j.jvb.2018.10.003.

Shores, M.L. and Shannon, D.M. (2007) 'The Effects of Self-Regulation, Motivation, Anxiety, and Attributions on Mathematics Achievement for Fifth and Sixth Grade Students', *School Science and Mathematics*, 107(6). Available at: https://doi.org/10.1111/j.1949-8594.2007.tb18284.x.

Sisk, V.F. *et al.* (2018) 'To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses', *Psychological Science*, 29(4). Available at: https://doi.org/10.1177/0956797617739704.

Smith, C. and Elger, T. (2012) *Critical realism and interviewing subjects*, *Studying organizations using critical realism: A practical guide*.

Sorhagen, N.S. (2013) 'Early teacher expectations disproportionately affect poor children's high school performance', *Journal of Educational Psychology*, 105(2). Available at: https://doi.org/10.1037/a0031754.

Spencer, S.J., Steele, C.M. and Quinn, D.M. (1999) 'Stereotype Threat and Women's Math Performance', *Journal of Experimental Social Psychology*, 35(1). Available at: https://doi.org/10.1006/jesp.1998.1373.

Steele, C.M. (1997) 'A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance', *American Psychologist*, 52(6). Available at: https://doi.org/10.1037/0003-066X.52.6.613.

Steele, C.M. and Aronson, J. (1995) 'Stereotype Threat and the Intellectual Test Performance of African Americans', *Journal of Personality and Social Psychology*, 69(5). Available at: https://doi.org/10.1037/0022-3514.69.5.797.

Stutchbury, K. and Fox, A. (2009) 'Ethics in educational research: Introducing a methodological tool for effective ethical analysis', *Cambridge Journal of Education*, 39(4). Available at: https://doi.org/10.1080/03057640903354396.

Suárez-Álvarez, J., Fernández-Alonso, R. and Muñiz, J. (2014) 'Self-concept, motivation, expectations, and socioeconomic level as predictors of academic performance in mathematics', *Learning and Individual Differences*, 30. Available at: https://doi.org/10.1016/j.lindif.2013.10.019.

Tapia, M. and Marsh, G.E. (2004) 'An Instrument to Measure Mathematics Attitudes', *Academic Exchange Quarterly*, 8(2).

Teddlie, C. and Tashakkori, A. (2009) *Foundations of Mixed methods research: Integrating Quantitative and Qualitative approaches in the Social and Behavioural Sciences*, *Sage Publications, Inc*.

Tiedemann, J. (2000) 'Gender-related beliefs of teachers in Elementary School Mathematics', *Educational Studies in Mathematics*, 41(2). Available at: https://doi.org/10.1023/A:1003953801526.

Timmermans, A.C., de Boer, H. and van der Werf, M.P.C. (2016) ‘An investigation of the relationship between teachers’ expectations and teachers’ perceptions of student attributes’, *Social Psychology of Education*, 19(2). Available at: https://doi.org/10.1007/s11218-015-9326-6.

Timmermans, A.C. and Rubie-Davies, C.M. (2023) ‘Gender and minority background as moderators of teacher expectation effects on self-concept, subjective task values, and academic performance’, *European Journal of Psychology of Education*, 38(4). Available at: https://doi.org/10.1007/s10212-022-00650-9.

Tourangeau, R. (2020) ‘Psychology of Survey Response’, in *Polling America: An Encyclopedia of Public Opinion, Second Edition: Volumes 1-2*.

UK Government (2018) *UK Data Protection Act 2018*, *https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted*.

Upadyaya, K. and Eccles, J.S. (2014) ‘How do teachers’ beliefs predict children’s interest in math from kindergarten to sixth grade?’, *Merrill-Palmer Quarterly*, 60(4). Available at: https://doi.org/10.13110/merrpalmquar1982.60.4.0403.

Wang, Z. *et al.* (2020) ‘The longitudinal role of mathematics anxiety in mathematics development: Issues of gender differences and domain-specificity’, *Journal of Adolescence*, 80. Available at: https://doi.org/10.1016/j.adolescence.2020.03.003.

White, P. (2009) *Developing research questions : a guide for social scientists*. Palgrave Macmillan.

Wigfield, A. and Cambria, J. (2010) ‘Students’ achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes’, *Developmental Review*. Available at: https://doi.org/10.1016/j.dr.2009.12.001.

Wigfield, A. and Eccles, J.S. (2000) ‘Expectancy-value theory of achievement motivation’, *Contemporary Educational Psychology*, 25(1). Available at: https://doi.org/10.1006/ceps.1999.1015.

Wolters, C.A. and Daugherty, S.G. (2007) ‘Goal structures and teachers’ sense of efficacy: Their relation and association to teaching experience and academic level’, *Journal of Educational Psychology*, 99(1). Available at: https://doi.org/10.1037/0022-0663.99.1.181.

Yeager, D.S. and Dweck, C.S. (2012) ‘Mindsets That Promote Resilience: When Students Believe That Personal Characteristics Can Be Developed’, *Educational Psychologist*. Available at: https://doi.org/10.1080/00461520.2012.722805.

Yin, R.K. (2013) ‘Case study research: Design and methods.’, *Applied Social Research Methods Series*, 18(2). Available at: https://doi.org/10.1097/00001610-199503000-00004.

Zachariadis, M., Scott, S. and Barrett, M. (2013) ‘Methodological implications of critical realism for mixed-methods research’, *MIS Quarterly: Management Information Systems*, 37(3). Available at: https://doi.org/10.25300/misq/2013/37.3.09.

Zimmerman, B.J. (2000) ‘Self-Efficacy: An Essential Motive to Learn’, *Contemporary Educational Psychology*, 25(1). Available at: https://doi.org/10.1006/ceps.1999.1016.

Zimmerman, B.J. and Martinez-Pons, M. (1990) ‘Student Differences in Self-Regulated Learning: Relating Grade, Sex, and Giftedness to Self-Efficacy and Strategy Use’, *Journal of Educational Psychology*, 82(1). Available at: https://doi.org/10.1037/0022-0663.82.1.51.

Appendices

Appendix A: Student Survey

Demographic questions						
Gender	Age	Ethnicity			Recent maths grade	
Sample questions adapted from Mathematics Self-Efficacy and Anxiety Questionnaire (MSEAQ) (May, 2009) and the Mathematics Anxiety Rating Scale-Short (MARS-Short) (Richardson and Suinn, 1972). 1 = strongly disagree to 5 = strongly agree						
I believe I can solve mathematics problems without too much difficulty		**1**	**2**	**3**	**4**	**5**
I feel nervous when using mathematics outside of school		**1**	**2**	**3**	**4**	**5**
I feel stressed when listening to mathematics teachers explain how to solve a problem		**1**	**2**	**3**	**4**	**5**
I get tense when I prepare for a mathematics test		**1**	**2**	**3**	**4**	**5**
I believe I am the kind of person who is good at mathematic		**1**	**2**	**3**	**4**	**5**
Sample questions adapted from Attitudes Toward Mathematics Inventory (ATMI) (Tapia and Marsh, 2004): 1 = strongly disagree to 5 = strongly agree						
Self-confidence	I am able to solve mathematics problems without too much difficulty	**1**	**2**	**3**	**4**	**5**
	Studying mathematics makes me feel nervous	**1**	**2**	**3**	**4**	**5**
Value	Mathematics is important in everyday life	**1**	**2**	**3**	**4**	**5**
	Mathematics is one of the most important subjects for people to study	**1**	**2**	**3**	**4**	**5**
Enjoyment	I have usually enjoyed studying mathematics in school	**1**	**2**	**3**	**4**	**5**
	Mathematics is dull and boring	**1**	**2**	**3**	**4**	**5**
Motivation	I am willing to take more than the required amount of mathematics	**1**	**2**	**3**	**4**	**5**
	I complete my tasks even if they are very hard or take me a long time	**1**	**2**	**3**	**4**	**5**
Sample questions adapted from the Fennema-Sherman Mathematics Attitudes Scale-Short Form (FSMAS-SF) (Fennema and Sherman, 1976) 1 = strongly disagree to 5 = strongly agree						

Males are naturally better at mathematics than females	**1**	**2**	**3**	**4**	**5**
Mathematics is more important for boys than for girls	**1**	**2**	**3**	**4**	**5**
Girls who enjoy studying mathematics are a bit weird	**1**	**2**	**3**	**4**	**5**
Females are as good as males in mathematics	**1**	**2**	**3**	**4**	**5**
Sample questions designed for probing educational and classroom expectations					
My maths teacher(s) pays more attention to the boys in my class	**1**	**2**	**3**	**4**	**5**
How far do you expect to go in your formal education?	College/ Sixth Form		University		Postgraduate
Do you plan to pursue a career that involves mathematics?	Yes			No	
If yes, what career are you considering? (Open-ended)					
Sample open-ended questions designed to gain insights					
Do you think there are differences in how boys and girls perform in mathematics? Why or why not?					
Have you ever felt that your gender influenced how others perceived your mathematical abilities? If so, how?					
How do you think your gender impacts your experiences in mathematics classes?					

Appendix B: Teacher Survey

Demographic questions						
Gender	Age	Years teaching experience	Subject(s) taught		Age range taught	
Sample questions adapted from the Teachers' Gender Stereotype Scale toward Mathematics (Nurlu, 2017). 1 = strongly disagree to 5 = strongly agree						
Environment	Boys/Girls are more likely to receive support from their families for pursuing mathematics.	**1**	**2**	**3**	**4**	**5**
	Boys/Girls are more likely to have role models in mathematical fields.	**1**	**2**	**3**	**4**	**5**
Career	Boys/Girls are more likely to choose careers that require strong mathematical skills.	**1**	**2**	**3**	**4**	**5**
	Mathematics is more important for Boys'/Girls' future careers.	**1**	**2**	**3**	**4**	**5**
Competence	Boys/Girls generally have higher natural abilities in mathematics.	**1**	**2**	**3**	**4**	**5**
	Boys/Girls are more competent in solving complex mathematical problems	**1**	**2**	**3**	**4**	**5**
Attribution	When Boys/Girls succeed in mathematics, it's usually due to hard work rather than innate ability.	**1**	**2**	**3**	**4**	**5**
	Boys'/Girls' success in mathematics is more often attributed to luck.	**1**	**2**	**3**	**4**	**5**
Sample questions that probe into teachers' expectations, adapted from Rubie-Davies, Hattie and Hamilton (2006). 1 = strongly disagree to 5 = strongly agree						
I expect male students to perform well in advanced mathematics courses.		**1**	**2**	**3**	**4**	**5**
I expect female students to perform well in advanced mathematics courses.		**1**	**2**	**3**	**4**	**5**
I believe male students are likely to pursue STEM careers in the future.		**1**	**2**	**3**	**4**	**5**
I believe female students are likely to pursue STEM careers in the future.		**1**	**2**	**3**	**4**	**5**
Sample questions that probe into teachers' awareness of gender diversity in mathematics education, adapted from Copur-Gencturk, Thacker and Quinn (2021). 1 = strongly disagree to 5 = strongly agree						

I am aware of potential biases in how mathematics is taught to different genders.	**1**	**2**	**3**	**4**	**5**
I actively work to create an inclusive environment for all genders in my mathematics classroom.	**1**	**2**	**3**	**4**	**5**
I believe gender stereotypes can influence students' mathematics performance.	**1**	**2**	**3**	**4**	**5**
I am conscious of how my own gender biases might affect my teaching practices.	**1**	**2**	**3**	**4**	**5**
Sample open-ended questions designed to gain insights					
Describe any differences you've observed in how male and female students approach mathematics.					
What strategies do you use to promote gender equality in your mathematics classroom?					
How do you address gender stereotypes in mathematics when they arise in your classroom?					
In your experience, how do societal expectations about gender roles influence students' attitudes towards mathematics?					

Appendix C: Student Interview Guide

Initial explanation of the study's purpose and informed concent process.

Sample Questions

General perceptions of gender in mathematics	Participant led
Self-efficacy and maths anxiety (May, 2009)	"How confident do you feel about solving mathematical problems? Can you give an example of a recent problem that challenged you?" "Do you ever feel nervous or anxious about maths? If so, can you describe when these feelings typically occur?"
Stereotype threat and belonging (Steele and Aronson, 1995; Nurlu, 2017)	"Have you ever felt that others might judge your maths abilities based on your gender? Can you describe that experience?" "Do you feel like you belong in your maths classes? Why or why not?"
Intersectionality (Becker and Hall, 2024)	"Besides being a boy/girl, are there other parts of who you are that you think affect your maths experiences? This could be things like your family background, where you live, or your cultural heritage."
Future aspirations	"Do you consider maths to be important for your future career aspirations? When do you think you will continue mathematics until?"
Support systems	"How do you find support for maths? Who would you go to? Would you be comfortable asking your teachers or parents?"

Conclusion involving thanks and any final comments

Appendix D: Teacher Interview Guide

Initial explanation of the study's purpose and informed concent process.

Sample Questions

General perceptions of gender in mathematics	Participant led.
Teachers' beliefs about gender differences in mathematics (Nurlu, 2017)	"In your experience, have you noticed any differences in how boys and girls approach mathematical problem-solving? Can you provide specific examples?" "Some research suggests that teachers may have different expectations for boys and girls in mathematics. What are your thoughts on this? Have you observed this in your own teaching or among colleagues?"
Classroom experiences and practices	"Do you use any gender-specific approaches in the classroom?"
Student experiences with stereotype threat and mathematics anxiety (Bandura, 1986; Steele and Aronson, 1995)	"Some students report feeling nervous or anxious about maths. Have you observed this in your classroom? How do you address these feelings when you notice them?" "Research suggests that some students might feel they don't belong in advanced maths classes due to their gender or background. Have you encountered this sentiment among your students? How do you approach this issue?"
Intersectionality in mathematics education (Becker and Hall, 2024)	"How do you think a student's gender might interact with other aspects of their identity, such as their socioeconomic background or ethnicity, in shaping their mathematics experiences?" "Have you observed any patterns in mathematics achievement or participation that seem to be influenced by multiple factors like gender, family background, and cultural expectations? Can you describe these patterns?"
Reflection on quantitative findings	Based on findings in the first phase
Strategies for promoting gender equity	"What are your current strategies for promoting gender equity in your classroom?" "What are the barriers to forming a truly equitable mathematics classroom?"

Conclusion involving thanks and any final comments

Appendix E: Observation Protocol

Number of males:	Number of females:

Tally of interactions with teacher:

	Negative Interaction	Neutral Interaction	Positive Interaction
Male			
Female			

Comment on areas:

Implementation of growth mindset	
Frequency of questions answered by gender	
Task completion rates by gender	
Other comments on gender inclusive practice	

Appendix F: E822 Ethical Appraisal Form

Masters: Education, Childhood and Youth

NB: it should be noted that The Open University is unable to offer liability insurance to cover any negative consequences students might encounter when undertaking 'in-person' data collection. It is therefore very important that you follow appropriate research protocols which should include seeking Gatekeeper permissions to undertake any data collection within your setting and adhering to ethical principles for the safety of yourself and your participants.

Because ethical appraisal should precede data collection, a completed version of this form should be included with TMA02 for those developing a Small-Scale Investigation (SSI) and as part of the EMA submission for those completing an Extended Literature Review and Research Proposal (EP) form of the Dissertation.

Fill in section 1 of this document with your personal details and brief information about your research.

For section 2, please assess your research using the following questions and click yes or no as appropriate. If there is any possibility of significant risk please tick yes. Even if your list contains all "no" you should still return your completed checklist so your tutor/supervisor can assess the proposed research.

Section 1: Project details

a.	Student name	William John Tongue	
b.	PI	[redacted]	
c.	Project title	Bridging the Mathematics Gender Divide: A Mixed Methods Study of Teacher Biases and Their Impact on Female Students	
d.	Supervisor/tutor	Dr Ray Chatwin	
e.	Qualification	Masters in Education	X
		Masters in Childhood and Youth	
f.	MA pathway (where applicable)		

g.	Intended start date for fieldwork	01/09/2024
h.	Intended end date for fieldwork	31/08/2025
i.	Country fieldwork will be conducted in *If you are resident in the UK and will be conducting your research abroad please check www.fco.gov.uk for advice on travel.*	United Kingdom

Section 2: Ethics Assessment		Yes	No
1	Does your proposed research need initial clearance from a 'gatekeeper' (e.g. Local Authority, head teacher, college head, nursery/playgroup manager)?	X	
2	Have you checked whether the organisation requires you to undertake a 'police check' or appropriate level of 'disclosure' before carrying out your research?[1]	X	
3	Have you indicated how informed consent will be obtained from your participants (including children less than 16 years old, school pupils and immediate family members)? Your consent letters/forms must inform participants that they have the right to withdraw from the study at any time.[2]	X	
4	Will your proposed research design mean that it will be necessary for participants to take part in the study without their knowledge/consent at the time (e.g. covert observation of people in non-public places)? If so, have you specified appropriate debriefing procedures? [3]		x

[1] You must agree to comply with any ethical codes of practice or legal requirements that maybe in place within the organisation or country (e.g. educational institution, social care setting or other workplace) in which your research will take place. If required an appropriate level of disclosure ('police check') can obtained from the Disclosure and Barring Service (England and Wales), Disclosure Scotland, AccessNI (Northern Ireland), Criminal Records Office (Republic of Ireland), etc.

[2] This should normally involve the use of an information sheet about the research and what participation will involve, and a signed consent form. You must allow sufficient time for potential participants to consider their decision between the giving of the information sheet and the gaining of consent. No research should be conducted without the opt-in informed consent of participants or their caregivers. In the case of children (individuals under 16 years of age) no research should be conducted without a specified means of gaining their informed consent (or, in the case of young children, their assent) and the consent of their parents, caregivers, or guardians. This is particularly important if your project involves participants who are particularly vulnerable or unable to give informed consent (e.g. children under 16 years, people with learning disabilities, or emotional problems, people with difficulty in understanding or communication, people with identified health problems). There is additional guidance on informed consent on the Masters: Education and Childhood and Youth website under Project Resources.

[3] Where an essential element of the research design would be compromised by full disclosure to participants, the withholding of information should be specified in the project proposal and explicit procedures stated to obviate any potential harm arising from such withholding. Deception or covert collection of data should only take place where it has been agreed with a named responsible person in the organisation and it is essential to achieve the research results required, where the research objective has strong scientific merit and where there is an appropriate risk management and harm alleviation strategy.

5	Does your proposed design involve repetitive observation of participants, (i.e. more than twice over a period of more than 2-3 weeks)? Is this necessary? If it is, have you made appropriate provision for participants to renew consent or withdraw from the study half-way through? [4]		X
6	Are you proposing to collect video and/or audio data? If so, have you indicated how you will protect participants' anonymity and confidentiality and how you will store the data?		X
7	Does your proposal indicate how you will give your participants the opportunity to access the outcomes of your research (including audio/visual materials) after they have provided data?	X	
8	Have you built in time for a pilot study to make sure that any task materials you propose to use are age appropriate and that they are unlikely to cause offence to any of your participants?	X	
9	Is your research likely to involve discussion of sensitive topics (e.g. adult/child relationships, peer relationships, discussions about personal teaching styles, ability levels of individual children and/or adults)? What safeguards have you put in place to protect participants' confidentiality?	X	
10	Does your proposed research raise any issues of personal safety for yourself or other persons involved in the project? Do you need to carry out a 'risk analysis' and/or discuss this with teachers, parents and other adults involved in the research?		X
11	Will financial inducements (other than reasonable expenses and compensation for time) be offered to participants?		X
12	Will the study involve recruitment of patients or staff through the NHS or the use of NHS data?		X

If you answered 'yes' to questions **12**, you will also have to submit an application to an appropriate National Research Ethics Service ethics committee (https://www.hra.nhs.uk/about-us/committees-and-services/res-and-recs/).

[4] Where participants are involved in longer-term data collection, the use of procedures for the renewal of consent at appropriate times should be considered.