Understanding the Psychological and Social Origins of Gender Disparities in Self-Beliefs, Motivation, and Educational Attainment

Brooke Van Zanden

BA (Hons) Psychology (Western Sydney University)

A thesis submitted in fulfillment of the requirements of the degree of Doctor of Philosophy

Institute for Positive Psychology and Education

Faculty of Health Sciences

Australian Catholic University

16th March, 2018

ACKNOWLEDGEMENTS

I remember when I was 17 I was told that I would have to repeat Year 12 because there was no way that someone like me would be able to finish high school successfully. As fate would have it, I've somehow finished a PhD. It's without question my thesis would not have been possible without the help and support of the people who have surrounded me these last few years.

To my supervisors Herb Marsh, Phil Parker, and Marjorie Seaton for pushing me further than I would be able to go myself; and for providing help, feedback and support that has helped me to improve so much throughout the course of my PhD. Without you this thesis would not have been possible!

To my PhD gang who have made doing a doctorate fun! I'm so lucky to have done a PhD surrounded by such a bunch of intelligent, funny, and supportive friends (even if it did mean I was the subject of 24/7 roasts) haha! I'd especially like to thank Michael Noetel and Katherine Owen for assisting me as screeners on my meta-analysis, and also Jasper Duineveld and James Conigrave for being my lifesavers with R. To Brit, my Honours Life Team, and friends who have supported me throughout the entire PhD journey and my psychology studies. Some of the years during my PhD were pretty tough for me. I know it sounds cliché but I honestly don't think I would have gotten through the PhD without my friends at my side. You are my family and my life team and I love you!

To Luke for always believing in me, even in my darkest of times. You kept me (relatively) sane throughout this degree…haha! I feel so grateful that I was lucky enough to have your support and encouragement throughout this journey. Thank you for being such an awesome role model to me, for being such a positive influence in my life, and for helping me realise that I am a Gryffindor! You have inspired me to continue in psychology to work towards improving the lives of other young people, to ensure that every young person feels as though they have the support they need to make their way through adolescence and adulthood, whatever disadvantage they may face.

To Mat, my best friend and partner. You are legit the best human I have ever met. I struggle to think of anyone who is as kind, supportive, encouraging, intelligent and accepting as you. You also happen to be the funniest person I know, AND you are a feminist who loves psychology!? Pretty sure I couldn't have imagined a better boyfriend even if I tried. My world is infinitely better with you in it. Although finishing this PhD is pretty exciting, it's nothing compared to the happiness and joy you give me everyday from being part of my life. I can't wait to go and be a psychologist with you. Bring on the next (thesis-free) chapter of our lives together!

STATEMENT OF AUTHORSHIP AND SOURCES

This thesis contains no material published elsewhere or extracted in whole or in part from a thesis by which I have qualified for or been awarded another degree or diploma.

No parts of this thesis have been submitted towards the award of any other degree or diploma in any other tertiary institution.

No other person's work has been used without due acknowledgment in the main text of the thesis.

All research procedures reported in the thesis received the approval of the relevant Ethics/Safety Committees (where required).

Brooke Van Zanden

Table of Contents

List of Figures

List of Tables

ABSTRACT

The second wave feminist movement during the mid to late 20th century saw rapid advancements in contraceptive, legal, and economic rights of women. However, despite recent advances in women's liberation, gender differences in educational outcomes (e.g., self-beliefs, attitudes, aspirations and educational attainment in science, technology, engineering and mathematics [STEM]) remain heavily entrenched. This thesis explores the notion that the gender gap in STEM has its origins in self-beliefs and task values of young people. However, less is known about how much of the STEM gender gap can be explained by these Expectancy Value Theory constructs. Moreover, there is a lack of research that utilises an intersectional lens to explore how social and cultural context moderates the size of gender gaps in self-beliefs and attitudes towards math and science. This thesis addresses these research gaps utilising meta-analytic, longitudinal, and interview data. Meta-analytic findings from 176 studies in Study 1 show that gender differences in expectancy value constructs are domain specific, and that there are significant moderation effects across social class, gender equality, and gender segregation in university enrollments. Study 2 explores the replicability of meta-analysis results from Study 1, and extends upon these results through an analysis of a large nationally representative database (n = 10,370) that includes ethnicity, geography, and educational attainment. Using the same database, results from Study 3 show that while EVT can account for some of the gender disparity in STEM enrollment, there is still a very large amount of difference that remains unexplained by current theory. Furthermore, results indicate that even when comparing male and female students of equal ability and attitudes, young women still are significantly disadvantaged in terms of STEM university enrollment. A content analysis in Study 4 explored whether open-ended interview data from young Australians who enrolled in a university STEM course (n = 447) versus those who chose to discontinue their STEM education after senior high school (n = 949) can add to current theory. Results point the role of dimensional comparison as critical to educational choices, but again, there were no major themes that arose that significantly deviate from current theory. Results are discussed in light of future directions for research, and implications for policymakers and educators.

INTRODUCTION

> "Life chances and opportunities remain circumscribed by gender, ethnicity, social origin, institutional structures, and the social and economic resources inherent in the connections young people have to their families and wider social context"
>
> (Schoon & Eccles, 2014, p. 6)

The second wave feminist movement during the mid to late 20th century saw rapid advancements in contraceptive, legal, and economic rights of women. However, despite recent advances in women's liberation, gender differences in educational outcomes (e.g., self-beliefs, attitudes, aspirations and educational attainment in science, technology, engineering and mathematics [STEM]) remain heavily entrenched. These gender differences in educational outcomes are crucial to the socioeconomic development of Australia and other countries, as post-industrial economies become increasingly reliant upon innovation and development from the STEM sector. As 50% of the population, women make up a significant portion of potential STEM workers; however, despite the fact that women have similar levels of achievement to their male counterparts, female students are disengaging from STEM (e.g., Watt & Eccles, 2008).

This thesis will explore the notion that the gender gap in STEM has its origins in childhood and adolescence: where female students develop lower self-confidence and lower task value in comparison to their male peers, even when controlling for ability. These lower expectancies for success, combined with poorer task values, act as a critical pathway for disengagement from STEM tertiary study and subsequent career choices. However, how much of the STEM gender gap can be explained by current theory? Moreover, to what degree do social and cultural contexts moderate the size of gender gaps in self-beliefs and attitudes towards math and science? This thesis will explore these questions through meta-analytic, longitudinal and interview data of young women and men.

Chapter 1 explores the persistent problem of gender segregation in study and work. The issue of vertical segregation (a phenomenon whereby women occupy lower status and lower paid roles compared to men) has attracted much attention in gender equality debates. In contrast, there has been less attention towards the effects of horizontal segregation (a phenomenon where women and men tend to work in separate industries or fields). In this chapter, the current statistics on horizontal gender segregation in education and in the workforce are discussed, with a specific focus on gender disparities in STEM. Finally, I outline the impact of gender segregation on economic prosperity, government policy, alongside other wider social implications arising from such highly gender differentiated workforces.

Chapter 2 discusses the causes of gender segregation in educational and career choices. Outlining the historical contexts of gender research in psychology and education, this chapter will explore the evolution of the field across time, highlighting the controversies of the field whereby sexism and science become deeply entangled. The historical context will then be relayed back to current debates within the field, reviewing current biological perspectives on gender differences and discussing current critiques of gender essentialist perspectives. The thesis will pose the question: if biological evidence is inconclusive, to what extent can psychosocial factors explain gender differentiation across the study and career choices?

Chapter 3 reviews Expectancy Value Theory, a key theory, alongside other advances in the study of gender and educational choices (e.g., Hyde's (2005) gender similarities hypothesis). Finally, current empirical evidence of achievement and attitude differences will be reviewed, concluding that gender differences are largest in self-beliefs and task values – and very small in achievement and performance. Chapter 3 extends upon the literature review by discussing current challenges and difficulties for researchers within the field, and highlighting areas requiring further research. This chapter focuses on two major areas for future research: 1) investigating the role of social and cultural context on expectancies for success and task value; and 2) identifying the degree to which current theory can explain young women and men's choices (i.e., to what extent can expectancy value theory account for choices in an empirical sense)?

Chapter 4 summarises the broad research aims and overarching research questions of the current thesis, outlining the key aims of exploring intersectionality with relation to gender and social/cultural context, and identifying the degree to which expectancy value theory can account for gender disparities in educational attainment. Research aims are discussed within the context of the theses' overarching methodologies including meta-analysis, longitudinal analysis of secondary data, and content analysis. Strengths and limitations of each methodology are discussed, while in-depth descriptions of methods specific to each study are provided in detail in Chapters 5, 6, and 7.

Chapters 5 and 6 present the method, results and discussion of Study 1: “Do Social Class and Cultural Context Affect the Size of Gender Gaps? A Meta Analytic Review of Gender Differences in Academic Self-Beliefs and Attitudes” and Study 2: “The Intersection of Gender, Social Class and Cultural Context: A Replication and Extension of Meta-Analysis Findings.” Results for Study 1 show that gender differences are highly domain specific, and that there are significant moderation effects across social class, gender equality, and gender segregation in university enrolments. Study 2 explores the replicability of meta-analysis results from Study 1, and extends upon these results through an analysis of a large nationally representative database that includes ethnicity, geography, and educational attainment. Results from Study 1 and 2 are discussed in light of an intersectional research agenda.

Chapter 7 presents the methods, results, and discussion of Studies 3 and 4: “Young Women Face Disadvantage to Enrolment in STEM Courses Regardless of Prior Achievement, Self-Beliefs and Attitudes”. Results from Study 3 show that while Expectancy Value Theory can account for some of the gender disparity in STEM enrolment, there is still a very large amount of difference that remains unexplained by current theory. Furthermore, results indicate that even when comparing male and female students of equal ability and attitudes, females still are significantly disadvantaged in terms of STEM university enrolment. A content analysis in Study 4 explores whether open-ended interview data can add to current theory. Results point to the role of dimensional comparison as critical to educational choices, but again, there were no major themes that arose that significantly deviate from current theory.

Chapter 8 synthesises the findings of studies 1-4, and discusses results, strengths, and limitations of the research in this thesis. Results are discussed within the context of previous theory, and future directions and challenges for researchers within the field are highlighted. Recommendations for policy makers and educators are made within the context of these findings.

CHAPTER 1
GENDER SEGREGATION IN EDUCATIONAL AND CAREER OUTCOMES: A CHALLENGE FOR GENDER EQUALITY IN THE 21st CENTURY?

In her best-selling book, Susan Pinker (2008) laments what she calls the *Sexual Paradox*; that is, despite Western women's access to contraception, economic freedom, and legislation promising equal opportunity – the working lives of women and men still follow vastly different paths. Pinker notes that men still seem to be drawn towards high status careers (e.g., science, business, law, technology); they are more likely to succeed in positions of leadership, and still seem to outperform women at the highest of levels. Despite gains from the feminist movement women *still* choose to embrace careers that involve nurturing and caring for others; while men continue to gravitate towards careers capitalising on skills of creation, innovation and leadership. This leads Pinker to pose the question – is this necessarily a problem that should be fixed? Is it time for a departure from traditional notions of gender equality, instead of embracing deeply ingrained gender differences in aptitude and interests?

This thesis uses Expectancy Value Theory (EVT) to explore whether gender differences in occupational choices can be traced back to childhood and adolescence. According to EVT (discussed in further detail in Chapter 3), young women disengage from STEM careers largely because of their comparatively lower levels of expectancy for success and task values, which in turn influence coursework choices at university and subsequent employment options. Importantly, these gender differences in self-beliefs and attitudes form early on, during childhood and adolescence. Thus, although the focus of this thesis is primarily on gender differences in educational outcomes, it is important to review the longer-term social and economic consequences of gender differences in educational outcomes because these differences are the critical precursors of gender segregation in the labour market.

Thus, this chapter will review the extant literature and statistics on gender differences in current educational and career outcomes to highlight the social and economic ramifications of gender roles and gender differences in self-beliefs and attitudes. Are gender gaps across different fields as large and stable as Pinker suggests? How has the education and labour market changed across time for men and women? Are trends for gender differentiation in the labour market universal, or, is there evidence to suggest that culture plays a critical role? And finally, are there negative social or economic implications of gender differentiated educational and occupational pathways, *or*, are gendered educational and career paths something we should embrace, as suggested by Pinker?

Out of the Kitchen, Into an Occupational Ghetto: Trends in Educational and Occupational Choices for Women and Men

The second wave feminist movement during the mid to late 20th century saw rapid advancements in women's rights. One of the most critical changes that occurred during the second wave feminist movement were the changes to the economic, legal and social fabric of society that resulted in greater representation of women in the workforce and greater representation of women studying at the higher education level. For example, the gender gap in university attendance and graduation has dramatically decreased to the point where the gaps in educational attainment and academic motivation have reversed in favour of women (Arnot, 2002; DiPrete & Buchmann, 2013). Furthermore, there have been steady and consistent gains in the rates of female labour force participation, showing dramatic changes in the composition of the workforce across the last half century (e.g., Brewster & Rindfuss, 2000). While these changes are impressive, there have been other aspects of gender equality in the workforce that have been harder to shift. In particular, women and men continue to work in starkly different settings, occupations, and levels of leadership (Charles & Grusky, 2004), and are drawn to radically different fields of study and training (Watt & Eccles, 2008). Despite all the legal, social, and economic changes of the feminist movement and the post-industrialisation of modern economies, the lives of women and men are still fundamentally shaped by gender.

There are two different types of gender segregation (i.e., the tendency for women and men to work in different occupations and to follow different educational pathways). Early work on gender segregation focused on the way the workforce was divided by gender (e.g., Hakim, 1979). *Vertical segregation*, highlighted by Hakim, refers to the hierarchy of status jobs; or the trend that sees men more likely than women to hold high status or high-power jobs (e.g., men might be more likely to be doctors, while women are more likely to be nurses). *Horizontal segregation* refers to the trend whereby women and men work in different career areas (e.g., women dominating service industry or caregiving jobs, while men dominate business, legal, and STEM fields; Blackburn, Browne, Brooks & Jarman, 2002). While horizontal segregation gets less media attention and public scrutiny, Blackburn, Brooks, and Jarman (2001) note that horizontal segregation is far more prevalent than vertical segregation. On the surface horizontal segregation may look innocuous, however, horizontal segregation is deeply entwined with the gender pay gap; whereby "women's jobs" (e.g., nursing, support work) are undervalued and underpaid in comparison to male dominated industries (e.g., STEM, law, business; Workplace Gender Equality Agency, 2016). In fact, it is estimated that industry segregation between genders can account at least 19% of the gender pay gap (Klynveld Peat Marwick Goerdeler [KPMG], 2016). Thus, while women might be participating in the workforce, they still remain confined to educational and career pathways that are underpaid and undervalued – the occupational ghettos of society.

Horizontal segregation in Australia and female STEM participation. Australia is highly gender segregated by industry and occupation - a pattern that has persisted over the last few decades despite other advances in gender equality (Workplace Gender Equality Agency, 2016). The problem has become so ingrained that the Australian Government has ordered a senate enquiry into gender segregation in the Australian workplace, as well as the economic consequences of workplaces that are stratified according to gender. In early 2017, the Director of the Workplace Gender Equality Agency remarked:

> "*Gender segregation is actually getting worse….We are not encouraging young women to look to industries other than health care and social assistance, and vice versa, we are not encouraging men to look at the more female dominated industries*"

Overall, women's labour is mainly focused on roles of caregiving and service to others; women hold 80.2% of jobs in health care and social assistance, 63.4% of jobs in education and training, and 58.4% of jobs in retail (Workplace Gender Equality Agency, 2016). In contrast, men dominate manufacturing, trades, and science and technology roles (Workplace Gender Equality Agency, 2016). Gender segregation becomes especially obvious in the area of STEM, particularly in engineering, technology, mathematics, and the physical sciences. For instance, recent surveys of Australian academic and research staff showed that excluding medical science (which traditionally has remained relatively even in gender balance), only 32.3% of STEM academics are female (Science in Australia Gender Equity, 2017). Numbers of female academics and researchers get even lower for math (22.8%), information technology (22.1%), and engineering (18.3%). This gender gap only gets worse in senior positions. For example, at a bachelor degree level, women comprise 33.1% of students studying STEM. This figure increases to 39.7% for women studying STEM at a PhD level. However, drops to 13.9% for women employed as senior STEM academics in Australian universities. Again, this problem is more pronounced for math and the physical/natural sciences (16.3% of senior academics are female), and engineering; where only 8.7% of senior academics are women. Although women face disadvantage in other disciplines in achieving roles of seniority (only 30.1% senior academics in non-STEM disciplines are female), the problem of female representation seems particularly bad in STEM careers.

Importantly, the gender gap in STEM participation does not just exist in the workforce; instead its origins can be traced back to choices made during high school. Research by Mack and Walsh (2013) has shown that there has been a growing gender disparity in STEM participation for New South Wales students across the last decade. For example, the number of girls not taking any math courses in Year 11 and 12 (students between ages 16-18) has doubled to 21.8% in 2011 from just 9.5% in 2001. Furthermore, only 1.5% of girls study advanced math alongside physics and chemistry (courses that are often prerequisites for many university courses including medicine and science degrees). Overall, Mack and Walsh’s (2013) findings show that gender disparities in STEM participation are emerging early and disparities are getting worse, not better; meaning that many young women miss out on developing key mathematical and scientific problem-solving and critical thinking skills as they embark on their journey to adulthood.

Does Cultural Context Shape Gendered Patterns of Study and Work?. Is horizontal gender segregation across career and study pathways a universal trend? The answer is complex. In a review of the literature, Charles and Grusky (2004) note that there are striking commonalities that occur globally for horizontal gender segregation (e.g., worldwide, men are more likely to engage in manual jobs while women are more likely to engage in non-manual jobs). However, there are critical differences in gender segregation across culture, particularly for STEM study and work. For example, Gharibyan and Gunsaulus (2006) report that a number of post-soviet nations have higher rates of female STEM participation than the rates typically observed in western nations; with female participation as high as 75% in some university computer science departments.

Although it seems counterintuitive, there is evidence suggesting that there is more, not less, horizontal gender segregation in post-industrial, rich, “gender equal” nations, as opposed to poorer, developing nations (see Charles and Grusky, 2004 for a brief review). For example, progressive, liberal Scandinavian countries actually have higher levels of gender segregation than countries like Japan, Portugal, Italy and Greece; all countries with far more traditional gender policies than Scandinavia (Anker, 1998; Blackburn, Jarman, & Brooks, 2000; Charles, 1990; 1992; Jacobs & Lim, 1992; Melkas & Anker, 2001; Roos, 1985; Rosenfield & Kalleberg, 1991).

Recent research has shown similar results. Blackburn, Brooks and Jarman (2001) showed that countries with low levels of Gross Domestic Product, such as many countries in Africa, actually have very small levels of gender segregation according to occupations. In contrast, Charles and Bradley (2009) showed that as economic prosperity increases, the numbers of women in engineering, math, and natural science declines; with more women choosing more feminine subjects in the humanities, social sciences and in health. Sikora and Pokropek (2012) found similar findings for adolescent career plans, where the gender gap in science career plans is stronger for students living in advanced, industrialised nations. Overall, patterns in horizontal gender segregation seem to show that as levels of wealth and prosperity rise, so does the tendency for women and men to take separate paths in their educational and career ambitions and outcomes.

What are the Social and Economic Implications of Horizontal Gender Segregation in STEM?

If horizontal gender segregation is associated with higher levels of gender equality and economic prosperity, should we embrace gender differences in occupational choices? And furthermore, what are the consequences of having highly differentiated career pathways for women and men; particularly within the areas of science and technology? In contrast to Pinker’s (2008) welcoming of gender differentiation across education and career pathways, many scholars have voiced their concerns about the continuation of gender segregation trends in STEM amongst western industrialised nations. These concerns typically revolve around three key arguments: 1) women as an untapped resource to meet growing demand in STEM jobs; 2) social justice for women in terms of the educational and career options available for young women; and 3) benefits of diverse study and work environments.

Women as an untapped resource to fix the “STEM Crisis”. Trends of student and workforce participation in science, technology, engineering, and mathematics (STEM) paint a bleak picture for science education: that is, science education in most post-industrial nations is on the decline. Numbers of students enrolling in senior high school science and mathematics courses are decreasing (Kennedy, Lyons, & Quinn, 2014; Mack & Wilson, 2015; Office of the Chief Scientist, 2014), meaning that many young people do not get exposure to STEM at the time they are making major decisions about what career path they will take. Not surprisingly, this trend continues in tertiary education, where enrolments in STEM courses are declining (U.S. Department of Commerce, 2012; Xie & Killewald, 2012). Furthermore, such low numbers of tertiary educated graduates means that there is an increasingly large shortage of suitably qualified workers in the expanding STEM field (Watt & Eccles, 2008; Xie & Killeweld, 2012). These downward trends, accompanied by a lack of qualified science teachers and reports of student disengagement (Office of the Chief Scientist, 2014), have prompted government investment in STEM education reform (e.g., National Science and Technology Council, 2013; National Innovation and Science Agenda, 2017; U.S. Department of Education, 2010).

The decline of science education in post-industrial nations is particularly troubling given that most modern economies are now heavily geared to rely on scientific and technological innovation for economic growth and prosperity (Carnevale, Smith, & Melton, 2011; Office of the Chief Scientist, 2014). Furthermore, STEM has been identified as an area of high future job growth, with the U.S. Bureau of Labor Statistics (2014) predicting that STEM jobs will grow by 13% between 2012 and 2022; a rate comparatively higher than non-STEM areas of employment. Furthermore, STEM workers are likely to be reimbursed at rates far higher than employees in non-STEM jobs. In fact, the U.S Bureau of Labor Statistics found that STEM workers were paid over double the median wage of U.S. workers ($76,000 compared to $35,080). However, despite the fact that STEM areas are fields of high job growth, trends of decreasing science participation show that young people and in particular young women, are leaving the education system ill-equipped to find employment in our increasingly technology-centred society (see Berman & Machin, 2000 for a review of skill biased technological change).

As half the population, women provide a large source of untapped potential that is currently underutilised in the battle to increase participation in STEM education and training. Consequently, there has been a strong push from governments of post-industrial nations to encourage the retention of female students in science education and careers (e.g., National Innovation and Science Agenda, 2017; National Science and Technology Council, 2013; Office of the Chief Scientist, 2014; U.S. Department of Education, 2010).

Social justice perspectives on gender gaps in STEM & the benefits of diverse study and work environments. Aside from economic benefits of STEM, the gender gap in STEM study and work is an important social justice issue. The next chapter of the thesis highlights how there is a lack of evidence to support the narrative that men and women are innately gifted in different areas; instead, it seems increasingly likely that at least a large proportion of decision making might be to do with girls' self-assessments of ability – which are drastically lower than their actual achievement. If boys and girls are equally able at science and math skills, is it fair that girls disengage from further study in STEM largely because of unwarranted low self-confidence in their abilities? Moreover, what are the social and educational consequences for girls who disengage from STEM study and work despite achievement that is on par with their male counterparts?

Some scholars, such as Lucy Sells (1980), have raised concerns about low levels of female participation in math coursework because of the importance of math as a "critical filter" to later opportunities. Sells argued that young women were often deprived of more diverse study and career opportunities because of early disengagement from math in high school or shortly after, thereby prematurely restricting vocational options and limiting access to high status and high-income careers. In essence, young women are deprived of career options that rely on mathematical and scientific knowledge, not because they lack ability, but because they disengage from STEM early on in their educational and vocational training.

Restricting educational and career opportunities can have important ramifications for women's economic wellbeing. For example, early disengagement in

STEM has been discussed as a possible mechanism behind the gender pay gap (Watt & Eccles, 2008). Many STEM jobs attract a high salary, and are highly prestigious; when women disengage from STEM areas of study they can miss out on the opportunity to gain employment in high-paying roles. Gender differences in pay are critical because women are more likely than men to have to support themselves and dependents without assistance (i.e., children and those requiring caregiving), thereby shouldering a greater level of financial responsibility (e.g., Coltrane & Adam, 1998).

Finally, numerous scholars have argued for the benefits of diverse study and work environments (e.g., Apfelbaum, Phillips, & Richeson et al., 2014; Intemann, 2009; Page, 2007). For instance, university faculties that are homogenous in gender and ethnicity may limit worker creativity and satisfaction (Apfelbaum et al., 2014; Page, 2007). Meanwhile, diverse environments in science further social justice but also expand on workplace talent and enhance ‘objectivity’ by allowing for scientific problems to be approached from people with diverse and unique experiences, allowing teams of scientists to see problems from multiple perspectives (Intemann, 2009).

Chapter Summary

In summary, horizontal gender segregation and the resulting gender gap in the STEM workforce has attracted significant government funding and attention, largely because of the substantial economic costs of a decreasing pool of talent of potential STEM workers. The social costs of horizontal gender segregation are important too, as social norms of work patterns mean that young boys’ and girls’ career and study options are constrained according to gendered norms of work. Women, in particular, bear the brunt of this gender segregation as their gender roles conveniently fit into the most underpaid and undervalued work of society. Solving this problem is difficult, and might require a two-pronged approach. In order to achieve fairer and more equitable workplaces, governments need to push for better pay and compensation for workers in the underpaid and undervalued “pink collar” service and care industries. In addition to this, we need to continue to encourage young women to remain engaged in STEM over the course of their education; and to create further research on the reasons *why* so many women opt out of STEM even when they are just as able as their male counterparts. Chapters 2 and 3 of this thesis will aim to answer the latter issue.

CHAPTER 2

A HISTORY OF SCIENCE AND SEXISM: PSYCHOLOGY'S COMPLICATED RELATIONSHIP WITH GENDER

Why do boys and girls choose such radically different educational and career pathways? Why are girls so unlikely to enter STEM fields even when government initiatives flag female participation in STEM education as a critical goal in furthering national growth and economic success? Why has horizontal segregation in educational and career choices remained so persistent despite the many gains in gender equality over the last century?

The answer to the above questions is perhaps one of the most controversial and hotly debated topics in psychology. Biological and evolutionary perspectives often explain the persisting patterns of gender segregation across the sciences and humanities as a result of genetic and hard-wired innate gender differences caused from sexual selection processes (e.g., Buss, 1995; Geary, 1998). Other researchers will argue the importance of gender socialisation, highlighting the differential way in which we treat girls and boys as paramount in understanding gender differences (e.g., Eagly, 1987; Ridgeway, 2009). These answers characterise the two perspectives that have dominated the long and controversial history of gender differences in psychology. As such, this chapter will review these perspectives within the historical context of scientific literature regarding gender differences in psychology, linking the historical context of gender differences to current debates and discourses surrounding research on gender in the present day. Finally, this chapter will discuss the empirical evidence on gender differences in academic achievement, attitudes and choice behaviours, providing a critical review of the current literature.

The Early Years: Sexism and Science

> "Man is more courageous, pugnacious and energetic than woman, and has a more inventive genius" (1897, p. 557).
>
> (Darwin, 1897, p. 557)

The study of gender and psychology has had a notoriously controversial and politically charged history. Early years of psychological research tended to either ignore women's experiences or assess them through a male-centric analysis (see discussion by Hare-Mustin & Marecek, 1990). Furthermore, when research on gender *did* occur, conclusions were often shaped by cultural norms and research results were used to justify sexist beliefs and ideas about women through pseudo-scientific prejudice (Squire, 1989). In other words, throughout history psychologists, scientists and philosophers have used scientific research to justify cultural beliefs that women are emotional and irrational beings primarily designed for reproduction and caregiving; all traits that are at odds with the characteristics of rationality, logic, and intellect needed for success in mathematics and the sciences.

The pseudoscientific history of psychology and gender, and the linking of femininity to emotionality and inferiority, dates back to the ancient civilisations of Egypt and Greece. In 1900BC Ancient Egyptians described depressive and anxiety symptoms as an exclusively female disorder that was caused by the uterus moving around the female body (Cosmacini, 1997; Sigerist, 1951). Treatments involved placing bad smelling substances near the mouth and nostrils, and pleasant smelling items near a woman's vagina to coax the uterus back into its original location (Cosmacini, 1997; Sigerist, 1951). Ancient Greeks held similar beliefs, and were responsible for coining the term 'hysteria' – claiming that mental health difficulties in women were the result of a wandering uterus (Sterpellone, 2002). Hippocrates claimed that women were physiologically cold and wet, and therefore prone to illness particularly if women did not fulfil their reproductive roles as child-bearers. This was in contrast to the male body which was dry, warm, and more resistant to such disorders (Sterpellone, 2002). Conveniently, women's medical advice was to get married, if not already, and focus on providing their husbands with pleasure within the confines of married life (Sigerist, 1951; Sterpellone, 2002). These early origins of the psychology of gender might seem outrageous to a modern day reader; however, Greek philosophers became highly influential in the evolution of modern western thought surrounding gender. Aristotle posited that men and women were related opposites (e.g., straight vs. curved, light vs. darkness, good vs. evil, completion vs. incompletion), and this idea of male and female duality whereby maleness was superior to femaleness became a central theme of Western philosophy and thought for centuries to come (Lips, 2017). Indeed, for many years these notions of superiority versus inferiority remained unchallenged, and there was little interest amongst Western scientists in investigating the psychology of sex differences because assumptions about gender were so firmly entrenched as common knowledge (Lips, 2017).

When gender comparisons became the focus of research in the 19th century, research focused on confirming assumed differences, rather than conducting rigorous scientific tests of the veracity of assumed gender differences. For example, in the 19th century, phrenologist Franz Joseph Gall measured male and female brains in order to demonstrate that female brains had more areas dedicated to "tender" feminine traits (see Shields, 1975 for a review of Gall's work in relation gender). Francis Galton's (1907) research on individual differences concluded that women were mentally inferior to men, work that was later critiqued for its sexism and racism (e.g., Shields, 1975; Gould, 1981; Lewin, 1984). Charles Darwin argued that lower variability in female intellectual ability explained male tendencies for genius and female tendencies towards averageness (1871/2016). Edward Clarke (1873/2006) claimed that women were unsuited to higher education because intellectual labour sapped energy from the ovaries to the brain, resulting in illness and infertility.[1] Gender differences in female versus male brain size also became a justification for the assumption of female mental inferiority (e.g., Romanes, 1887), however, when this conclusion was challenged, researchers shifted focus and started to search for the part of the brain that was larger in men than in women (Shields, 1975). Even when research showed evidence that women read faster and more accurately than men (Romanes, 1887), Romanes concluded this reflected the fact that women were good liars – rather than the obvious conclusion of skill in reading (see Caplan & Caplan, 2016 for a discussion). The

[1] In response to Clarke's (1873/2006) assertion that ovaries sapped energy from the brain, Lewontin (2000) dryly notes that testicles apparently created their own sources of energy.

above studies highlight a critical flaw in science that is often overlooked within positivism, the dominant way of understanding science and psychology. That is, that the social context affects the way scientists understand knowledge; choices in research questions, hypotheses, designs, and analyses; and the way scientists interpret and disseminate knowledge.

Psychology's complicated relationship with gender did not improve throughout the twentieth century. Research at the turn of the century continued to be plagued by pernicious assumptions of female inferiority. For instance, although Freud's work in psychoanalysis paved the way for effective new treatments grounded in talk therapy, and the importance of parent-child relationships in healthy psychological development (see Westen, 1998 for a review of Freud's contributions to psychology), Freud often ascribed women's distress to 'penis envy' and to neurotic personalities without questioning whether the oppressive environment women found themselves in could possibly contribute to psychological distress (Shafer, 1974). The following period between the 1920s-1970s was characterised by comparatively less research on gender than prior decades (Lips, 2017). Notable exceptions included the work of Karen Horney and Clara Thompson who both made important critiques of psychoanalytic theory, stressing the importance of social and cultural factors on women's psychological health and development (for a review of Horney and Thompson's contribution to psychology see Denmark & Fernandez, 1993). However, overwhelmingly psychology continued to be dominated by male researchers, investigating research with a decidedly male gaze (Squire, 1989).

Have We Really Changed? Gender Essentialist Views as a Mirror to the Past

With the significant inroads made by the feminist movement, the now-debunked claims of female inferiority made by 19th century scientists seem almost unbelievable. However, fast-forward to present day conversations amongst evolutionary psychologists and we see a similar pattern emerging to that of the 19th century research on gender. In the *gender essentialist* worldview, women and men are still diametrically opposed in skills and abilities, and these traits are predetermined perhaps even before birth. Evolutionary and biologically oriented psychologists with a gender essentialist perspective search for hormonal, neurochemical, structural, and functional brain differences to account for psychological differences between women and men (e.g., Cahill, 2003; Stevens & Hamann, 2012). These links between biological and gender difference can seem appealing – after all, they could explain the differences in our society that are stratified by gender. However, a closer look at the research in the area reveals evidence that is underwhelming, often inconclusive and riddled with methodological issues (Fine, 2010). Moreover, it is often the case that when biological differences are found, these differences are automatically linked to psychological differences without a clear causal pathway. This can become problematic because the logic between linking prenatal testosterone and STEM career choices is as tenuous as 19th century scientists linking skull size to intelligence; it is perhaps reasonable to speculate a relationship, but to claim for certain there is a causal link between the two is disingenuous and poor scientific practice.

Nonetheless, it is important to talk about biological explanations when reviewing the literature of gender differences across educational and career choices due to the resurgence of popularity in biological explanations after the 1970s' emphasis on gender socialisation, and the rapid advancement of technology enabling new techniques of exploring brain function and structure (Penner, 2008). Biological perspectives of horizontal career segregation generally focus on how evolution (e.g., Geary, 1998), hormones (e.g., Kimura, 1999), and brain structure and function (e.g., Baron-Cohen, 2003) affect mathematical ability when explaining why women and men choose such vastly different career and educational pathways.

Evolutionary explanations. Evolutionary explanations of gendered educational and career pathways often focus on the role of sexual selection in causing gender differentiation across personal traits and characteristics. For instance, Geary's (1998) work from an evolutionary perspective claims that sexual selection drives underrepresentation of women amongst top performers by driving greater variability among males compared to females. In other words, sexual selection resulted in males becoming more responsive to environmental conditions than females. Consequently, there are more men at the top in prosperous environments and more men at the bottom in adverse environments, thus creating greater male variability in mixed conditions of prosperous and adverse environments. This is just one example of evolutionary theory explaining gender differences, but the logic of sexual selection is a key theme across most evolutionary psychology theories.

Evolutionary psychology theory such as that of Geary (1998) has been criticised for its tendency to make bold claims that have no ability to be falsified through the traditional method (Gannon, 2002). That is, it is difficult to falsify Geary's claims as there is no way of experimentally testing sexual selection theory through random assignment of prosperous versus adverse environments. Furthermore, arguments and propositions in evolutionary psychology, like Geary's (1998), often rely on assumptions that have no universal consensus. Results that are inconsistent with original hypotheses are too often explained away in terms such as: being due to current environments being different than ancestral ones; animal data is often dismissed as irrelevant but then used when it supports results in humans; there is resistance to interpreting results with relation to other psychological paradigms (e.g., Geary fails to mention other possible social forces involved in gender differences); and a failure to acknowledge the role of political and ideological forces that might impact their own sub-disciplines (despite being keen to point out political influences in other sub-disciplines of psychology) (Gannon, 2002).

Hormonal explanations. Why are there so few successful female mathematicians? Some researchers (e.g., Baron-Cohen, 2009) argue that exposure to in-utero testosterone is to blame. Indeed, the relationship between hormones and mathematical aptitude has been well researched, ranging from the effects of testosterone in utero (see Brosnan, 2006), to the effects after birth (e.g., puberty, menstruation, contraception, menopause, and seasonal/circadian fluctuations of hormones) (Van Goozen, Cohen-Kettenis, Gooren, Frijda, & Van de Poll, 1994, 1995). Ceci, Williams, and Barnett's (2009) review on biological explanations for women's underrepresentation in STEM fields found mixed evidence for a hormonal link with mathematical ability. For example, Ceci et al. (2009) outlined mixed evidence from special populations with hormonal disorders such as congenital adrenal hyperplasia (CAH) and hypogonadotropic hypogonadism. A number of studies have found a relationship between spatial abilities and androgen levels, whereby girls with CAH who have an excess of androgen show significantly higher spatial scores than control girls, while boys with CAH show significantly lower spatial scores than control boys (e.g., Hines et al., 2003; Hines & Kaufman, 1994; Resnick, Berenbaum, Gottesman, & Bouchard, 1986). However, Ceci et al. (2009) note that there are a number of exceptions to the above (e.g., Caplan, McPherson, & Tobin, 1985; Malouf, Migeon, Carson, Petrucci, & Wisniewski, 2006; Ripa, Johannsen, Mortensen, & Muller, 2003; Schattman & Sherwin, 2007). Making matters even more complicated, Hines et al. (2003) found that differences in spatial ability amongst females do not covary according to the degree of androgen, and Resnick et al. (1986) found that boys with CAH who had high levels of early testosterone performed similarly to control boys.

Hormonal intervention evidence is also mixed (Ceci et al., 2009). Slabbekoorn, van Goozen, Megens, Gooren, and Cohen-Kettenis (1999) demonstrated that androgen therapy for individuals transitioning from female to male led to higher spatial ability compared to their ability prior to androgen therapy. Other studies have revealed a U-shaped association between activational levels of testosterone, mental rotation and mathematics scores (see Hampson & Moffatt, 2005 for a review). Again though, Ceci et al. (2009) note that there are numerous studies that did not show any relationship between hormones and ability (e.g., Christiansen & Knussman, 1987; McKeever, Rich, Deyo, & Conner, 1987; see Hogervorst, Bandelow, & Moffat, 2005 for a review).

Finally, Ceci et al. (2009) evaluated studies that explored the link between prenatal organising hormones (e.g., Baron-Cohen, Lutchmaya, & Knickmeyer, 2004; Knickmeyer & Baron-Coehn, 2006) and postnatal activating hormones (e.g., Moffat et al., 2002; Davison & Susman, 2001). Baron-Cohen et al. (2004) and Knickmeyer and Baron-Cohen (2006) observed an effect of male hormones in fetal and amniotic fluid on later spatial and mathematical ability. Fink, Brookes, Neave, Manning, and Geary (2006) also found a correlation between numeric competency and differences in finger length, a hypothesised marker of prenatal hormonal exposure (e.g., Manning, 2002; Sanders, Bereczkei, Csatho, & Manning, 2005). Again, Ceci et al. (2009) found that this finding has not been replicated (e.g., Finegan et al., 1992) and Puts et al. (2008) found only very small correlations between finger length ratios and spatial ability in their meta-analysis. In research on postnatal activating hormones, Moffat et al. (2002) showed a strong effect of testosterone of visual-spatial tests of ability. In contrast, Davison and Susman (2001) found a relationship between testosterone and spatial ability in boys, but not for girls. Moreover, Thilers, MacDonald, and Herlitz (2006) did not find any relationship between testosterone and spatial cognition at all.

Overall, reviews of the literature (e.g., Ceci et al., 2009; Kimura, 1999) have concluded that there is some evidence to suggest an optimal level of testosterone for spatial performance (falling at the low end of the male range and the high end of the female range). In other words, very high testosterone or very low testosterone is associated with poorer performance, rather than a linear relationship between testosterone and performance. Moreover, physical scientists' finger length ratios (a common measure of testosterone exposure in the womb) are actually closer to typical female ratios than to male ratios of a comparison group of male social scientists (Bronsan, 2006). Nonetheless, Ceci et al. (2009) emphasise the relationship between hormones and STEM success remains mixed as many studies mentioned above failed to find a U-shaped relationship between testosterone and ability.

Structural and functional brain differences. Some have hypothesised that differences in brain development might be caused directly by biological sex (Gur & Gur, 2007; Haier, 2007), however Ceci et al. (2009) warn that the aetiology of this relationship is unclear. For instance, sex differences in head volume, perimeter, and mass have been used to suggest that women are biologically less scientifically inclined (Scheilbinger, 1987). Rushton (1992a, 1992b) found that women had smaller brain sizes than men, even when controlling for body weight and size, leading some to conclude that male greater brain mass is genetic and is responsible for males having better mathematical and spatial abilities (Ankney, 1992). Ceci et al. (2009) caution against such sweeping generalisations, saying that although brain size could result in cognitive differences, researchers have provided no rationale as to why a difference in overall size would lead to deficits in STEM that does not extend to other cognitive domains. Furthermore, Ceci et al. (2009) wryly note that given that women achieve better school grades in math than men, the brain size sex difference would have to specifically account for standardised achievement test scores, but not STEM performance in other contexts. Overall, it seems as though the evidence for Ankney's (1992) brain size STEM deficit speculation is underwhelming at best.

Ceci et al. (2009) also reviewed evidence of brain imaging studies. Haier, Jung, Yeo, Head, and Alikire (2005) found that different brain areas were correlated to IQ scores for men and women; grey and white matter in frontal areas seemed to be more important for women, while grey matter in parietal areas were more important for men in predicting IQ. Furthermore, other studies showed that women and men use different brain structures when performing mental rotation (e.g., Gur et al., 2000; Haier et al., 2004, 2005; Haier et al. 2006; Hugdahl et al. 2006).

Cerebral theories of gender differences in mathematical and scientific aptitude have used brain structural and functional differences, as well as hormonal evidence to claim that these differences lead to sexed brains – a systemising male brain, and an empathising female brain. For instance, Simon Baron-Cohen (2003) is a researcher whose work has attracted considerable attention within the general public. He states that, “the female brain is predominately hard-wired for empathy. The male brain is predominately hard-wired for understanding and building systems.” According to Baron-Cohen, this hard-wiring can account for why men are ‘better’ scientists and engineers, while women are ‘better’ suited in professions that reflect traditional caregiving roles; counsellors, primary school teachers, nurses, carers, social workers, and therapists.

What evidence is there to support the notion of a ‘male’ and ‘female brain’? Baron-Cohen (2003) relies on an Empathy Quotient (EQ) and a Systemising Quotient (SQ) to diagnose ‘brain sex’. Questions measuring SQ include: ‘If there was a problem with the electrical wiring in my home, I’d be able to fix it myself’ and ‘When I read the newspaper, I am drawn to tables of information such as football league scores or stock market indices’. At the other end of spectrum are affective empathy statements such as ‘I really enjoy caring for other people’ and ‘I can easily tell if someone else wants to enter a conversation’. High scores on the systemising quotient indicate a systemising or ‘male brain’, while high scores on the empathy quotient indicate an empathising or ‘female brain’. Yet despite this, nearly half the women taking the test do not have a ‘female brain’ (Baron-Cohen, 2003). Moreover, it is highly questionable whether the EQ and SQ really tap into differential brain function; instead mirroring socialised behaviours stereotypically associated with male or female interests. Baron-Cohen claims to measure sexual differentiation in brains, but his measures merely tap into the degree an individual person fits into old-fashioned gender roles. Adding further doubt to the mix, Eisenberg and Lennon (1983) found that the effect size for gender differences in affective empathy becomes negligible when it is not obvious to the participant that affective empathy is being measured. Moreover, no gender differences were found for physiological facial/gestural measures of empathy. Essentially, Eisenberg and Lennon’s work showed that women and men actually differ more on how empathic they wish to appear, as opposed to how empathic they actually are. Thus, it seems that there is limited validity to a male female brain approach as proposed by Baron-Cohen, especially one that relies on self-report of traits and behaviour that are traditionally associated with masculinity or femininity.

Critiques and Questions re: Gender Essentialist Perspectives

> "Despite the many recent insights of brain research, this organ remains a vast unknown, a perfect medium on which to project, even unwittingly, assumptions about gender"
>
> (Fausto-Sterling, 2000, p. 118)

Gender essentialist perspectives that focus on biological causes of sex differences without acknowledging the role of social and cultural contexts have attracted criticism for basing their bold claims on inconsistent and contradictory results (e.g., Ceci et al., 2009; Fautso-Sterling, 1985; Fine, 2010). For instance, in their 2009 review Ceci et al. (2009) noted that research on hormonal causes behind women's underrepresentation in STEM showed a number of studies with contradictory results, failures to replicate, and trends that suggested that if there was any relationship at all between mathematical abilities and testosterone, it was definitely not linear as Baron-Cohen suggests (e.g., extreme male brain hypothesis, linking higher levels of testosterone to greater affinity with the sciences).

Sex difference researchers from gender essentialist perspectives have also been criticised for framing their research in reductionist terms, claiming that physiological processes can fully account for what happens on a psychological level (Magnusson & Marecek, 2012). Magnusson and Marecek note that biological accounts are usually discussed as if they are universally applicable, without scientifically testing these claims across diverse social and cultural contexts. In reality, the relations between biology, the environment, abilities and choices is far more complex than even original proponents of gender socialisation would have thought. For example, Fine (2010) argues that the brain's plasticity and malleability to the social environment means that any sex differences in brain structure and function may not necessarily be the result of innate sex differences – instead, structural and functional sex differences in the brain can arise from differential exposure to varying environments and experiences that occur as a result of gender socialisation.

Furthermore, Fine (2010) argues that biologically driven studies of sex differences have been plagued by poor methodologies, small unrepresentative samples, and over-interpretation of results, giving scientific authority to majority opinions. For example, one study reported that male newborns have a sex-bias towards things (moving mobiles) as opposed to emotions (faces), showing evidence of an innate sex difference (Connellan, Baron-Cohen, Wheelright, Batki, & Ahluwalia, 2000). However, in a critical review Nash and Grossi (2007) note that the study did not follow standard practice within infant research, whereby researchers present stimuli simultaneously to account for the poor attention span of infants. Additionally, because of poor vision, babies need to be placed in the same position, at the same angle in order for their responses to be comparable. Finally, there was the issue of experimenter expectancy effects (Nash & Grossi, 2007), where in this research it is critical to disguise the gender of the infants so the experimenter cannot subconsciously influence the integrity of the experiment by unintentionally giving cues or hints that would encourage babies to react to stimuli congruent with hypotheses. Unfortunately, Connellan et al. did not follow any of these practices noted above, thus leaving the data of the experiment wide open to bias and potential misinformation.

Despite methodological problems and over-exaggerated claims, the conclusions of evolutionary and biological sex difference research have trickled down to pop psychology. For example, Brizendine's (2006) New York Times best seller, '*The Female Brain*' makes the bold claim that women are so innately gifted in empathy, that they are able to know what their husbands are feeling, often before their husbands are conscious of their own feelings. According to Brizendine this is why females are destined to perform nurturing caregiving roles, while men remain innately gifted in the pursuits of scientific reason and rationality. Of course, Brizendine surmises that all of this is due to differences in exposure to testosterone during the prenatal period. And therein lies the problem. While these speculations of relationships between hormones, brain structure, and function might seem harmless at first, they become problematic when these narratives become accepted as scientific fact in popular culture, adding an air of scientific legitimacy to age old prejudices and stereotypes and potentially misinforming career advice for women and men (for a discussion of Baron-Cohen's career advice and critical policy decisions in education informed by gender essentialism see Fine [2010]). Biological and evolutionary perspectives of gender differences might be of value, but the problem is that overemphasis of results on methodologically poor studies, with small samples, and over-interpretation of findings within an overly deterministic lens can legitimise and further propagate unhelpful stereotypes that are detrimental to diversity of careers and skills that girls *and* boys feel they can aspire to.

Alternative Perspectives on Gender Differences (and Similarities)

Despite the failure of biological and evolutionary explanations to provide strong and consistent evidence to support gender essentialist claims regarding gender differences in educational and career pathways, the gender essentialist perspective still remains a dominant force in popular culture and within academia. Much of this seems to rest on the assertion that there are fundamental sex differences in ability and achievement. As such, much of the popular narrative on gender differences is that boys are intrinsically better at math and science. But how much evidence is there to support this claim?

Are men and women really from different planets?.

> “The general discussions of the psychology of sex, whether by psychologists or by sociologists show such a wide diversity of points of view that one feels that the truest thing to be said at present is that scientific evidence plays very little part in producing convictions”
>
> (p. 372, Helen Thompson Woolley, 1914)

The study of gender similarities dates back to Helen Thompson Woolley (1903), who, at the turn of 20th century asked the ground breaking question, “Do psychological differences actually exist between women and men?” To answer this question, Thompson administered a large battery of assessments and tests to undergraduate students. Tests assessed motor skills, sensory abilities, reaction times, ability to hit a target, rapidity of finger movement, pain thresholds, ability to discriminate heat and cold – abilities that were assumed to tap into the fundamentally different ways that men and women processed information. However, what Thompson Woolley actually found was that gender differences were markedly absent, thereby providing some of the first scientific evidence that questioned the age old assumption that men and women were vastly different in their skills and abilities.

Thompson Woolley was not the only psychologist who questioned whether gender differences had been overinflated in the scientific community. For example, Edward Thorndike (1914) argued that psychological differences were far too small, compared to within gender variation, to be considered important. Furthermore, in a pioneering review of gender research on mental traits, Leta Stetter Hollingworth (1918) concluded that there was little scientific evidence for psychological gender differences. Nonetheless, despite the work of Thompson Woolley, Hollingworth and Thorndike at the turn of the century, narratives of gender differences still dominated scientific discourse.

It was not until Maccoby and Jacklin (1974) in the 1970s, that scrutiny of sex difference research came to the forefront of the scientific community again. In their book, *The Psychology of Sex Differences*, Maccoby and Jacklin reviewed over 2,000 studies on gender differences across a variety of areas including abilities, memory, social behaviour, and personality. Again, their results revealed a similar pattern: gender differences previously assumed to be large and universal did not actually have much scientific weight behind them. With these findings, Maccoby and Jacklin dispelled many myths about gender differences that were prevalent at the time: that girls were more suggestible; that girls were better at rote learning and simple tasks, while boys excelled at higher level cognitive processing; and that girls had low achievement motivation. Out of thousands of studies, Maccoby and Jacklin only found evidence of gender differences in verbal ability, visual-spatial ability, math ability, and aggression. Nonetheless, psychology textbooks remembered Maccoby and Jacklin's work by their findings on these gender differences, rather than their overwhelming evidence of gender similarities (e.g., Gleitman, 1981).

The gender similarities hypothesis. The study of gender differences and similarities was revolutionised by the popularisation of the meta-analysis (Hyde, 2005). While previous reviews relied on the researcher reading and interpreting hundreds, if not thousands, of studies, meta-analytic procedures provided researchers with the opportunity to synthesise quantitative research findings across large numbers of studies through the computation of an effect size. Importantly, these new methods gave greater precision and accuracy to researchers reviewing quantitative research, allowing for stronger, more accurate conclusions to be made about the data (meta-analysis strengths and limitations are discussed in greater depth in Chapter 4).

The advent of the meta-analysis, and subsequent studies on gender differences using meta-analytic approaches to data analysis, gave way to Hyde's (2005) gender similarities hypothesis. Hyde's (2005) gender similarities hypothesis is based on the fact that most effect sizes for psychological gender differences are close to zero (d < +/-.10); a small effect size by Cohen's (1977) rule whereby small effects are (d = +/-0.20), medium effects are around d = +/-0.50, and d = +/-0.80 are considered large effects. In a review of meta-analyses published on gender differences, Hyde found that 30% of 124 published effect sizes were close to zero, and 48% of effect sizes were small. In other words, 78% of published gender differences in meta-analyses have effect sizes that are small or close to zero, meaning that the majority of gender differences are trivial, even if they happen to be statistically significant. This finding reflected a major problem within the field of psychology and gender, whereby researchers overemphasise the magnitude of effect sizes by focusing on statistical differences without looking at the actual size of effects, often leading to absolutist language (e.g., all men do X, and all women do Y; Chrisler & McCreary, 2010). Importantly, an independent evaluation synthesising 106 meta-analyses and 386 meta-analytic effects replicated Hyde's work, providing further compelling evidence to support the Gender Similarities Hypothesis (Zell, Krizan, & Teeter, 2015).

The evidence for Hyde's (2005) gender similarities hypothesis is particularly striking as it dispelled several pervasive myths about gender. For example, despite Carol Gilligan's (1982) influential theory on female and male moral reasoning, meta-analyses on gender differences in moral reasoning and orientation are only small (Jafee & Hyde, 2000). Similarly, although Maccoby and Jacklin's (1974) review concluded that there were reliable gender differences in math and verbal ability, meta-analyses revealed that the effect sizes for these differences were actually very small (see Hyde, 2005). Overall, Hyde's contribution to the study of gender within psychology has been important in that it has challenged long-held assumptions that differences between women and men are vastly different - an argument that has been echoed by many other feminist scholars (Fine; 2010, Lips, 2017).

Gender as dimensional rather than categorical. Finally, recent work (Carothers & Reis, 2013; Reis & Carothers, 2014) has extended upon the work of Hyde (2005) by investigating the degree to which gender differences reflected categorical differences between women and men, versus a more dimensional approach that favours gender differences on a continuum. Perspectives focusing exclusively on genetic, physiological, and biological mechanisms behind gender differences (as discussed earlier in this chapter), often rely on the assumption that sex differences are taxonic – reflecting the existence of two distinct categories of male/female. In contrast, research on how social and cultural forces influence the degree to which an individual develops gender-typical traits (e.g., Halpern, 2012; Wood & Eagly, 2012), typically reflects a dimensional perspective of gender.

The taxometric analyses employed by Carothers and Reis (2013), looks at not only the magnitude of gender differences, but also the distribution and pattern of differences across numerous variables. If gender was taxonic, as researchers such as Baron-Cohen would argue, then gender-typical behaviours would be expected to co-occur in all members of a sex, and little overlap would exist between men and women. In contrast, if gender is better understood as being dimensional, then a person's gender-typical behaviour on one variable would not imply being high on other gender-typical variables. Carothers and Reis found that gender differences for intimacy, femininity/masculinity, personality, empathy, relational interdependence, and science inclination were all dimensional. Only physical strength, sex-stereotyped leisure activities and anthropometric measurements were taxonic. Overall, results showed that while there were differences between men and women for psychological variables, these differences were not consistent or big enough to accurately ascertain group membership.

Although this thesis does not take a taxometric approach to analyses, the work of Reis and Carothers (2013; 2014) is important as it provides tangible, empirical evidence that dispels age-old assumptions that the psychologies of women and men are categorically and fundamentally different from one another, as often portrayed in gender essentialist viewpoints. Instead, the study of gender within psychology deserves a more nuanced approach that acknowledges that variations in gender-typical behaviours and psychological traits are likely to be the result of complex interactions between the social and cultural environment in addition to exclusively biological explanations.

Gender Socialisation Arguments: The Feminist Revolution Hits Psychology

> "Psychology has nothing to say about what women are really like, what they need and what they want, especially because psychology does not know"
>
> (Weisstein, 1968/1993, p. 197)

If the evidence for gender differences in biology *and* ability is inconclusive and at best underwhelming, what other mechanisms can explain women's underrepresentation in STEM? Could it be that social and environmental input is critically important to understanding women's underrepresentation in STEM? It was not until the mid to late 20th century that psychology began to face criticism for the way women were being discussed in research, and the lack of attention psychologists paid to social and environmental contexts of individuals. In the early 1970s at the peak of the women's liberation movement, a blistering critique of psychological research was published: "*Psychology Constructs the Female; or, The Fantasy Life of the Male Psychologist (With Some Attention to the Fantasies of His Friends, the Male Biologist and the Male Anthropologist)*". In this critique Weisstein (1968/1993) highlighted the unwillingness of psychologists to examine the impact of social and environmental forces when conducting male-female comparison research and reviewed empirical evidence from social psychology on social expectations and their effects on women. This prompted psychological researchers to incorporate social and environmental variables into gender research, and resulted in further research into the environmental inputs such as social expectations, situational demands, social rewards and penalties that could indeed produce gendered behaviour (Haaken, 1988; Sharps et al., 1994; Sherman, 1978).

Furthermore, female researchers began to highlight how the discipline of psychology was plagued with biases towards the experiences of men. The gender imbalance amongst psychology researchers was highlighted, and linked to a bias towards researching topics that were of social significance to men, while ignoring issues that affected women (Squire, 1989). Moreover, gender imbalances within samples were flagged as a critical issue within psychology (Squire, 1989). For instance, some of the major findings in social psychology were based primarily on male samples (Squire, 1989). These findings were often generalised to the rest of the population without any question of whether a broader, more diverse sample was needed to establish broad theories about human behaviour and cognition. (Magnusson & Marecek, 2012; Squire, 1989). Indeed, when researchers use all female, as opposed to all male samples, they are more likely to provide justification for their sample choice, and are more likely to provide a caveat that their findings do not generalise to the wider population (Ader & Johnson, 1994).

Feminist psychologists also highlighted the issue of ‘context-stripping’ in research – a practice whereby behaviour, thoughts, or feelings of an individual or group are treated as if they were separate from the context in which the individuals and groups usually existed (Squire, 1989). For example, if women from the Victorian era were shown to be less able to perform on tests of IQ, this was regarded as evidence of a deficiency in intellect caused by being female, rather than the effects of not having access to basic education and schooling like their male counterparts (Fine, 2010). Lips (2017) argues that context stripping is problematic because it is reductionist and produces a distorted picture of reality, that can sometimes be genuinely harmful to the subjects of the research focus.

Chapter Summary

Chapter 2 has provided a review of the historical context of the study of gender differences in psychology. The major counterpoint to EVT and other social cognitive theories explaining gender differences comes from evolutionary and biological gender essentialist theories. This chapter has reviewed the current evidence in the area, highlighting the inconsistencies and conflicting evidence within the biological and evolutionary framework of gender differences. Despite inconsistencies and over-exaggerated claims many gender essentialist views depict biology as the only driving force behind gender differences in STEM educational and occupational attainment. Essentially, modern day gender essentialist views appear to be making similar logical fallacies to that of their predecessors of the 19th century who used science to confirm sexist narratives about women and men, rather than interpreting evidence on its own merits. Alternative perspectives of gender such as Hyde’s (2005) Gender Similarities Hypothesis have provided evidence that gender differences are unlikely to have their origin in aptitude or ability. Furthermore, alongside the feminist movement in the seventies was an array of social-cognitive theories that began to take notice of the social causes of gender differences in psychology. Thus, this thesis will use EVT, a theory grounded in social and cognitive forces, to explain gender differences in young people’s educational attainment through expectancies for success and value.

CHAPTER 3

EXPECTANCY VALUE THEORY: CONNECTING THE DOTS BETWEEN THE SOCIAL ENVIRONMENT AND GENDER DIFFERENCES IN EDUCATIONAL OUTCOMES

Chapter 2 outlined the historical context and development of gender differences within the discipline of psychology. As mentioned earlier, the study of gender differences flourished alongside the popularisation of social learning perspectives. But how did these perspectives emerge? In Chapter 3, I detail the emergence of *Expectancy Value Theory* (EVT), the key focus of this thesis, charting the development of the theory, to modern applications of the theory relating to the issue of gender differences in educational and career choices. Current empirical evidence regarding the role of gender socialisation in the development of self-beliefs and values will be reviewed. Finally, I conclude by outlining several major research gaps in the area of EVT: a) a need for research to explore the degree to which EVT can account for current gender differences in educational attainment in university STEM course selection; b) a need for more qualitative research that explores alternative mechanisms to explain gender differences in tertiary educational attainment that cannot be fully explained by EVT; and c) a need to investigate the social and cultural contexts in which gender differences in expectancy for success and task value might vary.

From Rats in Mazes to Girls in Math: Charting the Development of EVT

In the mid to late twentieth century new theories emerging from gender socialisation perspectives and social learning began to blossom. These theories stemmed from the work of behaviourists, who used laboratory and experimental designs to deconstruct processes of motivation through schedules of reward and punishment. Behaviourist experiments focused on the observable, and focused on simple scenarios, often with animal subjects in order to capture cause and effect relationships between stimuli and behaviour (e.g., Skinner, 1953). However, a major critique of behaviourism was that it neglected the rich inner lives of individuals. That is, the motivation of human beings is more than simple responses to external reward and punishment. Instead, motivation is a complex construct tied in with cognition and thought, whereby expectancies for success, attitudes and values shaped by our experiences in the world can also play a critical role in behaviour (Lewin, 1938; Tolman, 1932; but also see Feather, 1982; Weiner, 1989/1980; and Wigfield & Eccles, 1992).

Kurt Lewin (1938) and Tolman (1932) were amongst the first theorists to highlight the cognitive components of *expectancy* and also *value* (or valence), and their role in predicting behaviour. Lewin discussed how positive and negative valence ascribed to a situation was critical in determining actions, and Tolman discussed how expectancy for success influenced motivation and subsequent behaviours in different domains. For instance, Tolman found that the behaviour of rats changed as a result of exposure to different alley openings (one encouraging further exploration for food, and one encouraging non-exploration because of the absence of food). Tolman theorised that rats developed an "*expectancy for success*" when faced with alley way openings that were similar to those that had previously led them to food. Their anticipation of success then motivated them to continue towards approach behaviour that would result in reward. This prompted speculation that expectancy for success was a critical mediator between stimuli and observed behaviour.

Can a theory based on rat behaviour truly explain the complexities of human motivation? Atkinson (1957, 1964) was one of the first researchers to translate these findings to achievement motivation to explain achievement behaviours such as striving for success, persistence, and to explain differences in choice of achievement behaviours. Atkinson was influenced by both Tolman's (1932) work on expectancy for success and Lewin's (1938) emphasis on how the valence of an activity is critical to determining later actions. To explore these constructs further, Atkinson conducted a series of experiments designed to test how individual differences in achievement related motives influenced behaviour in competitive achievement contexts. Atkinson proposed three achievement related motives: 1) *expectancy:* the cognitive anticipation that performing an action will be followed by a particular consequence; 2) *incentive:* the relative attractiveness of a specific goal, or the relative unattractiveness of an event that might occur as a consequence of an act (e.g., the amount of reward or punishment that occurs after an action); and 3) *motive*: the disposition to strive for satisfaction of attainment of incentives (e.g., achievement, affiliation, power) or to minimise pain or avoid failure (e.g., avoidant tendencies or aversions), originally measured using the Thematic Apperception Test (see McClelland, 1985). According to Atkinson, motivation is defined as:

$$motivation = (motive \times expectancy \times incentive)$$

Applying EVT to gender differences in educational outcomes. While EVT emerged in an experimental paradigm, modern day EVT is widely known for its application in explaining gender differences in educational outcomes (e.g., Eccles, 1994; Eccles, 2005; Eccles & Wigfield, 2002). There are several key differences that separate modern day EVT from Atkinson's (1957) experimental work. Firstly, while Atkinson argued that value and expectancy were inversely related, Eccles and colleagues' (1983) conception of EVT considers expectancy and value as constructs that are often reciprocally and positively related. Secondly, modern day EVT includes a substantial expansion on the original EVT constructs to include multiple subcomponents of task value, and also a clearer definition of each sub-component within the theory (see Figure 1). Thirdly, while Atkinson's early EVT work primarily focused on developing theory within an experimental/laboratory context, modern day EVT has extended its breadth to longitudinal and cross-sectional data. Finally, Eccles and colleagues' conceptualisation of EVT explicitly highlights the role of social and cultural context in determining self-beliefs and values. Indeed, according to Eccles and colleagues (e.g., Eccles & Hoffman, 1984; Eccles & Jacobs, 1986; Eccles, Jacobs, & Harold, 1990), gender differences in educational and career pathways are primarily the result of gender socialisation experiences that in turn affect young people's self-beliefs and values, leading to vastly different career choices and behaviours (see Figure 1). Thus, according to Eccles (1986; 1994), girls disengage from math and science, not because they lack ability or aptitude (as discussed in Chapter 2), but because of low expectancies for success and low task value in STEM subjects that are in large part shaped by social and cultural milieu.

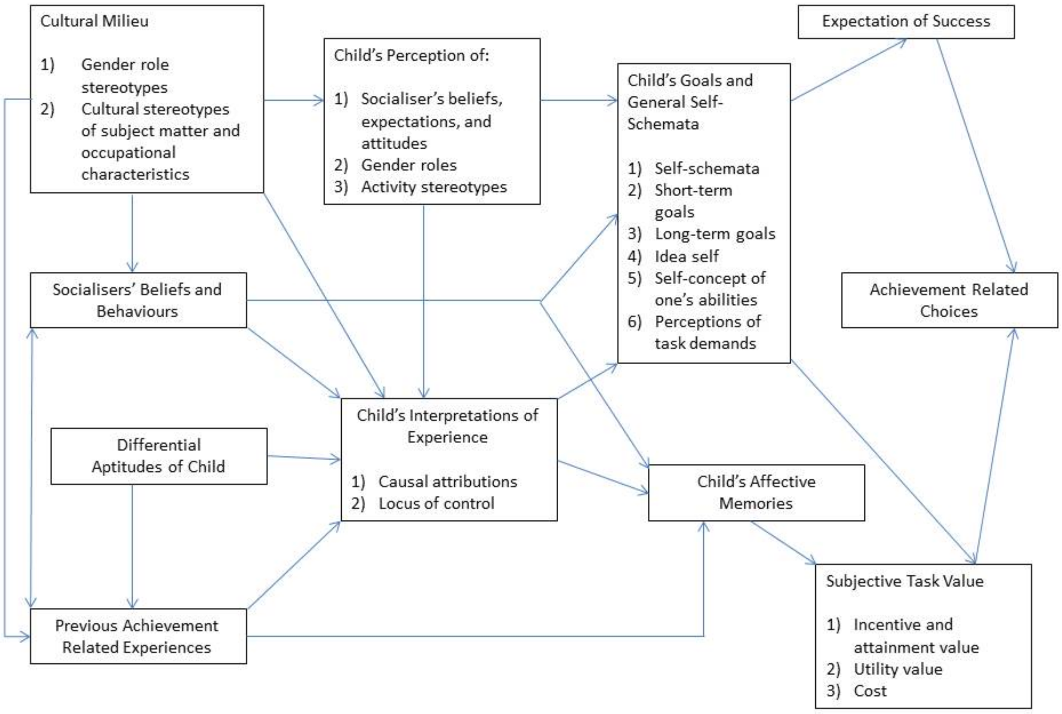

Figure 1. Expectancy value model based on Eccles, Adler, Futerman, Goff, Kaczala, Meece, and Midgley (1983).

Expectancy for Success

Eccles (1994) describes expectancy for success as an individual's sense of domain specific personal efficacy (e.g., I have an ability to do well in science). An individual's expectancy for success largely depends on their self-perceptions of their own abilities, as well as their estimation of difficulty for the course or task (Eccles, 1994). There are various expectancy constructs that are encompassed under the umbrella of expectancy for success. Constructs commonly used to refer to expectancy for success include domain specific self-concept, self-efficacy, expectancies for success, and self-perceived competence (Bandura, 1997; Pintrich, 2003). While these constructs have some theoretical and operational differences, they are unified by the principle that students who believe they can do well are more likely to be motivated and persist on a task (Bandura, 1997; Pintrich, 2003). Thus, researchers commonly use the term expectancy for success as a term that refers to domain specific self-concepts, self-efficacy and perceptions of competence (e.g., Gaspard, 2016).

From Bandura's (1977) social-cognitive theories on self-efficacy, to Eccles' (1994) work on EVT and expectancy for success, and finally to Marsh's (1990) multidimensional self-concept model, the importance of self-beliefs in predicting positive educational outcomes has long been recognised by researchers in the field. There is strong evidence to suggest that there are consistent gender differences in math self-concept/expectancy for success favouring boys (e.g., Skaalvik & Skaalvik, 2004), suggesting that self-beliefs in mathematical ability could be critical in explaining the gender gap in STEM attainment in higher education.

Despite strong emphasis on enhancement of self-beliefs as key goals of educators, Parker, Marsh, Ciarrochi, Marshall, and Abduljabbar (2014) note that there has been limited research which has considered the role of academic self-beliefs in predicting university enrolment. Nonetheless, the limited research available suggests that these factors are important in student educational attainment. Self-efficacy, for example, has been linked with higher academic aspirations (Bandura, Barbaranelli, Capara, & Pastorelli, 1996). Research has also shown that self-concept is a significant predictor of university enrolment, even when controlling for prior achievement (Marsh, 1991; Parker, Schoon, et al., 2012). More recent research has confirmed these findings in the STEM discipline. Math self-efficacy at Grade 12 has been linked to intention to major in a STEM field at college (Wang, 2013). Finally, Parker et al. (2014) showed that math self-concept remains a significant predictor of STEM university enrolment, even when other demographic and achievement factors are controlled for.

Task Value

Expectancy value built on previous literature by highlighting the importance of value judgements in shaping the education and career decisions of young people (Eccles, 1994). Students do not only perform self-assessments of abilities when they are making choices about educational engagement, instead they also rely heavily on their value judgments about the activity they are pursuing.

Value in EVT is a multidimensional construct that takes into account an individual's beliefs about a task or the outcomes' attainment value (importance), intrinsic value (interest), utility value (usefulness) and potential cost (Eccles et al., 1983). Although these constructs have been proposed as conceptually distinct, task values have often been analysed as a single factor with empirical studies often combining utility, attainment, and intrinsic value together (e.g., Anderman et al., 2001; Bong, 2001; Jacobs et al., 2002; Wigfield et al., 1997).

Attainment value. Eccles et al. (1983) define attainment value as the importance of doing well in a task. Attainment values develop from one's identity, as tasks become important when an individual views the task as central to their sense of self. Importantly, attainment value is often entwined with notions of masculinity and femininity, as individuals are driven to engage in tasks that reflect and confirm salient aspects of their self-schema (Eccles et al., 1983).

Intrinsic value. Intrinsic value (or interest) is the enjoyment that an individual gets out of doing a task. Intrinsic value is similar to Harter (1981) and Deci and Ryan's (1985) construct of intrinsic motivation. When a student engages in a task that evokes intrinsic value, they are typically able to persist and sustain interest in that task for a long time (Wigfield & Cambria, 2010). Differences in interests have long been touted as a key mechanism behind gender differences in educational attainment by researchers from both gender essentialist and gender socialisation perspectives (Baron-Cohen, 2003; Brizendine, 2006; Buss, 1991; Eccles, 1994; Ridgeway, 2009), and gender differences in math interest have been well documented (e.g., Preckel, Goetz, Pekrun, & Kleine, 2008). Furthermore, there is emerging evidence to suggest that interest and liking of math are the strongest predictors of math course selection in senior high school for Australian adolescents (Watt, Eccles, & Durik, 2006).

Utility value. In addition to interest, other value beliefs have also been found to be integral in predicting student choice behaviour. One of the most researched of these is utility value, commonly defined as the usefulness of a task in the immediate or long-term future (Wigfield & Cambria, 2010). Utility value resembles extrinsic motivation as it refers to how well a task matches an individual's future plans (e.g., taking statistics to complete a psychology degree; Wigfield & Cambria, 2010). However, Wigfield and Cambria note that despite some similarities to Ryan and Deci's (2000) concept of extrinsic motivation, utility value also can be deeply tied to an individual's sense of self. For instance, a task with high utility value may be connected to an individual's long held personal career goal. Utility value has been discussed somewhat less in the literature on STEM university entrance predictions. However, recent work has indicated that these values, alongside interest, are important and are more powerful predictors of university enrolment than achievement or prior enrolment in high school STEM (Maltese & Tai, 2011).

Cost and math anxiety. Finally, cost refers to what an individual has to give up to engage in a task (e.g., studying for an upcoming science test means I can't go out with my friends), and the anticipated effort that will be required for a task (Eccles et al., 1983). Despite the theoretical importance of cost, and its history of importance in micro-economics, it has been the least studied component of modern day EVT (Wigfield & Cambria, 2010).

Other possible constructs similar to the EVT construct of 'cost' that might explain gender differences in educational attainment include math anxiety, or emotional costs of entering STEM (e.g., Perez, Cromley, & Kapaln, 2014). Recent research has begun to explore other drivers of disengagement in STEM, particularly in terms of emotional cost, stress and anxiety. A recent longitudinal study charting college chemistry students' intents to leave a STEM course found that perceptions of effort cost (e.g., drawbacks associated with time and effort, lost opportunities, stress, and anxiety) predicted intent to leave a chemistry degree (Perez, Cromley, & Kaplan, 2014). Importantly, research has demonstrated that female students are more likely to report greater levels of math anxiety (e.g., Devine, Fawcett, Szűcs, and Dowker, 2012).

Development of Expectancies for Success and Task Value

Eccles and colleagues (e.g., Eccles & Wigfield, 2002) have implicated a number of factors that play a role in determining an individual's expectancies for success and value of a task. These factors include: an individual's cultural milieu (e.g., the cultural and social setting/environment in which a person lives); the beliefs and behaviours of socialisers (e.g., parents and teachers); differential aptitudes of individuals; previous achievement related experiences; individual perceptions of social beliefs; an individual's interpretations of experiences; affective memories; general goals; and self-concepts (see Figure 1). Importantly, modern EVT differs from Atkinson's (1964) classical EVT in that it places greater emphasis on socio-cultural forces that shape an individual's beliefs and self-perceptions (Eccles & Wigfield, 2002).

Schoon and Eccles (2014) highlight the importance of gender identity in shaping career expectations for success, and determining values that ultimately lead to different educational/occupational pathways for women and men. For example, Eccles et al. (1983) state that the attainment value, or level of importance an individual places on a task, is influenced by a person's existing and ideal self-schema. Drawing on self-schema theory (Markus & Wurf, 1987), Eccles et al. argue that tasks provide individuals with the opportunity to act in congruence with aspects of their existing and ideal self-schema. According to Eccles et al. people are more likely to assign higher attainment value to tasks that allow them to demonstrate important aspects of their self-schema (e.g., masculinity or femininity). Likewise, people are less likely to act in ways that disconfirm their already existing self-schemas. Schoon and Eccles (2014) state that self-perceptions develop in the context of an individual's experience of the cultural milieu they find themselves in. For example, a young male may be more likely to aspire to be a scientist because he has been exposed to many positive examples of male scientists throughout his life. His exposure to successful male role models in science may then have an impact on his self-concept (e.g., having an increased value for science based careers or feeling confident in his own ability to perform well in science based activities). Thus, in this framework the existence of stereotypes and cultural norms becomes integral in explaining an individual's self-concept as well as their values.

Socialisation as a Predictor of Self-Beliefs and Values: Empirical Evidence

At the core of modern EVT is the proposition that the social and cultural context is critical to the development of self-beliefs and values. However, to what extent can empirical evidence support the theoretical claims of EVT? Here, the current literature on the relationships between socialisers, gender identity, gender roles and gender typicality are reviewed.

The role of parents. There is now a wealth of evidence to support EVT's assertion that parents can have a significant impact on student motivation, grades, and achievement related behaviour (e.g., Ahmed, Minnaert, van der Werf, & Kuyper, 2010; Bouchey & Harter, 2005; Bowen, Hopson, Rose, & Glennie, 2012; Felson, 1989; Ferry, Fouad, & Smith, 2000; Frome & Eccles, 1998; Halle, Kurtz-Costes, & Mahoney, 1997; Luster & McAdoo, 1996; Parsons et al., 1982; Trent, Cooney, Russell, & Warton, 1996; Wang & Staver, 2001; Wentzel, 1998; Wilkins & Ma, 2003). For example, Frome and Eccles (1998) found that mothers' perceptions of ability are a stronger predictor of children's self-beliefs and task values than a child's actual grades. In keeping with gender stereotypes, mothers of daughters underestimated their child's math performance, while mothers of sons overestimated their math performance. Tenenbaum and Leaper (2003) have replicated these findings, showing that even when there were no gender differences in children's science performance, self-efficacy and interest, parents *still* are more likely to believe that daughters found science more difficult and less interesting compared to sons. Furthermore, fathers were more likely to use cognitively demanding speech with sons than daughters when helping them with a scientific task, perhaps reflecting parents' greater tendency to believe their sons to be more capable in science compared to daughters. Even more concerning is that these parent perceptions were a significant predictor of children's self-efficacy and interest in science.

The role of peers. As young people enter adolescence, peer perceptions become increasingly important to the formation of student self-beliefs, values, and attitudes towards science (Brown, 1990; Kindermann, 1993). For instance, Leaper, Farkas, and Brown (2012) found that perceived peer math and science support is a significant predictor of adolescent girls' math and science motivation, however, this finding is consistent across genders and a variety of academic subjects (e.g., Crosnoe, Riegle-Crumb, Field, Frank et al., 2008; Rice, Barth, Guadagno, Smith & McCallum, 2013; Robnett & Leaper, 2013; Stake, 2006, Stake & Nickens, 2005). Nonetheless, peer support may be particularly important for girls in non-traditional areas such as math and science (Crosnoe et al., 2008).

The role of teachers. Finally, EVT posits that teachers also play an important role in the formation of student self-beliefs, values and achievement related behaviour (Eccles et al., 1993). Indeed, teacher support and engaging instruction has been associated with higher math and science self-efficacy, and also better attitudes towards math and science during the transition to middle school and high school (e.g., Ahmed et al., 2010; Barth et al., 2011).

The Influence of Perceived Gender Roles and Identity on Self-Beliefs and Values

Measuring the impact of gender roles and stereotypes can be fraught with difficulty. Social norms relating to gender are difficult to 'see', and because of this, the effects of these norms on our self-beliefs and attitudes often occur outside of conscious awareness. Nonetheless, some studies have attempted to measure the impact of gender roles, identity, and stereotypes on self-beliefs through self-report measures, thereby lending support to EVT's proposition that social and cultural norms regarding gender can have powerful and important effects on student motivation. For example, Leaper et al. (2012) found that parental pressure to conform to gender roles was negatively related to student math/science motivation for adolescent females. Similarly, Dinella, Fulcher and Weisgram (2014) found that pressure to conform to gender roles positively predicts feminine career interests. In contrast Leaper et al. (2012) found that adolescent girls' exposure to feminism and endorsement of gender egalitarian values were positively related to math/science motivation. Gender egalitarian values have also been positively related to higher levels of academic achievement for adolescent girls (e.g., Valenzuela, 1993). Indeed, an awareness of feminist issues such as the female disadvantage in STEM seems to be related to better outcomes for young girls. Weisgram and Bigler (2007) showed that valuing of science amongst girls *increases,* when girls are taught about the discrimination of women working in STEM fields.

Other research has shown that sexist attitudes and gender typicality (the degree to which one relates to traditional gender masculine or feminine characteristics) is related to undergraduate men's ability beliefs and interest in areas incongruent with gender roles, and even young men's selections of college majors classified as traditionally feminine areas of study (Leaper & Van, 2008). In other words, men who endorse traditionally masculine gender characteristics and sexist attitudes are more likely to have poorer self-beliefs and interests in traditionally non-masculine areas of study, and are less likely to select college majors that are incongruent with their gender roles. Similar results were found in a sample of males and females, whereby gender typicality negatively predicted career interests that were incongruent with one's gender (Dinella et al., 2014).

In summary, studies examining the effects of parents, teachers, and peers show there is evidence that supports the socialisation hypothesis inbuilt in EVT. Furthermore, there is emerging evidence that suggests gender identity, typicality, exposure to feminist ideas, gender egalitarian views and sexist attitudes can predict student self-beliefs and values in areas of study that are incongruent with students' genders.

How Early do Gender Differences in Self-Beliefs and Values Emerge?

Having established the efficacy of expectancy value interventions, and possible causes of these beliefs and attitudes, a critical question for researchers to understand is at what age do these gender differences emerge? And furthermore, to what degree are these differences stable across age?

Earlier research has shown that children as young as 3.5 years are able to acquire knowledge of sex-role stereotypes (e.g., Reis & Wright, 1982). However, at what age do these differences begin to emerge for expectancy value variables? Several studies have attempted to answer this question (e.g., Frenzel et al., 2010; Jacobs et al., 2002; Nagy et al., 2010; Watt, 2004). Eccles, Wigfield, Harold, and Blumenfield (1993) and Wigfield et al. (1997) did not find any gender differences in math values among primary school aged children. However, differences in math enjoyment have been reported as early as Grades 3-5 (Frenzel, Pekrun, & Goetz, 2007; Lichtenfield, Frenzel, & Pekrun, 2007).

Research charting changes in EVT variables across age has provided mixed evidence regarding whether or not gender differences in expectancy value grow larger with age. Some studies have found support for a gender intensification hypothesis, whereby gender differences increase as children develop due to increased exposure to gender socialisation processes (e.g., Eccles, 1987; Hill & Lynch, 1983). For example, in a review of student attitudes towards science, Brotman and Moore (2008) note that several large quantitative studies show that girls' attitudes towards science are less positive than boys' attitudes, and often decline significantly with age. However, other studies have yielded conflicting results, with little evidence for changes in gender differences according to age (Fredricks & Eccles, 2002; Jacobs et al., 2002; Marsh, 1989; Watt, 2004). For instance, longitudinal evidence has shown consistent gender differences in math values across age, but a lack of evidence for gender differences in curved trajectories, meaning that the size of gender differences are likely to be stable across time (Fredricks & Eccles, 2002; Jacobs et al., 2002; Watt, 2004). An exception to this was work by Frenzel, Goetz, Pekrun, and Watt (2010) that showed a distinct period during Grade 7 where gender differences became significantly greater for girls in comparison to boys; showing that declines in boys' math interests levelled out, whilst girls had a steeper downwards trajectory for interest. Nonetheless, in this study interest levels plateaued later in adolescence, and the gender gap did not widen any further (Frenzel et al., 2010). Overall, it seems that there is consistent evidence for a male advantage in STEM subjects that occurs as early as elementary school, but there is some mixed evidence regarding the stability and size of this advantage across child, adolescent, and adult development. Thus, there is a need to synthesise current research to form stronger conclusions on how expectancy for success and task value interacts with age. This thesis will endeavour to address this question through a synthesis of current findings in a meta-analysis on expectancy value theory.

The Present Investigation: New Directions for Research on Gender and Expectancy Value Theory

The above review demonstrates the strong theoretical and empirical foundations of EVT, as well as longitudinal evidence for the importance of various EVT variables in predicting later outcomes (for a review of experimental evidence see Gaspard's 2016 dissertation on task value interventions, and O'Mara, Marsh, Craven & Debus, 2006 for their review on self-concept interventions).

There are, however, many unresolved questions in relation to gender and EVT. For instance, EVT research on the relationships between gender, self-beliefs and task values has been extensive; however, very little attention has been paid to the social and cultural contexts in which these gender differences may vary. For instance, do young girls from different social class experience gender disadvantages in math attitudes similarly, or are there critical differences in the size of the gender gap amongst different social classes? This issue will be a key focus of the thesis, and will serve as an overarching research question guiding Study 1a and Study 1b.

There has also been a dearth of research that looks at what extent traditional EVT variables can account for observed gender differences in tertiary educational attainment. In other words, when we compare young people of equal ability, expectancies for success, and task value, what is the size of the remaining residual unexplained gender effect? Following from EVT, it would be expected that these variables should account for a substantial amount of the gender effect, however, little research has actually explored this question in detail. Moreover, little research has explored qualitative responses of young people in relation to their choices to disengage from, or engage in, STEM. This issue will be the second focus point and research question of the thesis, guiding Study 2a and 2b.

Research Question I: Exploring Social and Cultural Contexts of Gender Differences in EVT Variables

Eccles and Wigfield (2002) recognise that one of the limitations of the current EVT research is the lack of consideration of how contextual influences shape student motivation, and even go as far to conclude that it is "difficult if not impossible to understand students' motivation without understanding the contexts they are experiencing" (pg. 128). Despite the wealth of research on gender differences and similarities in educational and occupational outcomes, there has been comparatively less research that has investigated the specific contexts in which gender differences in educational outcomes may vary. Indeed, this research gap is reflective of a wider problem within psychology, whereby psychology's focus on the individual in isolation to their contexts means that the impact of societal and structural inequality on marginalised individuals is often ignored (Fox, Prilleltensky, & Austin, 2009). Arguably, if we fail to consider how social, cultural, and political contexts of individuals feed into thoughts, appraisals, wellbeing, and values, then we are neglecting a potentially powerful and important influence on people's inner lives and experiences.

A Brief Introduction to the Intersection of Gender and Socio-cultural Contexts

Psychological research on gender has faced similar difficulties. Indeed, there has been criticism over the current lack of investigation into how gender intersects with other social and cultural categories (Else-Quest & Hyde, 2016a; Hyde, 2007). For example, in Eagly, Eaton, Rose, Riger, and McHugh's (2012) review on feminism and psychology, Eagly and colleagues found that only a small minority of studies on the psychology of women and gender have attended to the heterogeneity or diversity within gender by including analyses of gender across social class, sexual orientation and ethnicity. Eagly and colleagues noted that the intersection of gender with social class and sexual orientation was particularly under-researched, and thus flagged intersectionality as an area needing more research. Similarly, Hyde (2007) emphasised that much of what we currently know about gender and psychology is actually the experience of gender for American middle-class white college students (Hyde, 2007; see also Henrich, Heine, & Norenzayan, 2010). Consequently, little psychological research has investigated how gender effects in EVT are moderated by ethnicity, class, and nationality. Hyde (2012; 2013) flags this as an area in which new research is crucial investigating not only gender, but also the *intersection* of gender with social categories like class and ethnicity. While some evidence is available (see Chapter 5), there is a lack of research that brings together existing literature and data, to explore gender differences through an intersectional perspective. Addressing this research gap will be a primary aim of this PhD.

Although an in-depth discussion of intersectional research is beyond the scope of this PhD, it is important to provide a brief overview of the concept and its relevance to current questions in EVT research. Intersectionality is a broad approach that focuses on understanding inequality in terms of multiple social categories (e.g., ethnicity, gender, class), and focuses on how these overlapping identities are related to experiences of disadvantage (or privilege) and difference (see Choo & Ferree, 2010; Cole 2009 for a review). By and large, intersectionality emerged as a feminist perspective in response to critiques that the feminist movement only focused on the concerns of white women, and in doing so excluded women of colour (e.g., Crenshaw, 1989). Although there are many definitions and conceptualisations of intersectionality, at the heart of the approach is that social categories such as gender, social class, and ethnicity are entwined and interrelated; meaning that it cannot be assumed that the experience of gender is the same for all women and men, rather this experience may be contingent upon multiple other social identities.

While intersectionality is largely abstract, there are significant practical implications that an intersectional research agenda can have that make it an important research avenue to pursue. Namely, by utilising an intersectional research agenda there is a better chance of policies not being based on white, North American, straight women, to the exclusion of all others. Dating back to Brofenbrenner (1979), psychologists have understood that thoughts, feelings, and behaviours do not exist in a vacuum, instead the wider social and cultural context is critical to the individual experience. By using an intersectional perspective to guide our research questions and agenda, we can provide a more nuanced account of the role of power, privilege, social context, and how membership across multiple social categories impacts on the individual experience. Essentially, an intersectional perspective offers a unique opportunity to refine research, education, healthcare, and policy decision making to cater for women and men from many backgrounds, not just one.

Are Intersectional Research Questions Compatible with EVT and Quantitative Research?

Intersectional feminism emerged in large part from critical theory that challenged positivist assumptions about science and knowledge (see Bowleg, 2008 for a critique of quantitative methods in relation to intersectionality). This thesis largely relies on quantitative methods to explore questions about gender and educational attainment in the context of EVT. Thus, a controversial question remains: is intersectionality a theory that can be tested? And moreover, is it a perspective that is compatible with quantitative methods? Indeed, on the surface, quantitative research and intersectional feminism seem like strange bedfellows, but Else-Quest and Hyde (2016b) note that there are, in fact, many reasons why this should not the case. Else-Quest and Hyde argue that instead of conceptualising intersectionality as a falsifiable theory, it is better understood as critical perspective or approach that can be applied to existing psychological theories (e.g., EVT) to influence the choice and focus of research questions and hypotheses. Indeed, there have been several successful applications of intersectionality in quantitative methods across other disciplines, showing that quantitative research methods can provide new and novel insights into the study of intersectionality (e.g. Choo & Ferree, 2010; Spierings, 2012; Few-Demo, 2014).

A Review of Current Literature Relevant to Research Question I: Expectancies and Values Towards Math and Science and the Intersection of Gender, Social, and Cultural Context

What is the state of the existing literature in relation to research exploring the intersection of gender and socio-cultural contexts in expectancies for success and task value? The following section aims to provide a review of the current research in this area that is relevant to the first research question guiding this PhD. Here, I review how gender differences in educational variables intersect with social and cultural context, with particular focus on the relationship between national level gender equality, ethnicity, social class, and geography, and gender differences in expectancy value variables.

The Intersection of Gender and Ethnicity

As mentioned in Chapter 1, alongside a large gender gap in STEM, there is also an underrepresentation of students from ethnic minorities, especially for Latinos, African Americans, and Southeast Asians (National Science Foundation[NSF], Division of Science Resources Statistics, 2012). The representation of women from ethnic minorities is particularly troubling, given that only 8% of STEM jobs in the United States are held by women from ethnic minorities (NSF, Division of Science Resources Statistics, 2012).

What is the cause of this underrepresentation? One difficulty is that girls from some ethnic minorities often face negative stereotypical attitudes for both their gender (as discussed in Chapter 2) *and* their ethnicity. For example, Asian American and White Americans males are more likely to be stereotyped as having high abilities in STEM compared to males from other ethnicities (Sinclair, Hardin, & Lowery, 2006; Wenner, 2003). Furthermore, when young children are asked what a scientist looks like, the overwhelming majority of them draw a White male (Barman, 1997; Wenner, 2003). In line with Expectancy Value Theory, these stereotypes are powerful because they have the ability to shape young people's expectancies for success and attitudes towards STEM, and their mental representations of who can and cannot be a scientist. Thus, maybe the gap in STEM participation is not only about gender. Instead, it is critical to consider to role of gender in relation to ethnicity when exploring student motivation in STEM.

Out of the thousands of studies on attitudes towards science and math, only a small number have focused on the intersection of gender and ethnicity, and as discussed earlier in this chapter, most research has failed to include ethnically diverse samples in research on the STEM gender gap. Catsambis (1994; 1995) was one of the first researchers to explore this research question with regards to math and science attitudes. Catsambis found that there was slight variability in gender differences across Black, Latino/Latina and White students, but males tended to endorse more positive attitudes towards STEM. However, there was heterogeneity between ethnicities – with White students having the largest gender gaps for science attitudes, and Latino/Latina students exhibiting the largest gender gaps for math attitudes.

Other researchers have explored gender and ethnicity with regards to technology in addition to math. Zarrett, Malanchuk, Davis-Kean, and Eccles (2006) found gender differences favouring males in math and technology attitudes that existed regardless of whether students were African American or White. More specifically, gender differences in computer self-concept for programming and developing software, were medium for White students and small for African American students. Gender differences were almost non-existent for White students in computer self-concept in word processing and accounting, and very small for African American students (favouring females). Data about intentions to enrol in computer science majors show a similar pattern, whereby the gender gap for white students is (9:1), compared to (4.5:1) for other Latino, African American and Asian American students (NSF, Division of Science Resource Statistics, 2012). These findings suggest that gender gaps are the largest between White male and female students.

A similar pattern has occurred for math and science achievement, whereby the gender gap is greatest amongst White students. Data has shown that gender gap in achievement is largest for White and Latino/Latina students (Catsambis 1994; 1995**,** Coley, 2001; Else-Quest et al., 2013; McGraw, Lubienski, & Strutchens, 2006). Thus, there seems to be some evidence of an ethnicity by gender interaction amongst American samples, whereby gender gaps are the largest and most consistent for White and Latino/Latina students, and smaller for Asian American and African American students.

However, a recent study of adolescents by Else-Quest, Mineo, and Higgins (2013) has demonstrated that while there were some slight variations in the size of gender differences between Latino/Latina, Asian American, and Caucasian students for self-beliefs and values in science and math, these differences were not statistically significant. While these studies have provided valuable preliminary insights into the exploration of the intersection of gender and ethnicity for STEM attitudes, there has been little attempt to explore the ethnicity and gender interaction in settings outside the US. Thus, this thesis will endeavour to explore the gender ethnicity interaction in an Australian context, including Indigenous Australians.

National-Level Indicators of Gender Equality and Relations to EVT variables

Recent scholars have begun to assert that social and structural inequality is not merely a political issue; instead it is critical to individual functioning and wellbeing (Glick & Fiske, 2001**;** Jenkins, 2000; Lykes, 2000; Prilleltensky, 2008; Zurbriggen & Capdevila, 2010). One way to investigate the relationship between macro structures of power and inequality is to assess the relationships between individual level thoughts, feelings, and behaviours, and national level indicators of social equality. Thus, this section of the literature review will focus on documenting the current literature on gender differences in math performance and attitudes in relation to national indicators of gender and social equality.

Many studies have explored gender differences in math achievement and the relation with national level gender equality indices (e.g., Baker & Jones, 1993; Else-Quest, Hyde, & Linn, 2010; Fryer & Levitt, 2010; Guiso, Monte, Sapienza, & Zingales, 2008; Hyde & Mertz, 2009; Kane & Mertz, 2012; Penner, 2008; Riegle-Crumb, 2005). Overall, these studies suggest that indices of gender equality are likely to be related to smaller gender gaps in math achievement, with some possible exceptions in Middle-Eastern countries and domain specific measures such as labour force representation. Nonetheless, these studies provide initial evidence that the wider cultural context plays a critical role in determining gender differences in math performance.

Gender equality and math attitudes. As mentioned above, recent work has begun to unpack the relationships between national level gender equality indices and math achievement. But what about the relationship between gender equality indices, gender and expectancy-value constructs? In a meta-analysis of the 2003 TIMSS and PISA data, Else-Quest, Hyde, and Linn (2010) found that although broad measures of gender equality were found to predict greater cross-national variation in math achievement in the PISA 2003 database, they found that global measures of gender equality (i.e., the Gender Empowerment Measure [GEM] and the Gender Gap Index [GGI])[2] were not related to smaller gaps between male and female students for math value, extrinsic and intrinsic motivation, self-concept, self-efficacy and anxiety.

Similarly, recent research by Stoet, Bailey, Moore and Geary (2016) found that amongst the PISA 2003 and 2012 databases, the GGI (a global measure of gender equality) had no positive effect on the size of the gender difference in math anxiety. In fact, Stoet et al. showed that countries with higher levels of gender equality as measured by the GGI exhibited larger gender differences in math anxiety, r = -0.75, p = <.001 for the 2003 dataset, and r = -0.68, p = <.001 for the 2012 dataset.

[2] The GEM measures relative female representation in economics (e.g., gender representation in professional and management positions, and gender gaps in incomes) and politics (e.g., gender representation in parliament) see Klasen, (2006) for further information. The GGI is an index based on economic participation, economic opportunity, political empowerment, educational attainment, and health and wellbeing (see Lopez-Claros & Zahidi, 2005).

Importantly, gender effects for math anxiety were larger than what would be expected from the small differences in achievement alone (*d* = 0.11 in 2003 and *d* = 0.09 in 2012). Stoet et al. also showed that countries with smaller power distances had higher levels of math anxiety, and again this effect remained even when math performance was accounted for. Finally, Stoet et al. demonstrated that there was no correlation between the ratio of mothers working in STEM, and math anxiety or math performance, and that the gender difference for how much girls and boys perceived their parents to value math was larger in developed countries. There was some evidence for a similar pattern with the GGI, but this relied on the exclusion of Scandinavian countries as outliers. Stoet et al. concluded that these results provided strong counter-evidence to the gender stratification hypothesis, and queried the relevancy of national level gender equality in encouraging greater STEM participation. Importantly, a similar pattern of results has been found by Charles and Bradley (2009) and Charles, Harr, Cech, and Hendley (2014), where the gender gap for STEM outcomes increased amongst post-materialist and affluent countries (typically regarded as societies with higher levels of gender egalitarianism).

What can explain these results? When Else-Quest and colleagues (2010) analysed the data from PISA and TIMSS in relation to more specific areas of gender equality at the national level, they found that domain specific measures of gender equality associated with education were related to decreased gender differences in EVT variables. In contrast to global measures of gender equality, domain specific measures such as gender equity in school enrolment, ratio of women in research jobs, and parliamentary representation of women, demonstrated considerable cross-national variation in the size of gender gaps in math affect.

Furthermore, Mann and Di Prete (2016) also found evidence that the seemingly negative effect of high gender equality on gender differences in math attitudes might be more complex than previously anticipated. Using PISA 2006 data, Mann and Di Prete found that although the gender effect for STEM aspirations was larger for countries with high levels of gender equality (as measured by the GGI and GEM), these countries also had high levels of average achievement. Interestingly, high country average levels of achievement were associated with lower country level self-assessments of ability and larger gender differences in self-assessment of ability. Essentially, girls judged their own performance more harshly, compared to boys, as performance environments at the country level became more competitive.

However, Mann and Di Prete found that once this national performance context was accounted for, the gender effect for STEM aspirations was actually smaller in countries with higher levels of gender equality. This indicates that when controlling for achievement environment, national level gender equality is related to smaller gender gaps in attitudes towards STEM. Finally, Mann and Di Prete's research highlights that competitive performance environments may be another possible context related to larger gender differences in self-beliefs and values.

The Intersection of Gender, Social Class and Geography

The above shows that the intersection of gender, ethnicity and cultural difference has sparked the interest of a number of quantitative researchers in recent years. This is not surprising given the strong traditions of intersectional research and feminists who have examined relations between sexism and racism. However, there has been less research on how gender intersects with other important social categories (Eagly et al., 2012). For instance, how does social class affect the experience of being a woman? Or, how does the experience of gender compare across different geographies (e.g., Is the effect of gender on EVT-related variables different for women in high density urban cities, versus women in rural or remote areas)? Thus, in this thesis the interaction effect between gender and social class, and also geography will be investigated.

Social class, gender and math outcomes. There have been only a handful of studies that have included analyses of the interaction between gender and social class with relation to math achievement, and almost none in relation to EVT-related attitudes. Studies have shown that the small gender achievement gap in math becomes larger for American students from higher SES backgrounds, and that this effect occurred from elementary to high school (Lubienski, Crane, and Robinson, 2011; McGraw et al., 2006**)**. Interestingly, McGraw et al. (2006) found that this pattern only extended to white high SES students, with no evidence of the same trend amongst students from other ethnicities. Similarly, Fryer and Levitt (2010) found a larger gender difference in the decline of math performance during adolescence for young girls from the highest quintile of social class, girls with highly educated mothers, and girls who attended private schools.

Why is the gender gap in math performance larger for people in higher socio-economic statuses versus those from working class populations? One reason could be that children from wealthier, resource-rich families are provided with more gender socialisation opportunities. Indeed, there is evidence to suggest that children from higher socio-economic statuses are exposed to more gender specific parenting patterns and extra-curricular activities outside of school, compared to their peers from lower SES families (Lareau, 2003; Lubienski, Robinson, Crane, & Ganley, 2013). For example, research from Sáinz and López-Sáez (2010) on computer attitudes and behaviour showed preliminary evidence for this mechanism using self-report data from Spanish adolescents. In particular, Sáinz and López-Sáez found that the magnitude of the gender gap grew larger in higher SES adolescents, however, this interaction effect did not replicate for the affective measure of computer attitudes used in the study.

Experimental research has shown that teacher perceptions could also play a role in the greater gender differentiation amongst high SES students, showing that when socio-economic status and gender are manipulated, responses from teachers differ according to the gender and social class presented in the vignette (Auwarter & Aruguete, 2008a). The vignette in Aurwarter and Arugete's study featured a student who was struggling at school and failing math despite having an IQ that was not below average. Results showed that teachers were more likely to judge the personal characteristics (e.g., competence in math) of low SES girls favourably compared to their high SES counterparts, who received the harshest evaluations of their personal characteristics out of any demographic group. In contrast, low SES boys received more negative ratings in comparison to high SES boys who received the most favourable ratings for personal characteristics out of any demographic group. Furthermore, high SES girls were the least likely demographic to be referred to academic support or assistance (e.g., math tutoring). Interestingly, the interaction effects between social class and gender did not extend to the responses of school counsellors (Auwarter & Aruguete, 2008b).

Locale, gender and math outcomes. Research on how gender and locale might interact is extremely rare. In a study of gender differences in cognitive abilities amongst Peruvian children, Stevenson, Chen, and Booth (1990) found that for math achievement, there was a trend of an increasing gender gap in urban areas compared to rural areas. For example, young unschooled or first grade girls from the urban area of Lima experienced greater disadvantage relative to unschooled or first grade boys from Lima (d = 0.26). In contrast, gender effects for girls and boys from more rural areas of Peru were (d = 0.05 and 0.09) amongst non-schooled and first grade children. This trend extended to first, second and third grade older children to some degree (d = 0.29 in urban area, compared to (d = 0.15 and -0.07 in rural areas). Sáinz and López-Sáez (2010) showed similar results in the area of computer studies, finding that the gender gap in the behavioural dimension of computing attitudes (i.e., time spent using computers) was larger for urban adolescents, and smaller amongst rural adolescents.

Research Question II: How Much of the Gender Gap in STEM Educational Attainment Can Be Explained by EVT?

The second overarching research question guiding this PhD, is the degree to which EVT, the dominant theory, can explain the gender effect for educational attainment in STEM. Indeed, EVT has provided strong evidence that self-beliefs and values are significantly related to achievement related behaviours and educational attainment. Longitudinal studies have shown that expectancy for success (Guo, Marsh, Morin, Parker, & Kaur, 2015; Marsh, 1991; Parker et al., 2014; Parker, Schoon et al. 2012; Wang, 2013), utility value (Maltese & Tai, 2011), interest (Watt, Eccles, & Durik, 2006), and perceived emotional costs (Perez, Cromley, & Kaplan, 2014) can predict college and senior high school entry for students.

Clearly, there is strong evidence relating expectancies for success and task values as significant predictors of young people’s educational attainment. But to what extent can current theory explain the gender gap in university STEM enrolments? To the author’s knowledge, research has yet to investigate the predictive power of expectancies for success *and* task value combined together in predicting senior high school and university STEM enrolment within the context of gender. By utilising multiple components of task value, this thesis will be able to explore which EVT factors are the most crucial in predicting educational attainment in STEM. Moreover, the amount of time between time waves means that this thesis can explore the extent to which values in middle adolescence can predict outcomes during early adulthood. Furthermore, there is a dearth of research that examines the degree to which EVT-related variables can actually account for the gender effect in STEM. In other words, after achievement and EVT-related attitudes are controlled for, what is the remaining or residual effect of gender on STEM enrolment? Exploring this research question will help educators to better understand the degree to which expectancies for success and values can explain gender differences in course selection decisions, but also open the conversation to explore other possible predictors that may not already be well addressed by current theory.

If EVT-related variables cannot adequately explain the gender effect on educational attainment in STEM, researchers must investigate whether there are alternative mechanisms that could better account for the gender effect. One way to do this is through the utilisation of qualitative research methods. Indeed, the dominance of quantitative correlational studies in EVT literature means there is little understanding of specific individual experiences that occur in the process of engaging or disengaging from STEM education. Most current and past studies focus on comparing levels of expectancy and value beliefs between student groups (e.g., boys and girls), and using expectancy value constructs in statistical models to predict later educational attainment choices and measures of achievement. However, there has been less research that has allowed young people to voice their own opinions about what factors they perceive to be personally significant in their decision to engage or disengage from science. Notably, there has been a call for an increase in qualitative and descriptive analyses of expectancies and value in order to add another dimension to the current literature, and to better describe the meanings young people ascribe to expectancy and value processes (e.g., Eccles, 1994; Tiedemann, 2000). Thus, Study 2b of this thesis will analyse student responses on open-ended interview data in order to uncover any alternative mechanisms to EVT that may help researchers to better understand the issue of gender disparities in educational attainment.

Chapter Summary

This chapter has provided an outline of EVT and has highlighted current gaps in the literature. Namely, a) the need to explore the extent to which gender differences in EVT variables generalise across or are moderated by individual student characteristics such as those emphasised in intersectional research; and b) the need to explore how much of the gender gap in STEM enrolment and attainment can be explained by EVT, and the need for more qualitative research to explain residual gender effects. Consequently, addressing these research gaps is a central aim of the current thesis. Chapter 4 will delve into the various methods that will be employed in order to answer these questions.

CHAPTER 4
OVERVIEW OF THE THESIS: BROAD RESEARCH AIMS AND METHODOLOGIES

Chapters 2 and 3 reviewed the current literature on Expectancy Value Theory, and the under-researched area of gender gaps in academic self-beliefs and values viewed through an intersectional lens. Three critical research gaps in the literature were identified: 1) a lack of understanding of how gender differences in self-beliefs, attitudes, and educational attainment might vary according to social and cultural context – particularly with regards to social class; 2) a lack of research that explores much of the gender gap in STEM enrolment and attainment can be explained by EVT-related variables; and c) the need for more descriptive studies using qualitative data to explain why young women are so poorly represented in STEM.

The overarching aim of this thesis is to address these research gaps by investigating gender differences in academic self-beliefs, attitudes and patterns of educational attainment with a diverse range of methodological tools including meta-analysis, large scale longitudinal data analysis, and analysis of open ended qualitative interview data. Secondly, the thesis aims to answer calls for more quantitative research on the intersection of gender with social and cultural contexts (Eagly et al., 2012; Else-Quest and Hyde, 2016). Extending beyond current intersectional research that is largely dominated by research on the intersection of ethnicity and gender differences and similarities in achievement, this thesis explores how factors like social class, Indigenous status, and geography influence the size of gender effects in academic self-beliefs, attitudes and educational attainment.

Thus, this chapter provides a broad overview of the strengths, limitations and challenges of utilising a multi-method approach with an intersectional perspective to explore the theses' overarching research questions: 1) Do social and cultural context impact on the size of gender effects in self-beliefs, attitudes, and educational attainment; 2) To what extent can current theory explain gender differences in STEM attainment at a university level; and 3) What other factors can explain residual or unexplained gender effects in STEM university attainment? Detailed methodological information will be presented in each individual study chapter within the thesis; as such, this chapter focuses on a general description of the overarching strengths and limitations of the methods used in the thesis; as well as an introduction to the

Longitudinal Study of Australian Youth (LSAY) database that is the focus of Studies 2-4.

Specific Research Questions and Contributions of Studies 1 and 2

Each study of the PhD builds to towards better understanding of: a) the degree to which gender differences in educational attainment in STEM can be explained by current expectancy value theory; and b) the role of social and cultural context in shaping the magnitude of gender differences in academic self-beliefs and attitudes.

Study 1 uses meta-analytic techniques to ask the question, “does social and cultural context affect the size of gender gaps in academic self-beliefs and attitudes?” This study pools gender effects across the current literature to provide overall estimates of gender effects in expectancy value across math, science, and verbal domains. Utilising study characteristics, Study 1 uses moderation analyses to further understand how national level gender inequality, national level gender segregation in university enrolments, social class and belonging to an ethnic minority influences the magnitude of gender gaps in self-beliefs and attitudes. Study 2 attempts to replicate the results of the meta-analysis, controlling for achievement, in a large-scale nationally representative database.

Study 3 explores the degree to which gender differences in STEM educational attainment can be explained by current theory (e.g., EVT and achievement). Additionally, Study 4 investigates whether student responses from open ended interview data can provide any alternative mechanisms behind the gender STEM enrolment gap that are currently overlooked by quantitative literature. Results are discussed in light of the methodological strengths and limitations of each respective method (e.g., meta-analysis, longitudinal analysis, and content analysis). Thus, the next section will provide a brief overview of the strengths and limitations of each method used in this thesis, as well as the benefits of utilising these methods in conjunction with one another.

Best of Both Worlds: Combining the Strengths of Meta-Analysis, Large Scale Secondary Databases, and Qualitative Interviews

Using a variety of methods to answer a research question has several advantages. First, different research methods have different strengths and weaknesses. By using a diverse range of research methods to answer a research question, researchers are able to utilise these different strengths in a complementary way. For example, rather than relying on one method, conclusions made from the research can be bolstered when the various strengths of multiple methods work in unison to counteract weaknesses specific to particular kinds of analyses. Secondly, utilising a variety of methods enables the researcher to approach the research question from new and novel standpoints, as different methods lend themselves to answering different sorts of questions (e.g., some methods are better suited to answering descriptive rather than predictive questions). Thus, the following sections will outline the strengths and limitations of the methods used in this thesis: meta-analysis, large-scale quantitative analysis, and content analysis of qualitative interview data. Finally, the chapter will end by discussing the benefits of utilising these methods in tandem with one another.

Meta-analysis: Strengths and Limitations

Since Gene Glass's popularisation of the meta-analysis (e.g., Glass, 1976; Smith & Glass, 1977), its use as a research tool has significantly increased amongst psychologists. Meta-analyses enable researchers to systematically review, and synthesise numerous findings from the existing empirical literature (Card, 2012). The meta-analysis extends upon traditional narrative reviews by computing a pooled effect size across studies, usually summed up by Cohen's d (1977). Effect sizes are typically interpreted in line with Cohen's recommendations: small effect sizes are $d = < .20$; medium are $d = .50$; and large are $d = .80$, although these recommendations are encouraged as a guide rather than absolute cut-offs for interpreting effect size. Pooling the effect sizes of multiple studies allows for powerful conclusions to be made about the consistency and size of effects in the literature, and as such meta-analyses have become somewhat of a gold standard of research evidence (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Strengths of meta-analytic techniques. Meta-analyses have several strengths that make them a powerful tool for researchers in psychology. Out of any statistical tool, meta-analyses arguably provide the strongest and most robust tests of generalizability and replicability of results – all critical markers of good science (Marsh, Bornmann, Mutz, Daniel, & O'Mara, 2009). Furthermore, meta-analytic techniques allow researchers to examine how study characteristics explain differences across heterogeneous effect sizes in the literature. Marsh et al. (2009) note that even for the highest quality primary databases this is difficult to achieve.

Meta-analyses also extend upon traditional narrative reviews by adding systematic rigour and quantitative precision to assessing the magnitude of effects in the literature (e.g., in this case the size of gender differences in expectancy and success). Traditional narrative reviews often lack transparency in the review processes (e.g., how articles were searched, selected, and interpreted), meaning that subjective biases of the researcher have a greater opportunity to influence the conclusions of the review unchecked (Card, 2012). In contrast, the transparent documentation of meta-analyses search strategies, eligibility criteria, and interpretation of effects means that readers can have greater confidence in the method used to reach conclusions in the review.

Furthermore, Card (2012) notes that the process of consolidating many (often inconsistent) findings, and interpreting these findings appropriately can go beyond the limits of a researcher's information processing abilities. In contrast, the quantitative computation of pooled effect sizes in a meta-analysis allows for a more accurate and precise way of consolidating findings that is not impaired by the limits of human information processing capacities.

There are also several statistical advantages to meta-analyses. Borenstein, Hedges, Higgins, and Rothstein (2009) note that narrative reviews have no reliable mechanism for synthesising *p* values from various studies. Non-significant effects are often interpreted as zero effects in narrative reviews, meaning that reviews can potentially miss important effects. This is problematic because larger effects can often become insignificant when there is not enough statistical power to reach significance. By calculating pooled effects across a large number of studies and participants, meta-analyses can address this by computing effect sizes based on extremely large samples, thereby increasing the statistical power needed to study effects of interest (Walker, Hernandez, & Kattan, 2008). If the goal is to test the null-hypothesis (i.e., that an effect does not exist), then meta-analyses provide a mathematically powerful mechanism to do so, due its strength of statistical power.

Similarly, narrative reviews can often overestimate small effects that are significant, because of an overemphasis on statistical significance and an under-emphasis on looking at the actual magnitude of effects (Borenstein et al., 2009). For example, Hyde (2005) states that a problem with gender research is that it has overinflated the size of gender differences because the magnitude of effect sizes has been neglected as researchers are more likely to pay attention to the statistical significance or difference as opposed to the size. The focus of meta-analyses on the *magnitude* of effect sizes, as opposed to the *significance* of effects, means that meta-analyses provide researchers with a valuable tool to go beyond statistical significance in understanding the scale of difference rather than just the degree to which an effect occurred by chance.

Finally, the systematic and rigorous design of meta-analysis search strategies, study selection, and study evaluation criteria mean that meta-analyses provide researchers with an excellent opportunity to: identify research gaps with confidence; evaluate the current literature; and finally, pose pertinent research questions for future studies to investigate (Noble, 2006).

Limitations of meta-analyses. Despite numerous strengths of meta-analyses, there are also a number of limitations and controversies researchers should be aware of. Indeed, Walker et al. (2008) note that although meta-analyses are powerful tools that can summarise the findings of many studies, they are also controversial because there are several conditions that are critical to a reliable meta-analysis, and violations of these conditions can lead to inaccurate and misleading results.

Meta-analyses can also be plagued by publication bias. Historically journal editors have favoured studies that report large, statistically significant effects, typically because these studies attract greater publicity, excitement and citations than studies that do not find an effect (Walker et al., 2008). For instance, in the field of gender there is likely to be a publication bias towards statistically significant gender differences, while reports of gender similarities are less likely to be accepted for publication. This means that meta-analyses are more likely to be biased towards larger effect sizes, when in reality the overall effect might be much smaller. However, O’Mara (2009) has provided evidence that suggests publication bias might not be a major problem for gender differences in that many studies included in gender differences meta-analyses do not have a primary focus on gender, and that the focus of the study (e.g., gender as a primary theme) did not significantly moderate effect sizes.

Nonetheless, most guidelines of meta-analysis argue that it is important for researchers to take preventative measures to attenuate for the effects of publication bias. Walker et al. (2008) suggest that researchers should include ‘grey literature’ as well as published studies because theses, dissertations and conference papers are more likely to report non-significant findings. Researchers can also include funnel plots in their analyses as a technique of identifying publication bias. Symmetrical funnel plots tend to indicate that there is little evidence of publication bias (as effects are spread evenly), while asymmetrical plots are often a marker of publication bias (Walker et al., 2008). However, researchers should be careful not to over interpret funnel plots as asymmetrical distributions can also reflect high levels of heterogeneity across studies. Nonetheless, funnel plots are an important method for identifying potential problems with publication bias across effect sizes. Thus, the meta-analysis in this thesis includes both unpublished and published studies to counter the effects of publication bias, and provides funnel plots alongside meta-analysis results to help readers draw conclusions about the degree to which data might be affected by such bias.

The inclusion of grey literature, however, is complicated by the fact that unpublished studies might be poorer quality compared to studies in peer-reviewed publications – ‘a garbage in, garbage out’ dilemma. If a meta-analysis includes studies of poor quality, results are likely to be unreliable as studies of poor quality are given equal weighting to studies of high quality in the computation of effect sizes.

Some researchers object to the idea of summarising large amounts of information using a single number (Walker et al., 2008). Indeed, the emphasis on number-crunching in meta-analyses can result in important qualitative differences between studies being missed (Walker et al., 2008). Indeed, a critical component of meta-analysis is testing the study-to-study variation in effect sizes and potential moderators that explain this heterogeneity when it exists (Marsh, et al., 2009). There is also the issue of comparability in meta-analyses. For instance, the meaningfulness of an effect size can be completely obscured if the effects are based on a variable that is too broad reaching and different (e.g., computing effect sizes on math attitudes in general instead of gender differences across different types of math attitudes).

Marsh et al. (2009) have also argued that while meta-analyses provide a strong tool for identifying overall trends in effect sizes, they are weak at testing the generalisability of effects at the individual person level. Although it is possible to test moderation for study-level characteristics, meta-analyses are inherently weak in regards to testing *a priori* predictions of effect sizes in relation to individual level factors, in part because researchers do not have access to data at the individual level of the participant. This means that while meta-analyses can provide descriptive information regarding the size of effect sizes, they are ill suited to testing relationships between individual-level characteristics and effect sizes. Thus, meta-analyses can be limited in that they are constrained by the information that is reported by other researchers.

Large Scale Quantitative Analysis

Strengths of large scale quantitative analysis. Due to the methodological limitations mentioned above, Marsh et al. (2009) argue that meta-analyses and primary analyses should be considered as complementary methodologies, in that both methods help each other in the interpretation of results and identifying new areas of enquiry.

For example, although the meta-analysis provides a robust test of effect sizes across the literature, it cannot account for effects once important variables are controlled for, or held constant (Marsh et al., 2009). A key strength of quantitative database analysis is that most databases allow for the testing of important effects controlling for other critical variables (Marsh et al., 2009). For instance, researchers interested in gender differences in expectancy and value can further the understanding of the role of gender in determining differences, by controlling for achievement at the individual student level so the true effect of gender (once achievement is accounted for) can be revealed. Finally, due to the availability of individual level characteristics, primary analyses can also better test the generalizability of effects across individual characteristics, whereas meta-analysis moderation effects are forced to rely on study level qualities when testing for effects across different groups of people or individual characteristics (Marsh et al., 2009).

Large-scale quantitative databases are often more suitable for evaluating patterns of relationships between effects, while meta-analyses results are more descriptive (Marsh et al., 2009). For instance, this thesis utilises longitudinal data that enables an examination of the degree to which gender differences in expectancy for success and value predict or explain young people’s choice of study at a later date, by assessing the extent to which attitudes and self-beliefs at age 15 can predict patterns of university enrolment. In summary, adding a longitudinal quantitative component to the thesis allows for a test of real world applicability that goes beyond merely describing effects.

Introducing the LSAY database. Studies 2-4 of this thesis are based on the LSAY database, a large nationally representative database that tracks cohorts of Australian youth from the ages of 15 to 25, and is funded by the Australian

Government. The particular focus of LSAY is to better understand young Australians training, work, and social development as they move from the transition between adolescence and adulthood. The LSAY research program has been collecting data on Australian youth since the mid 1970s (then known as the Youth in Transition Program), but for this thesis the 2003 LSAY cohort is the focus of the analyses because of its emphasis on tracking STEM educational attainment at senior high school and university levels of study.

A critical strength of LSAY is its national representativeness. The survey spans across Australia and has been designed to ensure that young people from underrepresented communities have been adequately sampled in the current dataset. Participants were selected through random selection of 50 students from each school from a sample of 355 schools that was designed to represent a diverse range of locations, experiences, and demographics.

The first wave of data in the Y03 2003 cohort is integrated with the OECD's PISA survey. Approximately 12,500 students aged 15 participated in this first wave for PISA, and 10, 370 of these students completed the follow up questions specific to LSAY. In the first wave of data, participants completed assessments of academic achievement and a background questionnaire about family life, educational and vocational plans, and attitudes towards school. Students were followed up in a telephone interview, where further questions about school and work were answered. Subsequent waves of data have been collected annually via telephone interview. In the telephone interview, young people were asked about their school experiences, transitions from school, post-school education, work, health, living arrangements, finances, and general life attitudes. More detailed information on the variables critical to the current study can be found in Study 2.

Strengths and limitations of the LSAY database. LSAY is a publicly available secondary database. There are several advantages and disadvantages to secondary data. Because of their large size and scope, secondary databases have been critical to the field of developmental psychology (e.g., Elder, 1998). Indeed, the size, scope, and complex design of secondary data provides researchers with an excellent opportunity to test questions and hypotheses on data that would simply be impossible to collect and organise as a small team of researchers. There are also ethical reasons for utilising secondary data where possible. Many argue that secondary data is an ethical and effective choice for researchers as it often avoids unnecessary, invasive and expensive collection of new data, and as such secondary data analysis has become a key aim of government funding bodies (e.g., Commonwealth of Australia, 2011; 2012).

However, there are many challenges for researchers who choose to use secondary data. Secondary data analysis can often be hampered by the fact that secondary data is unlikely to align perfectly with the variables needed to test a researcher's hypotheses or research questions. As such, research using secondary data analysis has to be designed around such constraints, sometimes resulting in limitations of the study. For instance, the LSAY 2003 database does not include science self-beliefs and attitudes alongside our longitudinal outcome of university STEM entry, and therefore I was unable to test the role of these beliefs and attitudes in addition to math attitudes in predicting STEM entry. Nonetheless, secondary data provides researchers with a unique opportunity to have access to large-scale data that simply would not be feasible given the time and financial constraints of smaller scale research projects. Rather than relying on a small cross-sectional sample with no long-term follow up, this study is bolstered by strong design, national representativeness, and a powerful test of student attitudes across several years.

Interview Data

Strengths and limitations of qualitative interview data. The final study of the PhD is based on the qualitative interview responses of a subset of young people who participated in the wider LSAY study. Open-ended interview responses provided a valuable opportunity to better understand the phenomenon of young women's underrepresentation in STEM. However, the data presented a number of unique challenges. Firstly, the interview responses given by participants were often brief and to the point, meaning that traditional qualitative analyses requiring deep and rich accounts of the subject matter at hand were unlikely to be an appropriate choice of analysis. Secondly, the size and scale of the sample meant that a method that is able to synthesise and summarise large amounts of information was required. Furthermore, in order to answer the proposed research questions a strategy was required that would enable comparisons between categories according to frequency. While automated text-mining analyses are becoming an increasingly popular tool to analyse large datasets, it was deemed that a frequency count of words without context taken into consideration would not suffice the aims and objectives of the thesis. Finally, due to the nature of the research questions in this thesis, I decided to choose a study that could be adapted to suit both an inductive and deductive theory-driven analysis. To fulfil these requirements, a content analysis methodology was used to analyse the qualitative interview data (this method will be discussed in depth in Chapter 7).

What are the benefits of using qualitative data in conjunction with purely quantitative research? At its core, a mixed method approach allows for multiple ways of seeing, hearing, interpreting and knowing the research problem (Greene, 2007). Quantitative survey research relies on the ability of survey items to elicit the true opinions and feelings of participants; therefore, quantitative results are inherently limited by the fact that what we can know is constrained by the questions survey items ask (Guba & Lincoln, 1994). In contrast, open-ended interview questions that allow for responses to be recorded qualitatively allow for researchers to explore a topic in greater depth, and participants are able to respond unshackled by a predetermined survey scale. Thus, qualitative responses are ripe for providing new and novel insights about a phenomenon that may have not already been considered in quantitative research.

Utilising multiple sources of data can also strengthen the validity of results through convergence or corroboration; or divergence or dissonance (Greene, 2007; Johnson & Onwuegbuzie, 2004; Teddlie & Tashakkori, 2010). Interpreting results from different sources of data allows for better insight and understanding into complex research problems that may otherwise go unnoticed (Teddlie & Tashakkori, 2010).

Chapter Summary

This chapter provided a broad overview of the strengths, limitations and challenges of using a multi-method approach to exploring the theses key research questions. In particular, the strengths and limitations of meta-analysis, large-scale quantitative analysis, interview data, and secondary data analysis were discussed. Overall, by utilising a mix of methods the generalisability and validity of conclusions drawn from results are enhanced, as the strengths of one method compliments the limitations of another. Finally, this chapter provided an introduction to the LSAY database that is the focus of Studies 2-4.

CHAPTER 5
THE INTERSECTION OF GENDER, SOCIAL CLASS, AND CULTURAL CONTEXT: A META-ANALYSIS

Chapter 5 presents the method, results and discussion of Study 1a and 1b: "Do Social Class and Cultural Context Affect The Size of Gender Gaps? A Meta Analytic Review of Gender Differences in Academic Self-Beliefs and Attitudes" (Study 1). Results for Study 1a show that gender differences are highly domain specific, and that there are significant moderation effects across social class, gender equality, and gender segregation in university enrolments. Results from Study 1a are discussed in light of an intersectional research agenda.

Study 1: Do Social Class and Cultural Context Affect the Size of Gender Gaps in Expectancy for Success and Task Value? A Meta-Analytic Review of Gender Differences in Academic Self-Beliefs and Attitudes

Research Aims

The central aim of this study is to address the dearth of research exploring gender differences in relation to social class and cultural contexts, as outlined in Chapter 3. To do this, I utilised a meta-analysis approach to review the current literature which enabled me to analyse average effects of gender across a large number of studies and a diverse range of participants. Hyde (2013) notes that there are few meta-analyses on gender that have tested gender differences across other social categories such as ethnicity and class, making a meta-analytic approach to intersectionality a method ripe for exploring. However, there are a number of challenges for researchers seeking to address these research gaps. For example, Hyde states that a key problem in fulfilling this research agenda is that researchers often fail to include adequate information about the ethnicity and socioeconomic status of their participants, and very rarely report gender differences for different demographics within a study (e.g., gender differences across each category of social class as opposed to averaging gender differences across the whole sample). Thus, researchers' failure to report demographics and their failure to include diverse samples can make it difficult to accumulate enough statistical power to successfully perform moderation analyses in meta-analyses. Thus, this study hopes to not only provide an overview of current EVT literature, but also to explore the compatibility, strengths, and limitations of exploring intersectional research questions with quantitative methods.

Research Questions and Hypotheses

To fulfil the aims of: a) reviewing the current literature on gender differences across expectancy for success and task value for math, science and verbal domains; and b) exploring the compatibility of an intersectional approach with traditional quantitative methods the following research questions and hypotheses were posed:

Research Question 1: From the current literature, what is the overall magnitude of gender differences for EVT variables in mathematics, science, and verbal ability?

Research Question 2: How much heterogeneity (study-to-study variation) exists across effect sizes for gender, and are there any moderators (e.g., publishing date, percentage of sample from an ethnic minority, social class of sample, and gender equality indices) that can explain heterogeneity between studies?

Research Question 3: Does the size of gender differences change across different intersections of social categories? (e.g., What is the effect of gender in middle class versus working class samples?)

Research Question 4: What are the current gaps in the literature on gender differences across expectancy for success and task value variables?

Hypothesis 1: Gender differences will follow a gender stereotypical pattern, whereby math and physical sciences will exhibit the largest gender differences favouring males, while verbal domains and biological sciences will exhibit the largest gender differences favouring females.

Hypothesis 2: Although gender stereotypical patterns will emerge, the pooled effect size for each academic domain will be within the small-moderate range. However, these effects will be larger than the small gender effects on achievement observed in the current literature.

Method

Eligibility criteria

To be included in this review, studies were required to report a quantitative relationship between gender and a measure of domain specific (e.g., math, science) expectancy for success AND at least one measure of domain specific task value. All studies were required to have full-text English results to meet eligibility criteria to ensure that the data extracted were accurate and representative of the study in question. Effect sizes from meta-analyses and other review articles were excluded. There were no restrictions on publication date or type, participant age or any other demographic factors (see Appendix A for a full list of the eligibility criteria used in this study).

Information sources

Searches were conducted within Psychinfo, Education Resources Information Center (ERIC), and Web of Science. Combinations of key words relating EVT terms were used to identify eligible studies in April 2015.

Search

The search strategy aimed to encompass a variety of terms and concepts that tapped into the constructs of expectancy for success and value. Keywords and the full search strategy are available in Appendix B.

Study Selection

All potentially eligible studies were exported into a single Endnote library where duplicate studies were removed. Next, three researchers independently screened titles and abstracts and excluded records where all researchers agreed that titles and abstracts did not meet eligibility criteria. Finally, full-text versions of the remaining articles were obtained and independently screened for eligibility. Discrepancies regarding inclusion were resolved by discussion between the researchers.

Data Collection Process

Three researchers extracted the data from eligible studies. Extracted data included the year of publication, gender split of participants, the domains in which expectancy/value was measured, country in which the study was conducted, socio-economic status of the sample, ethnic minority percentage of the sample, expectancy and value measures used, mean age, and the statistical result that examined the effect of gender on expectancy and value.

Summary Measures

I used standardised mean differences, correlation coefficients, *t* values, and *f* values to calculate the Cohen's *d* effect sizes for each study. All summary measures were converted to Cohen's *d* using Rosenthal's (1991) and (1994) conversion formulas. Effect sizes (*d*) were reported in keeping with Cohen's (1988) general guidelines for interpreting effect sizes; .2 (small), .5 (medium), and .8 (large).

Analysis

Traditional meta-analyses have used fixed and random effects models to analyse data. However, these approaches are limited in that they assume independence (Field, 2003; Marsh et al., 2009), meaning that only one effect size per study can be included in the meta-analysis. Traditional methods of dealing with this (e.g., average effect sizes, or reporting only one effect from a study) are problematic in that they can lose vital information, and limit the testing of moderators (Cheung, 2014).

One way to overcome these challenges is the utilisation of structural equation modelling and multilevel modelling approaches to meta-analysis (Goldstein, 1995; Marsh, Bommann, Mutz, Faniel, & O'Mara, 2009; Raudenbush & Bryk, 1985, Van Den Noorgate & Onghena, 2003). Thus, in this meta-analysis I took a structural equation approach to multi-level meta-analysis. I conducted all analyses in R using the package metaSEM (Cheung, 2011), using unconditional mixed-effects models to calculate overall pooled effect sizes (pooled *d*) and their respective 95% confidence intervals (CIs). Significant effects were reported when the 95% CIs did not cross zero.

To test heterogeneity in pooled effect sizes, I used the I^2 statistic (Higgins, Thompson, Deeks, & Altman, 2003). When effect sizes were heterogeneous (i.e., I^2 was above 25%), moderator analyses were conducted to explore the degree to which study and sample characteristics could explain heterogeneity in the pooled effect size. For each moderation analysis, I reported the proportion of explained variance of heterogeneity that can be accounted for by the inclusion of a moderator variable (R^2), and the heterogeneity between effect sizes in each category (I^2). However, one complication of this was that typically at least 4 effect sizes are required in each moderator sub-category in order to calculate accurate results (Fu et al., 2011). Thus, I only included moderation analyses on variables that had enough data to reach reliable conclusions.

Moderators

Moderators included social class of samples (e.g., working class, middle class, and upper class majority sample in each study), % of participants belonging to an ethnic minority within a study[3], average age of study participants (elementary school, middle school, high school, young adult, and adult), national level gender inequality as measured by the Gender Inequality Index (GII), a ratio of national gender segregation between arts and science graduates (used as a measure of gender equality more closely related to the topic of interest), publication date, population type (e.g., advanced or elective student populations), publication type (e.g., peer-reviewed versus theses) and reliability (>.70 versus <.70). I attempted to collect data at the within study level where possible (e.g., separate effect sizes for gender for a different age group within a study), however, this was dependent on whether studies reported individual effect sizes for subgroups within a study. Further information about the coding of moderator variables is available in Appendix C.

Publication Bias

Funnel plots were examined to assess for publication bias (Sterne, Egger, & Moher, 2011). The *x*-axes of the funnel plot represented the effect size (as measured by Cohen's *d*), and the *y*-axes represented standard errors. Studies with lower standard error (and therefore greater precision) usually sit around the top of the funnel plot, while studies with high standard errors will fall on the bottom of the plot (Higgins & Green, 2011). A symmetrical, inverted funnel-shaped plot is typically associated with low risk of publication bias. In contrast, asymmetrical plots are associated with high levels of publication bias. Plots with greater asymmetry usually indicate a higher risk of bias. High publication bias can mean that meta-analysis effect

---

[3] Note: This meta-analysis originally aimed to analyse effects according to ethnicity as opposed to ethnic minority; however, there were not enough data to perform these analyses. Thus, a limitation of this study is the heterogeneous nature of ethnicities and cultures within an "ethnic minority". Ethnic minority was coded to reflect the percentage of the sample that reported as belonging to an ethnic minority within the country that the study was conducted in. Thus, this category does not reflect one particular ethnicity. Instead, the code reflects the proportion of participants in a study that identified as an ethnicity other than the dominant ethnicity of the country the study was conducted in. Results have been included for the sake of transparency as analyses stemmed from a priori hypotheses. However, readers are advised that results should be interpreted in light of the aforementioned limitation.

sizes are inflated (Egger, 1997; Villar, 1997). The triangular area of the plot represents the region within which 95% of the studies would be expected to fall with the absence of heterogeneity and bias. It is recommended that funnel plots should only be conducted when there are 10 or more effects, as smaller numbers do not provide a strong enough test of asymmetry (Higgins & Green, 2011).

Results

Study Selection

Study selection results are displayed in Figure 2. Through searches of electronic databases and grey literature (e.g., theses, dissertations and conference papers) 6,456 records were identified. After reviewing the titles and abstracts of these 6,102 non-duplicate records, 757 potentially relevant full-text records were obtained and reviewed. After full-text review, 176 studies met inclusion criteria and were included in the meta-analysis.

Study Characteristics

Study characteristics are detailed in supplementary materials. Publication dates ranged from 1966 to 2016. Most of data came from the last 20 years (*ES* <1980s = 15; 1908s = 19; 1990s = 65; 2000s = 80; $\geq$2010 = 76). Participant mean age ranged from 7.04 years to 33.30 years. Studies were categorised according to age group: elementary school age (*n* = number of study clusters*)* = 33), middle school age (*n* = 73), high school age (*n* = 85), young adult (*n* = 54) and adult (*n* = 6).[4]

Most studies were conducted in the United States; however, the review includes studies from Asia, Africa, South America, Oceania and Europe. Countries were classified according to the United Nation's Gender Inequality Index 2014 (GII). Most effect sizes were from countries with either very high gender equality (*ES* = 27), high gender equality (*n* = 60) or medium (*n* = 148) level of gender equality, with a smaller number of effects from countries with low ratings of gender equality (*n* = 13).

During data extraction, information was collected about the social class of the sample. Most studies were described as either lower SES (e.g., majority working class or lower-middle class, *n* = 35) and middle class (*n* = 28). There were only 9 effect

[4] Note: For the sake of clarity, I have reported the number of study clusters with each demographic as opposed to (*k* – number of studies), to allow for clearer reporting on the actual number of effect sizes included in the meta-analysis.

sizes based on high SES samples. The percentage of participants identifying as belonging to an ethnic minority was skewed towards samples that were lower in ethnic diversity. There were 58 effect sizes from samples with less than 25% of participants belonging to an ethnic minority, however, 17 effect sizes came from 25-49% ethnic minority samples, and 36 effect sizes were from samples with a majority of students identifying as belonging to an ethnic minority.

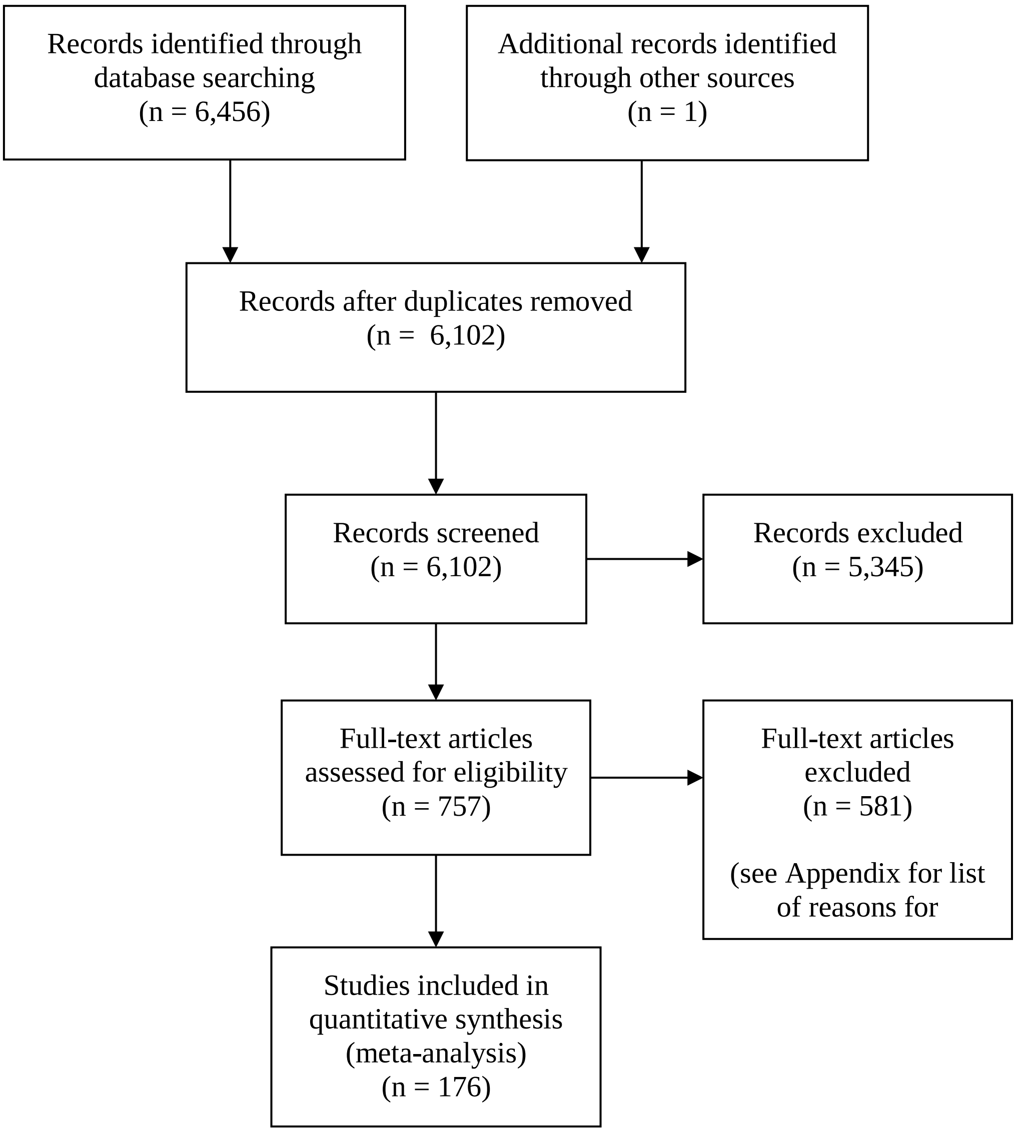

*Figure 2***.** Flow diagram of meta-analysis identification, screening and eligibility, and inclusion processes.

Synthesis of Results

Math expectancy for success and value. In support of Hypothesis 1, boys had higher levels of expectancy and value in math (see Table 1 for full results of pooled effect sizes across all domains). The overall pooled effect of gender on math expectancy for success was *d* = -0.27, 95% CI [-0.31, -0.23], indicating that boys had a small advantage in terms of their perceived ability to do well in math. Effects for the different components of math task value were smaller, but were still in the hypothesised direction. Math task value (*d* = -0.14), 95% CI [-0.21, -0.06], and math intrinsic value (*d* = -0.17), 95% CI [-0.22, -0.12], favoured boys the most out of all the task value components. Whereas, gender differences for math utility value (*d* = -0.08), 95% CI [-0.13, -0.02], math attainment value (*d* = -0.02), 95% CI [-0.15, 0.10], and math cost (*d* = 0.08), 95% CI [-0.05, 0.21], were negligible. Overall, there was a substantial degree of heterogeneity across the effects for gender expectancy/value (I^2 ranging from 0.84-0.93), with the exception of math cost[5].

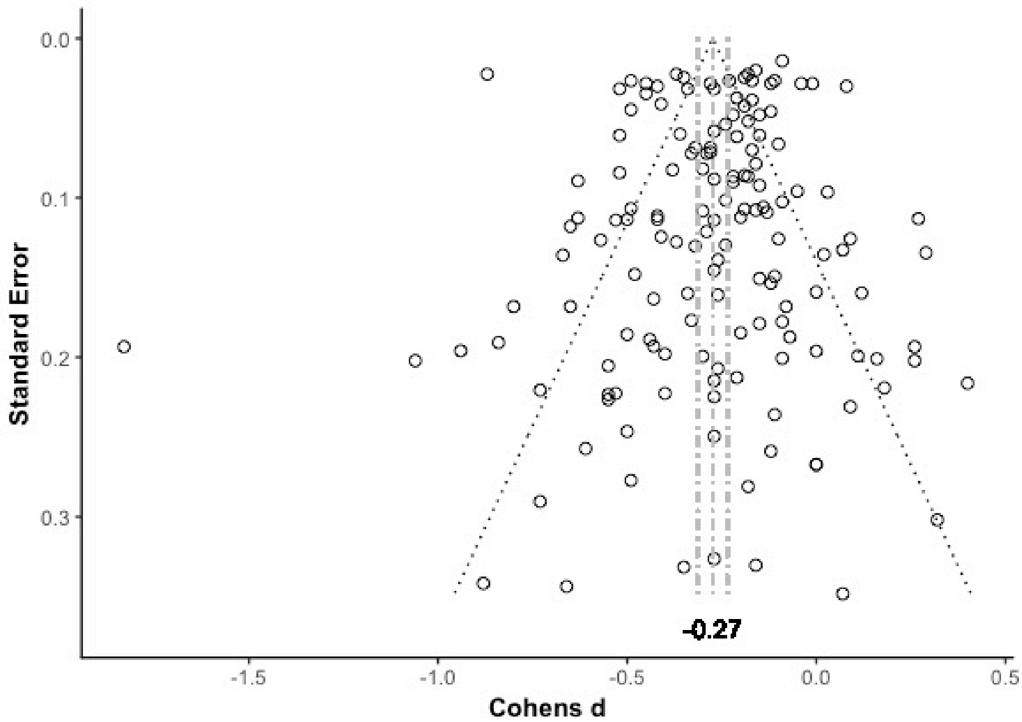

Figure 3a. Math expectancy for success funnel plot.

[5] Heterogeneity estimates for Math Cost were regarded with caution because of the small number of studies in analyses (see von Hippel, 2015 for a discussion of I^2 biases in small meta-analyses.

Table 1

Results of Gender and Expectancy Value Meta-Analyses Across Math, Science, and Verbal Domains

Variable	*ESs*	*D*	*Lower 95% CI*	*Upper 95% CI*	t _2	t _3	I^2_2	I^2_3	*Q Statistic*
Math expectancy for success	150	-0.27	-0.31	-0.23	-	0.04	-	0.92	2046.31
Math task value	44	-0.14	-0.21	-0.06	-	0.05	-	0.93	567.35
Math intrinsic value	79	-0.17	-0.22	-0.12	-	0.04	-	0.91	1020.50
Math utility value	60	-0.08	-0.13	-0.02	-	0.03	-	0.89	563.94
Math attainment value	8	-0.02	-0.15	0.10	-	0.02	-	0.84	23.92
Math cost*	4	0.08	-0.05	0.21	-	0.00	-	0.00	3.22
Science expectancy for success	58	-0.18	-0.26	-0.10	0.02	0.04	0.33	0.60	794.65
Science task value	28	-0.01	-0.08	0.06	-	0.03	-	0.88	424.73
Science intrinsic value	39	-0.21	-0.32	-0.11	0.02	0.04	0.31	0.62	575.22
Science utility value	16	-0.05	-0.12	0.02	-	0.01	-	0.80	141.87
Science attainment value	10	-0.05	-0.13	0.02	-	0.01	-	0.64	30.61
Computing expectancy for success	22	-0.44	-0.60	-0.28	-	0.13	-	0.97	198.76
Computing task value	9	-0.22	-0.38	-0.07	-	0.04	-	0.91	30.44
Computing intrinsic value	14	-0.48	-0.69	-0.26	-	0.13	-	0.97	81.37
Computing utility value	10	-0.21	-0.35	-0.07	-	0.02	-	0.86	21.02

(continued)

Variable	*ESs*	*D*	*Lower 95% CI*	*Upper 95% CI*	t _2	t _3	I^2_2	I^2_3	*Q Statistic*
Engineering expectancy for success*	7	-0.24	-0.40	-0.08	-	0.02	-	0.81	10.54
Engineering intrinsic value*	7	-0.22	-0.32	-0.11	-	0.00	-	0.00	2.03
Engineering utility value*	7	-0.04	-0.22	0.14	-	0.01	-	0.78	18.77
Physical sciences expectancy for success	16	-0.43	-0.56	-0.29	-	0.05	-	0.93	60.00
Physical sciences task value*	3	0.14	-0.38	0.65	-	0.17	-	0.98	13.63
Physical sciences intrinsic value*	8	-0.27	-0.36	-0.19	-	0.00	-	0.25	8.05
Physical sciences utility value*	7	-0.05	-0.21	0.11	-	0.01	-	0.68	10.69
Physical sciences cost*	3	0.32	0.21	0.43	-	0.00	-	0.00	5.04
Biological sciences expectancy for success*	5	0.03	-0.14	0.19	-	0.01	-	0.60	5.92
Biological sciences intrinsic value*	7	0.23	0.06	0.40	-	0.03	-	0.90	26.30
Biological sciences utility value*	4	0.09	-0.12	0.30	-	0.01	-	0.78	5.64
Verbal expectancy for success	65	0.17	0.11	0.23	-	0.04	-	0.92	353.17
Verbal task value	22	0.48	0.34	0.62	-	0.10	-	0.96	315.92
Verbal intrinsic value	35	0.32	0.24	0.40	-	0.04	-	0.90	196.98
Verbal utility value*	7	0.27	0.23	0.31	-	0.00	-	0.00	7.92
Verbal attainment value	9	0.28	0.22	0.35	-	0.00	-	0.42	13.58

Notes. d = Cohen's *d*; *ESs* = number of effect sizes; *CI* = confidence intervals; see text for descriptions of heterogeneity and homogeneity measures. * Note I^2 should be interpreted with caution in small meta-analyses where the number of effects is considered 'small' (e.g., ~ 7 or under). See von Hippel (2015) for a discussion of I^2 biases in small meta-analyses.

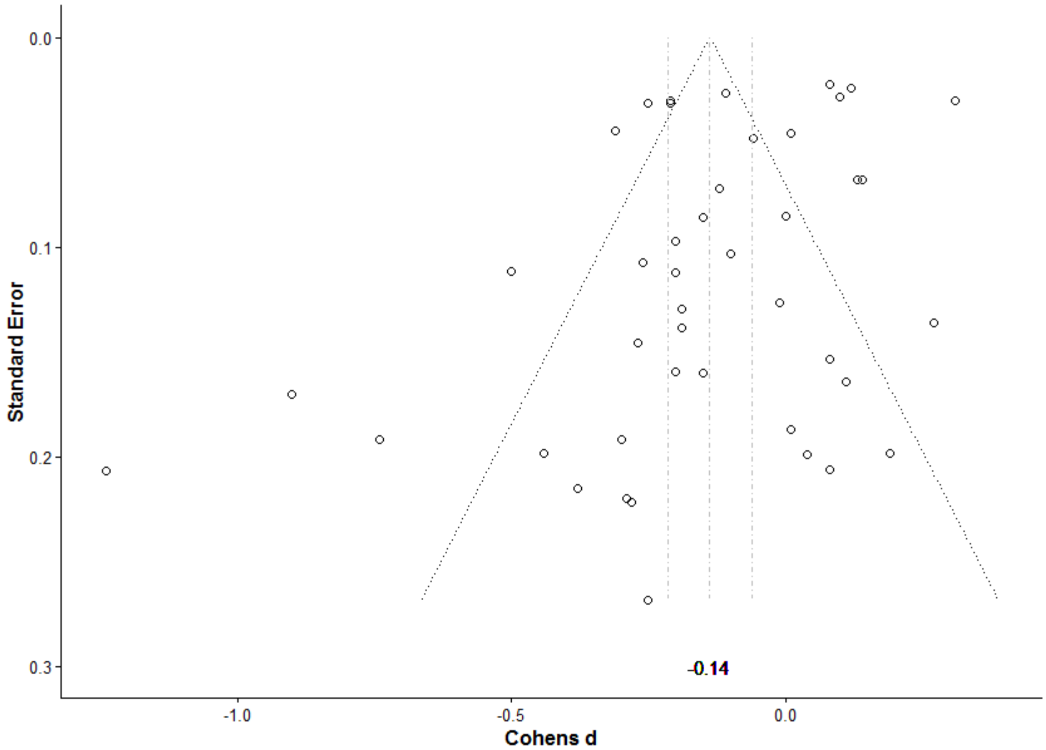

Figure 3b. Math value funnel plot.

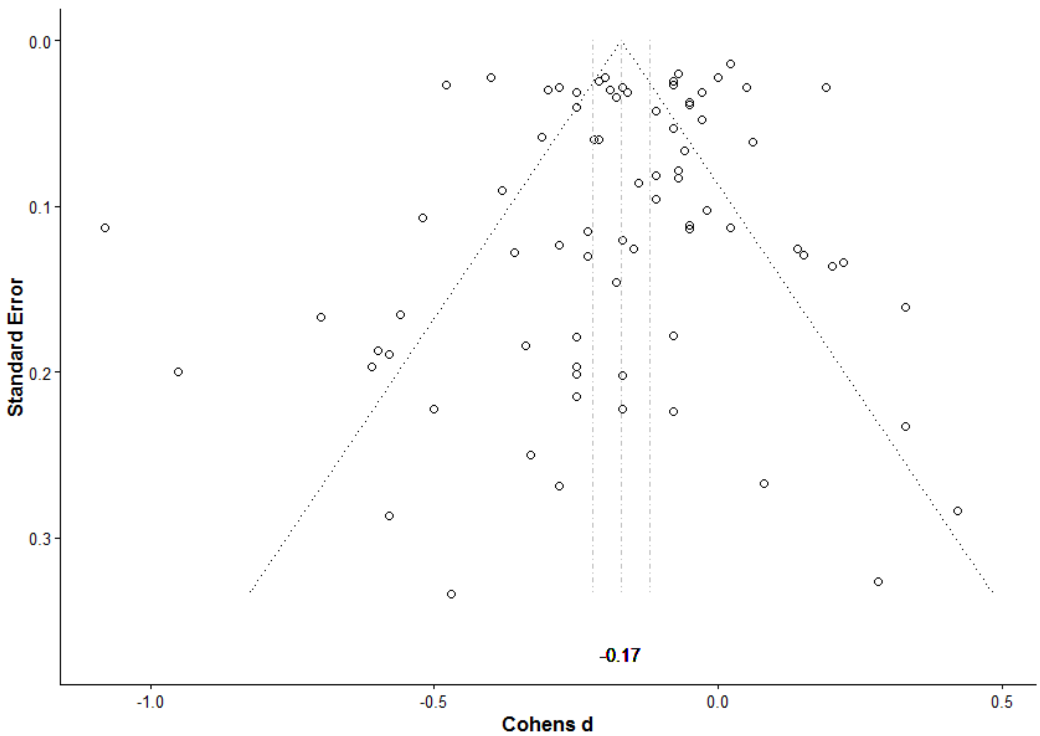

Figure 3c. Math intrinsic value funnel plot.

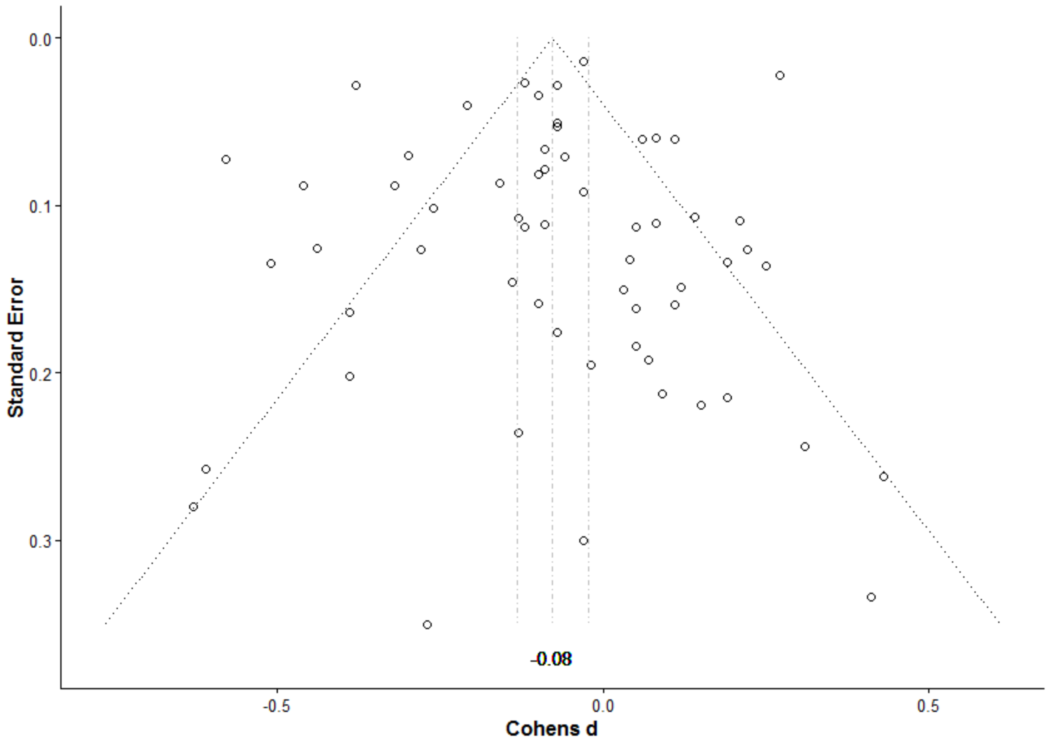

Figure 3d. Math utility value funnel plot.

Science expectancy for success and value. Gender differences were weaker for general science expectancy/value than they were in more precisely defined areas of STEM (e.g., general science versus physical or biological sciences). The overall pooled effect of gender on science expectancy for success was *d* = -0.18, 95% CI [-0.26, -0.10], indicating that boys had a slight advantage in terms of their perceived ability to do well in science. The strongest effect was science intrinsic value *d* = -0.21, 95% CI [-0.32, -0.11], whereby boys were more likely to report higher intrinsic value in science. There was little or no difference between genders in science task value (*d* = -0.01), 95% CI [-0.08, 0.06], science utility value (*d* = -0.05), 95% CI [-0.12, 0.02], and science attainment value (*d* = -0.05), 95% CI [-0.13, 0.02]. Overall, there was a large amount of heterogeneity across effects with I^2 ranging from 0.60 - 0.88. There were not enough studies to provide a meta-analysis on science cost responses, but available effect sizes indicated that there was a small effect of females being more likely to report higher levels of science cost.

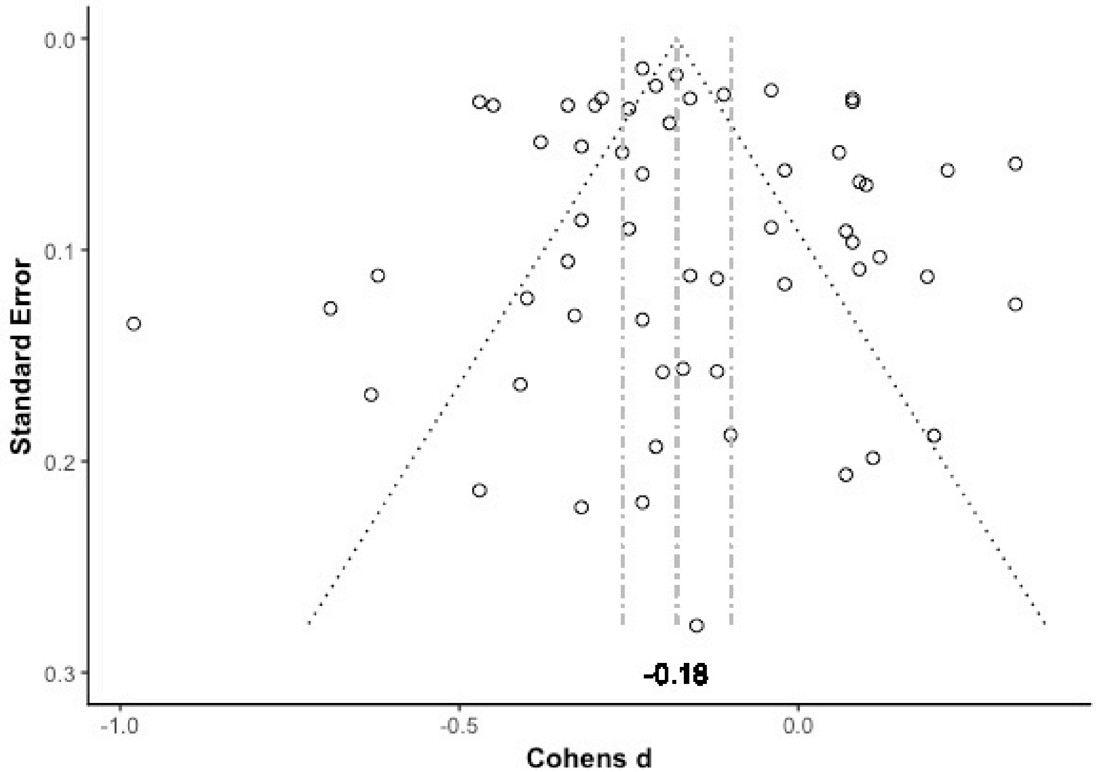

Figure 4a. Science expectancy for success funnel plot.

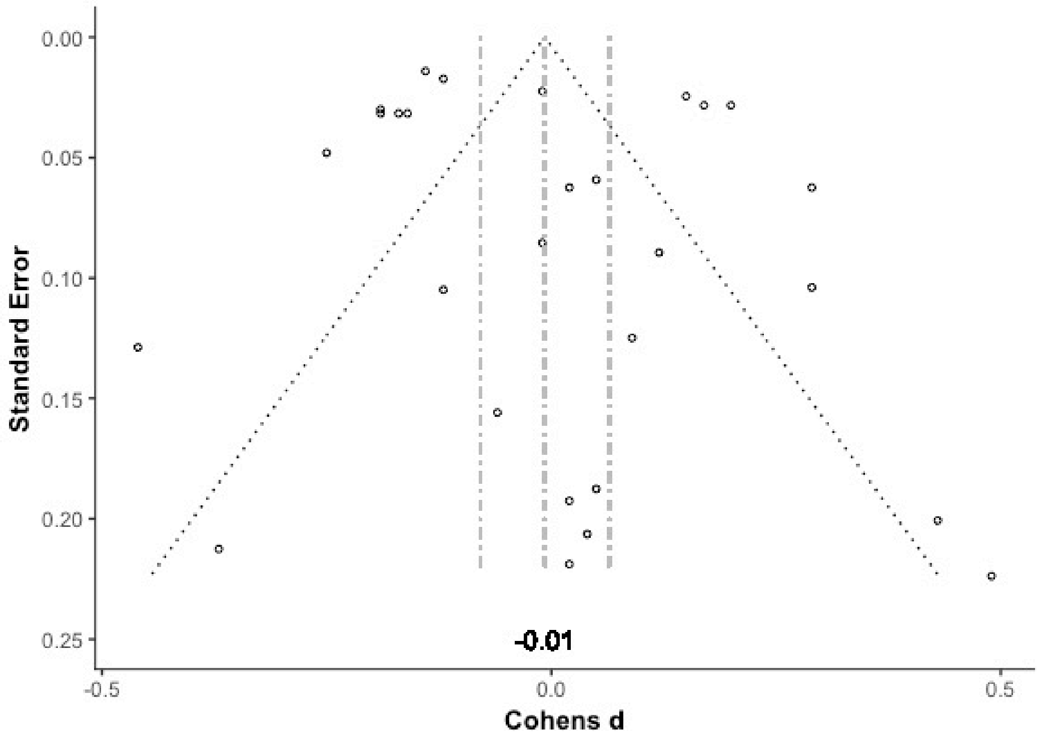

Figure 4b. Science task value funnel plot.

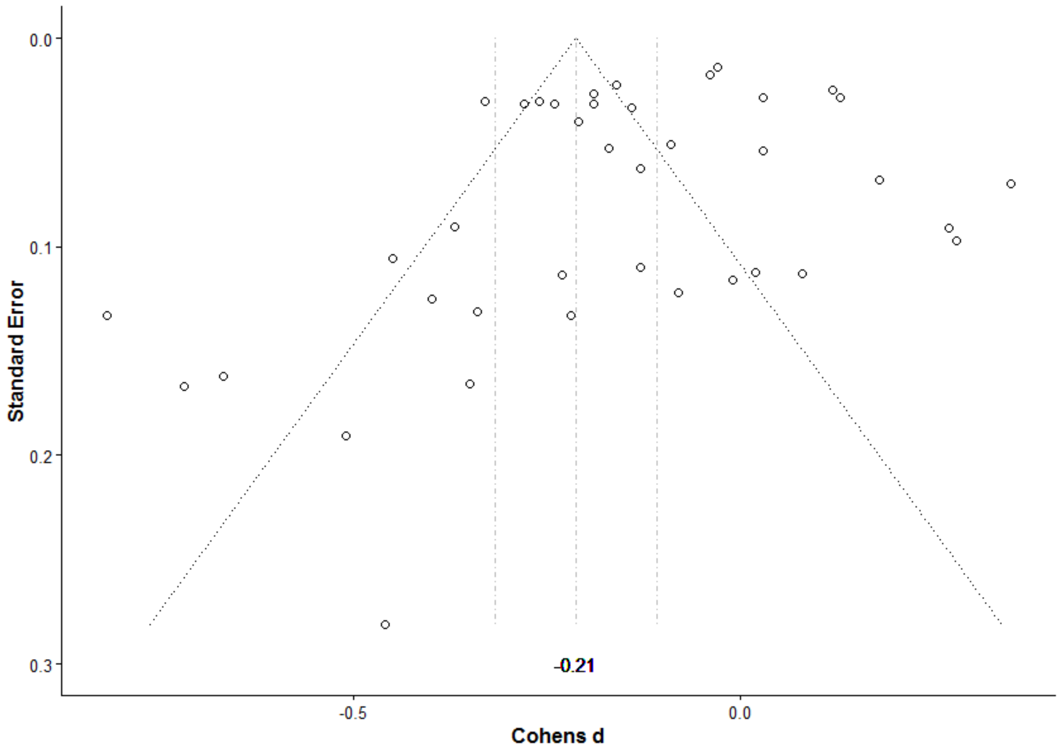

Figure 4c. Science intrinsic value funnel plot.

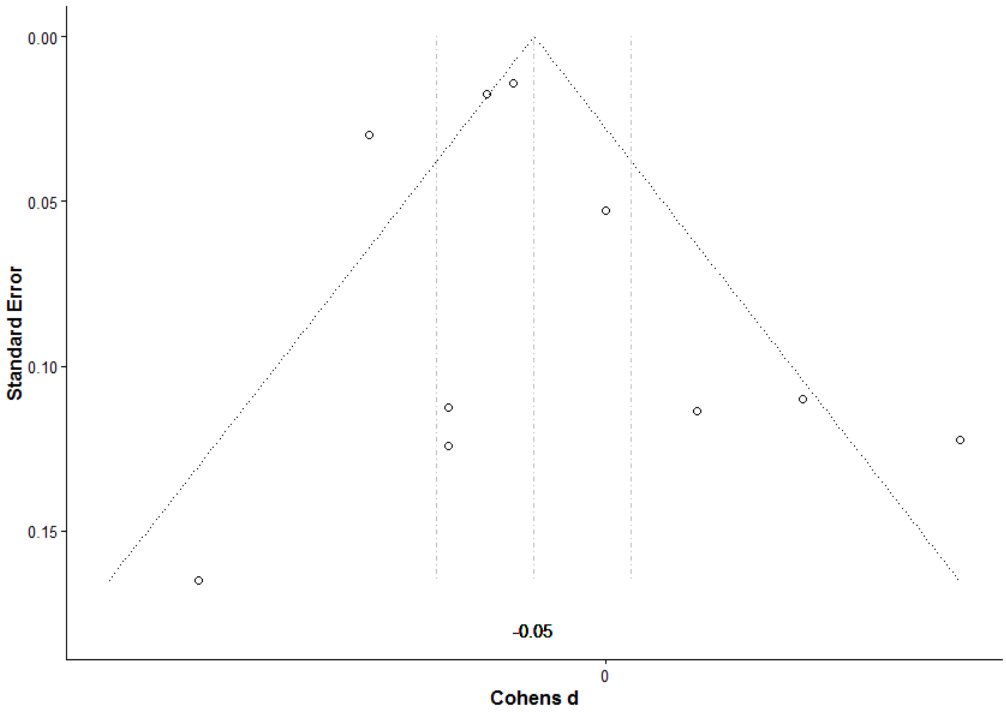

Figure 4d. Science attainment value funnel plot.

Computing expectancy for success and value. In support of Hypothesis 1, effects of gender on computing expectancy for success and value were in favour of males. The overall pooled effect of gender on computing expectancy for success was *d* = -0.44, 95% CI [-0.60, -0.28], demonstrating a medium sized effect favouring males in terms of their perceived ability to do well in computing. In line with the aforementioned results, intrinsic value showed larger gender effects in comparison to other components of task value (*d* = -0.48), 95% CI [-0.69, -0.26]. Effect sizes for computing task value (*d* = -0.22), 95% CI [-0.38, -0.07], and computing utility value (*d* = -0.21), 95% CI [-0.35, -0.07] also revealed a small effect for females being less likely to rate computing as high in task value and career/practical values in comparison to their male peers. There was a large degree of variance across effect sizes with I^2 ranging from 0.86 - 0.97.

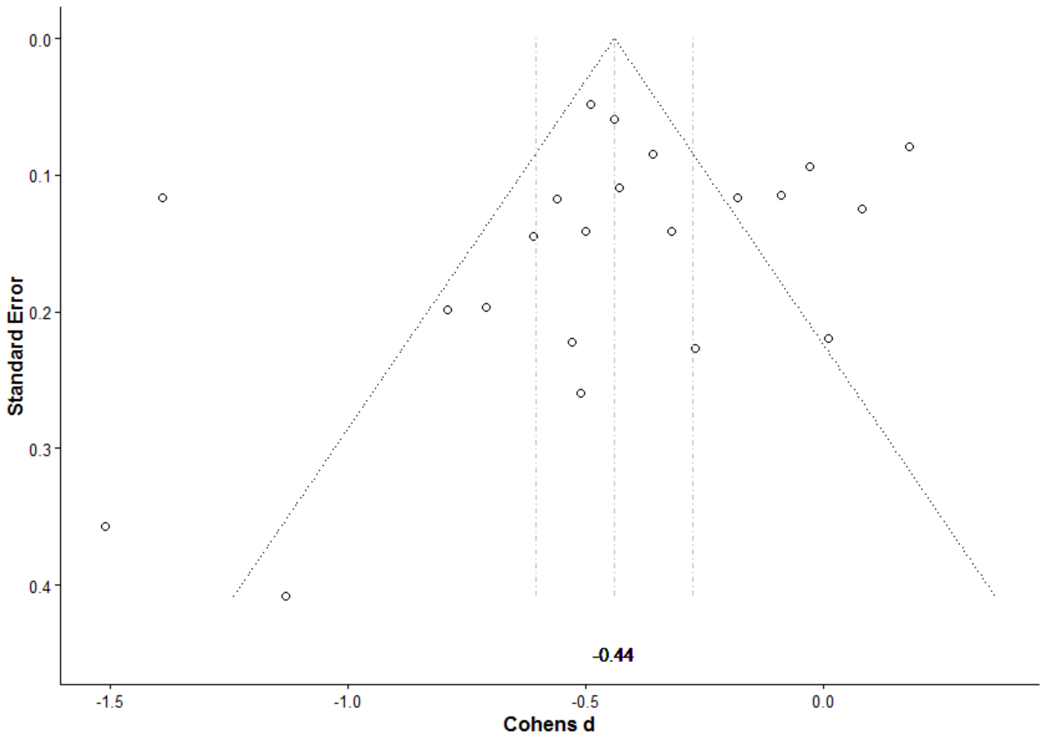

Figure 5a. Computing expectancy for success funnel plot.

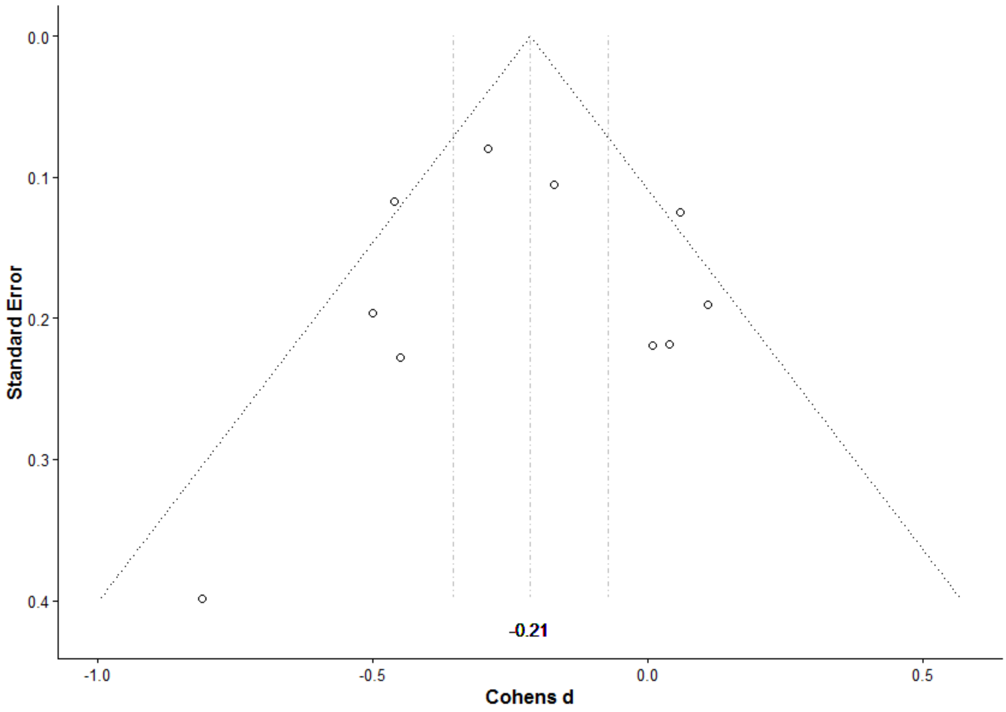

Figure 5b. Computing utility value funnel plot.

Engineering expectancy for success and value. Again, in support of Hypothesis 1, effects of gender on engineering expectancy for success and value were in favour of males. The overall pooled effect of gender on computing expectancy for success was *d* = -0.24, 95% CI [-0.40, -0.08], demonstrating a small effect favouring males in terms of their perceived ability to do well in engineering. Engineering intrinsic value showed a similar effect (*d* = -0.22), 95% CI [-0.32, -0.11], however the effect of gender on engineering utility value was negligible (*d* = -0.04), 95% CI [-0.22, 0.14]. I^2 scores varied considerably, most likely due to the small number of studies on engineering expectancy and value.

Physical sciences expectancy for success and value. The overall pooled effect of gender on expectancy for success in the physical sciences was *d* = -0.43, 95% CI [-0.56, -0.29], demonstrating a medium effect favouring males in terms of their perceived ability to do well in the physical sciences. Again, out of all the task value components intrinsic value showed the largest effect: (*d* = -0.27), 95% CI [-0.36, -0.19]. In keeping with the above results, the effect for utility value was extremely small (*d* = -0.05), 95% CI [-0.21, 0.11]. In contrast to predictions, physical science task value showed a small, but non-significant effect in favour of females (*d* = 0.14; 95% CI [-0.38, 0.66]. Physical science cost was showed a medium effect for gender, with females reporting higher levels of cost for engaging with physical sciences (*d* = 0.32), 95% CI [0.21, 0.43]. I^2 scores varied considerably, but values were interpreted with caution due to the low number of studies including physical science.

Biological sciences expectancy for success and value. In contrast to hypotheses, there was almost no effect of gender on expectancy for success in the biological sciences (*d* = 0.03, 95% CI [-0.14, 0.19]. However, in support of hypotheses there was a positive effect of being female in terms of interest in biological sciences (*d* = 0.23), 95% CI [0.06, 0.40]. Again, there was a smaller non-significant effect for utility value (*d* = 0.09), 95% CI [-0.12, 0.30]. I^2 scores ranged from 0.60 to 0.90 indicating considerable heterogeneity across the effect sizes.

Verbal expectancy for success and value. In support of hypothesis I, effects of gender on verbal expectancy for success and value were in favour of females. The overall pooled effect of gender on verbal expectancy for success was *d* = 0.17, 95% CI [0.11, 0.23], demonstrating a small effect favouring females. Verbal task value showed the largest effect (*d* = 0.48), 95% CI [0.34, 0.62], followed by verbal intrinsic value (*d* = 0.32), 95% CI [0.24, 0.40], In comparison to task value and intrinsic value, utility (*d* = 0.27), 95% CI [0.23, 0.31] and attainment value (*d* = 0.28), 95% CI [0.22, 0.35] had comparatively lower effects. I^2 scores varied considerably for verbal expectancies for success and value, however, values were likely affected by the low number of studies included in analyses.

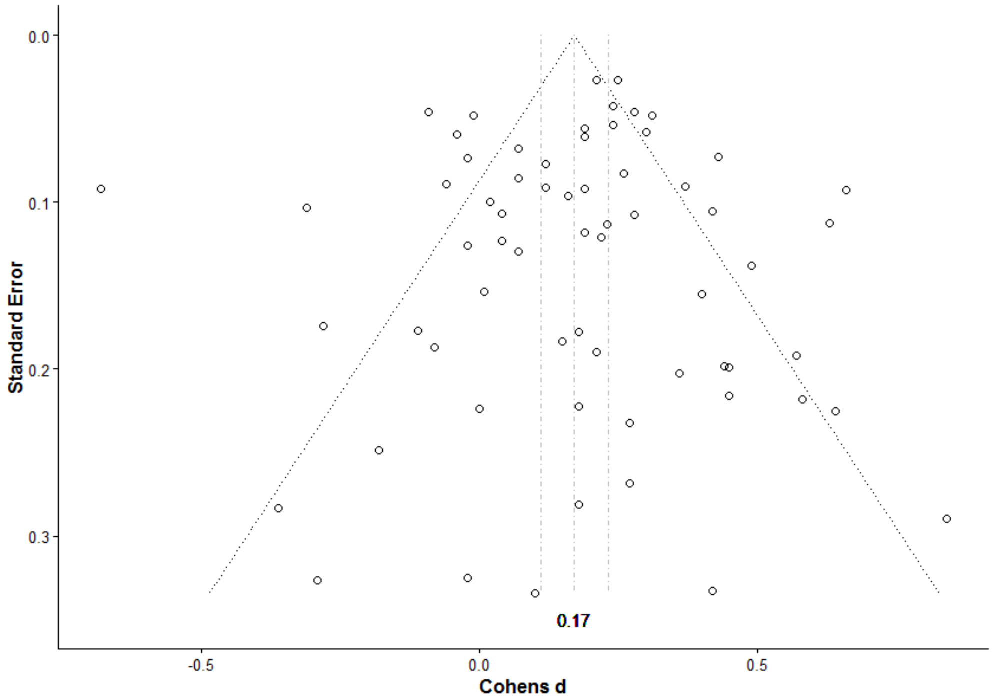

Figure 6a. Verbal expectancy for success funnel plot.

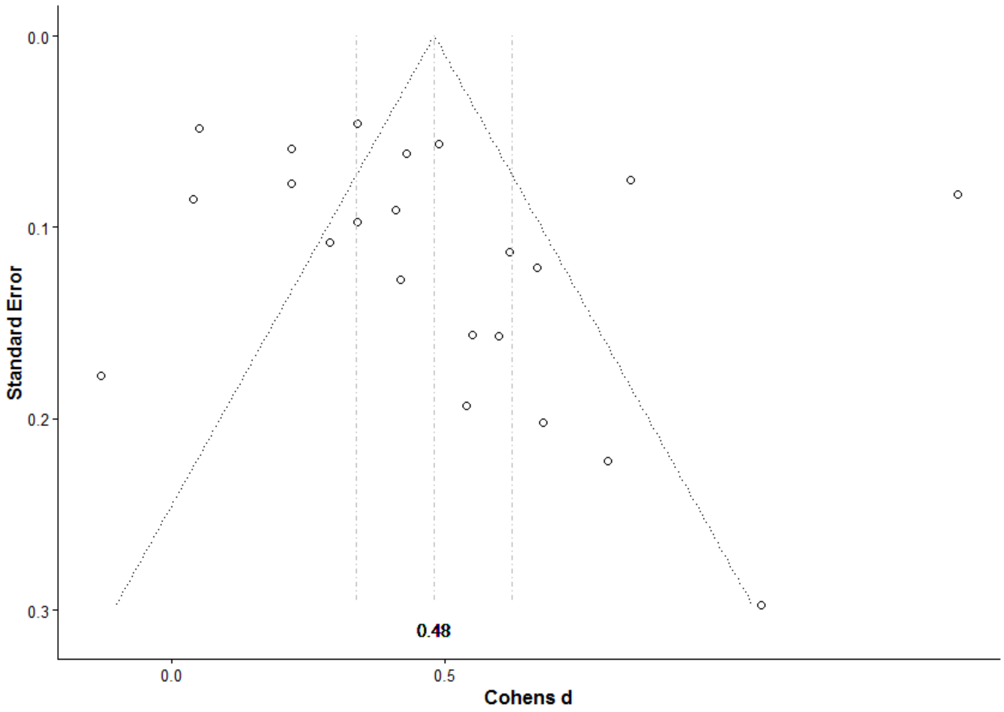

Figure 6b. Verbal value funnel plot.

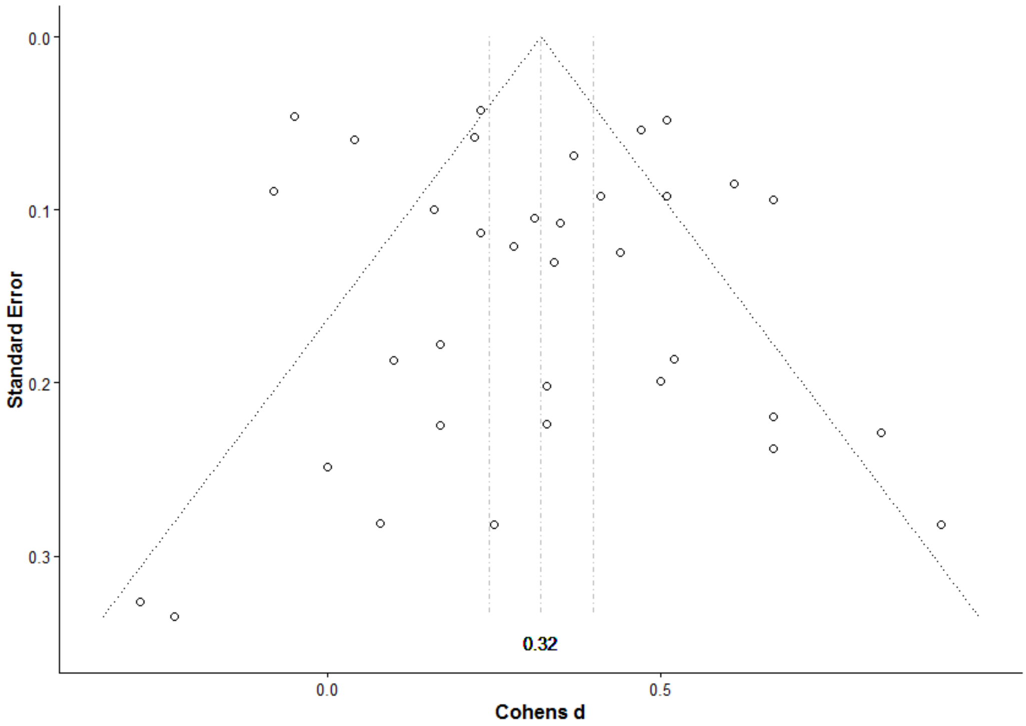

Figure 6c. Verbal intrinsic value funnel plot.

Moderator Analyses

Social class. Gender gaps were largest in high SES samples, and lowest in low SES samples for nearly all math variables. For instance, for math expectancy the gender gap rose from a small effect for majority working class samples: (d = -0.19), 95% CI [-0.28, -0.10], to a slightly larger effect for majority middle class samples (d = -0.25), 95% CI [-0.33, -0.17], and finally, to a large effect for the most affluent samples (d = -0.77), 95% CI [-1.24, -0.31]. Importantly, although confidence intervals in each category overlapped, the moderation effect for social class was statistically significant at p = 0.00. This effect was similar for math task value (p = 0.00); majority working class samples: (d = -0.03), 95% CI [-0.12, 0.06], compared to majority high SES samples: (d = -0.67), 95% CI [-1.00, -0.34]; and also math intrinsic value; majority working class samples: (d = -0.05), 95% CI [-0.11, 0.01], compared to majority middle class samples: (d = -0.08), 95% CI [-0.20, 0.04], and majority high SES samples: (d = -0.47), 95% CI [-0.75, -0.19]. Math utility also showed a trend towards the same direction; majority working class samples: (d = 0.05), 95% CI [-0.04, 0.14], compared to majority middle class samples: (d = -0.10), 95% CI [-0.20, -0.01] . Less data was available to analyse the effect of social class on other domains.

Science expectancy for success showed a similar statistically significant moderation effect for social class; majority working class samples: (*d* = -0.14), 95% CI [-0.27, -0.01], compared to majority middle class samples: (*d* = -0.56), 95% CI [-0.86, -0.27]. Finally, there was no clear pattern for social class across verbal domains, with no significant moderation effects for any of the variables. However, it should be noted that there was substantially less data available to test effects of social class in verbal domains, particularly for the higher end of the SES bracket.

Age. Overall, age was not a significant moderator of gender effects, with there being no significant moderation for gender effects as a function of age for most variables for samples ranging from elementary school age to young adulthood. The one exception was in science expectancy for success (*p* = 0.03), whereby the gender gap grew larger for older samples. In this instance, the pooled effect for gender on science expectancy for success was almost non-existent for samples comprised of elementary school children: (*d* = 0.05), 95% CI [-0.03, 0.14], however, rose to (*d* = -0.31), 95% CI [-0.51, -0.11] for young adult samples.

Publication date (era). Again, for nearly all variables, there was little variation in effect sizes according to the era in which the studies were published, with the exception of math utility and verbal intrinsic value. For math utility value, the gender difference that favoured males in the 1980s (*d* = -0.18), 95% CI [-0.30, -0.06] all but disappeared from the 2000s onwards (*d* = 0.03), 95% CI [-0.05, 0.10] for the 2000s, and (*d* = -0.02), 95% CI [-0.15, 0.10] for 2010 and beyond. Interestingly, this effect was reversed for verbal intrinsic value, whereby gender differences favouring females in the 1990s were larger (*d* = 0.40), 95% CI [-0.30, 0.51], compared to results from 2010 onwards (*d* = 0.21), 95% CI [0.08, 0.34]. Limited data for earlier time periods prior to 2000s meant that it was only possible to test the effect of publication date for a subset of variables.

Ethnic minority percentage of sample. Overall, there was no clear pattern across the proportion of participants who identified as belonging to an ethnic minority. The only exception was for math utility value. For math utility value, samples with the lowest levels of ethnic diversity had larger gender gaps favouring males: (*d* = -0.13), 95% CI [-0.23, -0.03], compared to samples with the highest level of ethnic diversity: and (*d* = 0.16), 95% CI [0.07, 0.24], where the gender effect favoured females.

Sample type. There was little evidence for any moderation effect across sample type (e.g., normal samples versus elective, university or gifted students). The one exception was for computing expectancy for success, whereby students in elective or advanced courses showed larger gender differences (*d* = -0.61), 95% CI [-0.83, -0.39], versus students from 'normal' populations (*d* = -0.23), 95% CI [-0.41, -0.05].

Reliability. There was mixed evidence for gender differences varying across psychometric ratings of reliability. Math expectancy for success and verbal intrinsic value were the only variables that had a significant moderation effect for reliability, however these effects went in opposite directions (e.g., math expectancy for success showed that studies with low reliability were more likely to report smaller gender differences, whilst verbal intrinsic value showed the opposite).

Moderation Analyses of Country Level Indicators

National gender inequality. There was mixed evidence for a moderation effect of national gender inequality (as measured by the GII) across variables. Math intrinsic value exhibited a significant moderation effect, showing that the gender gap in math intrinsic value actually got smaller as gender inequality increased at the national level. For instance, the average pooled effect size for countries with the lowest levels of gender inequality was: (*d* = -0.32), 95% CI [-0.45, -0.18]. In contrast, the average pooled effect for countries with high levels of gender inequality was (*d* = -0.04), 95% CI [0.00, 0.02]. This pattern also occurred for science expectancy for success and science task value, however, there were no clear patterns across other variables.

Gender segregation in educational attainment. Again, there was mixed evidence for a relation between student attitudes and gender segregation in educational attainment (as measured by the difference between national levels of female enrolment in science versus humanities for each country). Science expectancy for success and science task value showed a significant effect, both showing larger gender differences corresponding to higher levels of gender segregation. For instance, countries with high levels of gender segregation across university enrolments had an average pooled effect for science expectancy for success: (*d* = -0.30), 95% CI [-0.37, -0.23], compared to countries with medium levels of gender segregation (*d* = -0.11), 95% CI [-0.20, -0.02], and countries with low levels of gender segregation (*d* = 0.06), 95% CI [-0.15, 0.27]. Verbal intrinsic value also showed a significant effect, whereby boys from countries with higher levels of gender segregation in university enrolments were more likely to have large deficits in interest in verbal areas ($d = 0.52$), 95% CI [0.37, 0.67] and ($d = 0.36$), 95% CI[0.25, 0.46], compared to boys from countries with medium levels of gender segregation in university enrolments ($d = 0.25$), 95% CI [0.10, 0.40]. Nonetheless, the remaining variables did not exhibit similar patterns for gender segregation across university enrolments in the sciences and humanities.

Table 2a

Meta-Analyses and Moderation Analyses for Math

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Math expectancy for success		149	-0.28	-0.32	-0.24	0.04		0.92	2050.30
Social class	***p* = 0.00**						0.32		
Low SES		25	-0.19	-0.28	-0.10	0.02		0.65	56.87
Middle SES		19	-0.25	-0.33	-0.17	0.00		0.15	23.97
High SES		6	-0.77	-1.24	-0.31	0.31		0.94	85.09
Reliability	***p* = 0.01**						0.08		
<.70 or not reported		55	-0.19	-0.24	-0.14	0.02		0.71	215.39
≥.70		93	-0.32	-0.37	-0.26	0.05		0.94	1598.16
Math task value		44	-0.14	-0.21	-0.06	0.05		0.93	567.35
Social class	***p* = 0.00**						0.65		
Low SES		8	-0.03	-0.12	0.06	0.00		0.00	7.29
High SES		5	-0.67	-1.00	-0.34	0.11		0.85	36.13

(continued)

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Math intrinsic value		79	-0.17	-0.22	-0.12	0.04		0.91	1020.50
Social class	***p* = 0.02**						0.38		
Low SES		6	-0.05	-0.11	0.01	0.00		0.00	13.17
Middle SES		17	-0.08	-0.19	0.03	0.02		0.54	26.75
High SES		5	-0.47	-0.75	-0.19	0.08		0.80	23.22
National Gender Equality	***p* = 0.03**						0.20		
Highest level of gender equality		13	-0.32	-0.45	-0.18	0.05		0.94	181.78
High level of gender equality		19	-0.17	-0.23	-0.11	0.01		0.64	51.93
Medium level of gender equality		39	-0.13	-0.21	-0.05	0.05		0.91	522.59
Low level of gender equality		5	-0.04	-0.11	0.02	0.00		0.46	14.23
Math utility value		60	-0.08	-0.13	-0.02	0.03		0.89	567.23
Social class	***p* = 0.03**						1.00		
Low SES		14	0.05	-0.04	0.14	0.00		0.00	13.77
Middle SES		4	-0.10	-0.20	-0.01	0.00		0.00	2.28
Era	***p* = 0.02**						0.29		
< 1980s		14	-0.18	-0.30	-0.06	0.04		0.72	49.09
1980s		8	-0.20	-0.37	-0.03	0.04		0.87	47.84
1990s		17	-0.09	-0.19	0.02	0.02		0.72	95.84
2000s		14	0.03	-0.05	0.10	0.01		0.86	174.28
Current		7	-0.02	-0.15	0.10	0.02		0.77	23.43

(continued)

% Ethnic minority	***p* = 0.03**						0.33		
Lowest		21	-0.13	-0.23	-0.03	0.04		0.85	67.24
Low		4	0.09	-0.07	0.25	0.02		0.91	128.71
Highest		9	0.16	0.07	0.24	0.00		0.00	12.74

Note. Non-significant results are available in Supplementary Materials. Math utility confidence interval for social class was calculated using a fixed level meta-analysis.

Table 2b

Moderation Analyses for Science

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Science expectancy for success		58	-0.18	-0.26	-0.10	0.04		0.60	794.65
Social class	***p* = 0.01**						0.64		
Low SES		7	-0.14	-0.27	-0.01	0.00		0.13	6.41
Middle SES		4	-0.56	-0.86	-0.27	0.07		0.83	26.44
Age	***p* = 0.03**						0.17		
Elementary age		4	0.05	-0.03	0.14	0.00		0.00	2.51
Middle school age		22	-0.09	-0.18	0.00	0.04		0.92	455.49
High school age		25	-0.24	-0.33	-0.14	0.05		0.95	184.42
Young adult		5	-0.31	-0.51	-0.11	0.04		0.81	47.01
National Gender Equality	***p* = 0.00**						0.27		
High level of gender equality		17	-0.24	-0.36	-0.12	0.06		0.94	190.19
Medium level of gender equality		32	-0.15	-0.23	-0.08	0.03		0.86	190.22
Low level of gender equality		6	0.11	0.00	0.22	0.01		0.82	44.33
Gender Segregation in Graduates	***p* = 0.01**						0.33		
Low levels of gender segregation		4	0.06	-0.15	0.27	0.04		0.91	73.45
Medium levels of gender segregation		28	-0.11	-0.20	-0.02	0.04		0.90	186.55
High levels of gender segregation		6	-0.30	-0.37	-0.23	0.01		0.72	66.48

(continued)

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Science task value		28	-0.01	-0.08	0.06	0.03		0.88	424.73
National Gender Equality	***p* = 0.00**						0.72		
High level of gender equality		7	-0.14	-0.16	-0.12	0.00		0.00	13.51
Medium level of gender equality		14	0.02	-0.07	0.12	0.02		0.78	104.72
Low level of gender equality		6	0.16	0.13	0.19	0.00		0.00	9.84
Gender Segregation in Graduates	***p* = 0.01**						0.52		
Medium levels of gender segregation		15	0.05	-0.04	0.14	0.02		0.81	100.20
High levels of gender segregation		4	-0.15	-0.18	-0.12	0.00		0.10	4.96

Note. Non-significant results are available in Supplementary Materials. Science value confidence interval for high gender equality was calculated using a fixed level meta-analysis.

Table 2c

Moderation Analyses for Computing Sciences

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Computing expectancy for success		22	-0.44	-0.60	-0.28	0.13		0.97	198.76
Sample Type	***p* = 0.02**						0.27		
Special sample (e.g., elective STEM)		13	-0.61	-0.83	-0.39	0.12		0.92	82.58
Standard sample		9	-0.23	-0.41	-0.05	0.06		0.95	84.51

Note. Non-significant results are available in Supplementary Materials.

Table 2d

Moderation Analyses for Verbal

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Verbal intrinsic value		35	0.32	0.24	0.40	0.04		0.90	196.98
Era	$\boldsymbol{p = 0.02}$						0.33		
1990s		18	0.40	0.30	0.51	0.01		0.65	34.99
2000s		7	0.43	0.36	0.51	0.00		0.00	6.66
Current		10	0.21	0.08	0.34	0.04		0.88	76.50
Gender Segregation in Graduates	$\boldsymbol{p = 0.04}$						0.65		
Medium levels of gender segregation		7	0.25	0.10	0.40	0.01		0.78	14.68
High levels of gender segregation		6	0.52	0.37	0.67	0.01		0.80	9.42
Highest levels of gender segregation		9	0.36	0.25	0.46	0.00		0.00	8.80
Reliability	$\boldsymbol{p = 0.01}$						0.37		
<.70 or not reported		25	0.39	0.31	0.47	0.02		0.74	65.84
≥.70		10	0.19	0.07	0.31	0.03		0.88	47.67

Note. Non-significant results are available in Supplementary Materials. Verbal intrinsic value confidence interval for 2000s was calculated using a fixed level meta-analysis.

Discussion

This study used a meta-analytic approach to examine gender differences and similarities across a wide range of EVT constructs in a variety of different domains. Additionally, the meta-analysis provided an opportunity to synthesise current literature through an intersectional lens, by exploring the degree to which gender differences varied across a number of cultural and social contexts. These findings shed light on a number of issues pertinent to the study of gender and attitudes towards math, science, and verbal domains. As expected, gender differences followed a gender stereotypical pattern across academic domains. Finally, there is emerging evidence to suggest that there are important differences and similarities across social and cultural contexts. Below are the most important findings from the analyses and review.

Gender Differences and Similarities Across EVT Constructs

Research Question 1 asked what the overall magnitude of gender differences for EVT variables was for math, science, and verbal ability. Pooled effect sizes revealed that there were significant gender differences in EVT related constructs across multiple domains. Funnel plots revealed that effect sizes were largely symmetrical, indicating a lower risk of publication bias. Supporting Hypothesis 2, 17 out of the 31 pooled effect sizes in this study fell within the small range (approx. $d = .20$). Four effects were medium in size (approx. $d = .50$). These larger effects were for computing expectancy for success and intrinsic value, physical sciences expectancy for success (favouring boys) and verbal task value (favouring girls). However, in contrast to a priori hypotheses, 10 out of the 31 effects were near zero ($< d = .10$). Gender differences failed to reach statistical significance for math attainment value and cost, science task value, utility value, and attainment value, physical sciences task value and utility value, and biological sciences expectancy for success and utility value. Thus, there are important gender similarities and differences across EVT constructs.

A key finding was that effect sizes for gender differences in academic attitudes were highly domain specific. That is, there was a great deal of variation in effect sizes spanning across math, science, and verbal domains. However, more importantly there were critical differences in the magnitude of gender effects within sub-disciplines of science, a domain that is often measured in general terms. For instance, gender differences in general science ranged from small effects favouring males to zero effect sizes. In contrast, gender differences in the physical sciences and computing were comparatively much larger, and small differences favouring females were evident in the biological sciences. Thus, researchers wanting to research the "STEM gender gap" should consider domain specific measures relevant to their specific target areas, as there is substantial variation across the different sub-disciplines of science that seem to counterbalance one another when measuring general science. In support of Hypothesis 1, results are consistent with patterns of educational attainment in STEM showing the gender gaps to be largest in the "hard" sciences (e.g., physics, math, engineering, computing), and smallest in the biological sciences (National Center for Education Statistics, 2015). Essentially, the size of the "STEM gender gap" largely depends on how STEM is defined. Thus, interventions to reduce the gender gap in STEM might be better placed to focus specifically on improving self-beliefs and attitudes in the physical sciences and computing as opposed to biological sciences, where there are more similarities than differences between male and female student attitudes.

Moderation Effects for Social and Cultural Contexts

Gender and social class. Arguably the most interesting findings were found within from the moderation analyses. Social class showed a number of significant effects across math related variables and science expectancy for success, whereby the gender gap favouring males increased as social class became higher. This finding reflected the results of Fryer and Levitt (2010), Lubienski, Crane, and Robinson (2011) and McGraw et al. (2006) who found that the gender achievement gap in math for elementary and high school students is larger amongst high SES students compared to low SES students.

Why is there such a large gender gap for people in higher socio-economic statuses versus those from majority working class populations? Due to the inherent weaknesses of a meta-analysis approach to moderation analyses such as limited moderator data at the individual level (see Marsh et al., 2009 for a discussion), results should be interpreted with caution. Nonetheless, the corroboration of these results with previous data suggests an underlying effect that warrants further investigation using large-scale primary data.

A potential hypothesis behind the gender/social class relationship could be that children from wealthier, resource-rich families are provided with more gender socialisation opportunities. Indeed, there is evidence to suggest that children from higher socio-economic statuses are exposed to more gender specific parenting patterns and extra-curricular activities outside of school, compared to their peers from lower SES families (Lareau, 2003). An unanticipated negative effect of this is that exposure to gender stereotypical activities results in greater gender stereotypic differentiation in the self-beliefs and attitudes of children. As children gain more experience in gender-congruent activities, and less experience and familiarity with tasks and activities incongruent with their gender, their self-beliefs and attitudes consequently become confined to a gender stereotypical pattern reflecting their exposure (or lack of) to different experiences. Interestingly, this pattern occurs despite the fact that parents from high SES are more likely to claim to hold gender egalitarian views (Marks, Lam, & McHale, 2009).

Another alternative explanation is that high SES environments often have higher levels of average achievement and this can translate into a more competitive environment for students. A growing body of research has shown that on average women respond less favourably to competitive environments than men (e.g., Bönte 2015; Gneezy, Niederle, & Rustichini, 2003; Niederle & Vesterlund 2007, 2010). Furthermore, in the context of education, Alon and DiPrete (2015) showed that the intensity of competition, as signaled by admission standards into university STEM courses, had a larger deterring effect on female applicants compared to males.

It is, however, intriguing that the same pattern did not extend to science and verbal domains. One reason for this is that there was limited data available to test social class moderation effects, particularly in scientific and verbal domains. Further research should endeavor to include social class in participant demographics, and to recruit a wider range of participants outside of the middle class demographics that has traditionally been oversampled in psychological research.

Gender and ethnicity. I explored the relationship between gender and ethnicity by comparing the size of gender differences across samples with differing levels of ethnic diversity. Importantly, there were more similarities than differences in effect sizes amongst samples of high ethnic diversity and samples of low ethnic diversity, with the exception of some differences showing that participants from samples with low ethnic diversity were more likely to have gender differences favouring males in math utility value. These low ethnic diversity samples were primarily white. Overall, results reflect previous research that has provided mixed evidence regarding the interaction between gender and ethnicity. While some research has suggested that gender differences in attitudes and achievement become larger amongst white students (e.g., Catsambis, 1994; 1995; Cooley, 2001; NSF, Division of Science Resource Statistics, 2012; McGraw et al., 2006; Zarrett et al., 2006), other studies have not replicated this effect (e.g., Else-Quest et al., 2013; Zarrett et al., 2006). Importantly, these results are limited by the fact that many different ethnicities may exist within the category of ethnic minority, and therefore, important differences may be overlooked by the use of such a broad category.

Moderation Effects for Country Level Indicators

Gender inequality. There was mixed evidence with regards to the role of gender inequality, as measured by the GII, in moderating the size of gender effects. Some moderation analyses indicated that for science variables and math intrinsic value, lower levels of gender equality resulted in larger gender differences in attitudes, replicating the results of previous studies that have used other global measures of gender equality (e.g., Charles & Bradley, 2009; Charles et al. 2014; Else-Quest et al. 2010; Stoet et al., 2016). What can explain the relationship between gender equality and attitudes? On the surface, it might be tempting to associate 'gender equality' with poorer outcomes for girls in terms of self-beliefs and attitudes. The relationship, however, is more complex than this. As previously discussed in Chapter 3, recent work by Mann and DiPrete (2016) has suggested that the negative effect of gender equality on gender differences disappears once the country level achievement is included in statistical models. Hence, these results reflect the pattern seen in social class, whereby affluent resource-rich and highly competitive contexts exacerbate gender differences in attitudes towards STEM.

Gender segregation in educational attainment. In line with the suggestions of Else-Quest and Grabe (2012), domain specific indicators of gender inequality may provide different insights into gender effects in attitudes than global measures. Arguably, using a measure specific to gender segregation across academic pursuits is better suited for research on gender differences in attitudes, as it is more closely related than broad gender equality. While results in this study were mixed, there were some significant moderation effects that were in contrast to results from the global GII measure. In particular, gender differences for science task value, science expectancy for success, and verbal intrinsic value were associated with the ratio of women enrolled in STEM university courses compared to university courses in the humanities. As there was increased gender segregation across university courses, effect sizes favoured males in science and girls in verbal domains. Interestingly, this pattern did not emerge for math. Nonetheless, the use of measures of gender segregation across university enrollments remains a promising area for further research, and connecting individual attitudes to enrolment patterns at a national level.

Moderation Effects for Study Characteristics

Age. Another key finding was that the effect of gender was unrelated to average participant age, with the exception of science expectancy for success. Again, this finding is surprising, given that the age range of participants in the meta-analysis ranged from 7 to 33 years. One conclusion of this finding could be that gender socialisation has a limited role in determining the degree to which self-beliefs and attitudes are differentiated according to gender. Indeed, the gender stratification hypothesis maintains that gender differences should become larger across development because of greater exposure to gender socialisation. An alternative to this explanation could be that gender socialisation is so heavily entrenched during early childhood, that gender differences emerge at ages below what this study examined. Regardless, these findings show that gender stereotypical patterns in self-beliefs and attitudes are heavily entrenched even in participants who are still in the early childhood phase of development. One gap in the literature is that there is an over-representation of school age children, but limited attention to children below school age. Given that gender differences are already established amongst elementary aged children, researchers need to start to focus on the emergence of gender differences in self-beliefs and attitudes in children before they engage in formal schooling. While there are likely many methodological challenges in working with such a young population, it seems as though this work is critical to furthering our understanding of gender differences in self-beliefs and attitudes (see Marsh, Ellis, & Craven, 2002 for a discussion about the self-concept measurement of preschoolers).

Publication date (era). Another way of exploring the role of cultural context in relation to gender differences in attitudes and beliefs is assessing the relationship between publication date and effect sizes. For nearly all variables there was no statistically significant effect for publication date. The only exception was for math utility value , whereby effect sizes have diminished since the 1980s to almost zero difference in the current day. A similar pattern emerged for verbal intrinsic value, whereby gender differences favouring girls decreased somewhat from the 1990s to the current day. The lack of change in gender differences in attitudes for other variables paints a potentially damning picture of gender equality in educational attitudes, but I am cautious to over-interpret these findings. Firstly, there was limited data available for the time period before the 1980s, with most studies being published within the last 15-20 years. Thus, this meta-analysis was unable to provide conclusions of whether the gender gap in self beliefs and attitudes has changed since much earlier decades in the twentieth century. Nonetheless, it is concerning that there has been little change since the 1980s for most variables. Indeed, this is in line with recent research that has shown the persistence of gender stereotypes across time, showing that people's perceptions of gender stereotypes have been largely stagnant and resistant to change since the 1980s (e.g., Haines, Deaux, & Lofaro, 2016).

Sample type. Moderation effects across sample types also revealed some critical insights into contextual factors and gender differences. Analyses showed that for nearly every variable there was almost no difference between 'normal' samples of students versus samples drawn from students taking advanced or elective courses. Nonetheless, there was one significant effect for engineering expectancy for success that showed that the gender effect (favouring males) was largest in elective/advanced samples. These findings are concerning as it shows that even for female students who are motivated and gifted enough to engage with STEM studies at a higher level, there are still substantial challenges they face in terms of lower self-beliefs and attitudes compared to their male classmates. Thus, educators should be aware that even amongst high-achieving and highly motivated populations, female students still have lower confidence and poorer attitudes towards STEM relative to their male peers.

Boys and verbal domains. Finally, this meta-analysis also highlights the 'flip side of the coin' with regards to gender differences in self-beliefs and attitudes. Analyses showed that although boys have advantages in STEM fields, they also have lower expectancy for success and task value in verbal domains. Thus, educational policies that discuss gender equality in education need to also recognise that boys have lower self-beliefs and attitudes in verbal domains. If we are to encourage truly gender equal educational settings we need to view gender equality in education holistically, considering both verbal and STEM domains when discussing differences in self-beliefs and attitudes so all children develop perceptions of their abilities commensurate with their achievement.

Conclusions, Limitations and Directions for Future Research

This study provided gender differences and similarities across a diverse range of expectancy value constructs in a number of domains. This study also explored the use of meta-analytic techniques within the framework of an intersectional perspective to gender differences. By gathering data about the social class, ethnicity, and country of origin for each study, more insight was gained into how social forces such as class and different aspects of gender equality are associated with gender differences in self-beliefs and attitudes across math, science, and verbal domains. Results showed a relationship between social class and gender that consistently appeared across math attitudes, showing that more affluent females are potentially prone to greater disadvantage in academic self-beliefs and attitudes relative to their male peers. There was a complex relationship between gender equality, attitudes, and gender segregation in educational outcomes by showing that the gender ratio of females enrolled in STEM versus humanities was a stronger predictor of gender differences, than the broader GII metric commonly used by other researchers. However, there were a number of difficulties that place some limitation on the conclusions that can be drawn from the study, particularly with regards to using meta-analytic methods to explore intersectional questions.

Firstly, this review has revealed that many studies did not include basic demographic information about their samples, particularly in relation to social class and ethnicity. This lack of information made it difficult to gain enough power to perform moderation analyses on all variables included in the meta-analysis. As a result, the ethnicity analysis had to be restricted to the percentage of participants as belonging to an ethnic minority, as opposed to particular ethnic groups. Results indicated that there is likely to be more similarities than differences amongst ethnic groups for gender differences, however, this conclusion is limited in that amongst any given ethnic minority there is a wide variation of cultures and practices, that make the category of ethnic minority extremely heterogeneous, and therefore it is difficult to ascribe reliable conclusions from the data on ethnicity. Furthermore, categorisation of gender equality was also limited in that it was difficult to match GII ratings specifically to the year of publication because of power needed to perform moderation analyses and the GII's relatively recent history. Nonetheless, gender inequality of nations (and other country-level indicators), is unlikely to dramatically change across the years (Stotsky, Shibuya, Kolovich, & Kebhaj, 2016). Thus, results from this study are tentative in that they are based on less precise information compared to other research using data that matches directly to the year in which the gender indices were reported.

This review also revealed that there is a substantial over-representation of white American middle class participants in the literature, confirming wider critiques of psychological research (e.g., Henrich et al., 2010; Hyde, 2007). As such, most of the literature on expectancy value theory and gender is actually based upon a very narrow group of people within society. Future research should be mindful of the need to extend research agendas to sample beyond such narrow demographics, and to promote further research that is inclusive of a diverse range of populations and experiences. Additionally, researchers should aim to report more detailed demographic information relating to the ethnicity and socioeconomic status of participants.

Finally, this review highlighted gaps within the expectancy value literature. In comparison to other subjects, math was over-researched. The domain specific aspects of STEM (e.g., engineering, physical sciences, biological sciences, and computing) were under-researched in comparison to general measures of math and science. Given that gender differences are likely to be better understood using domain specific STEM measures, it is critical that researchers researching the gender STEM gap consider moving away from general science and math measures, in favour of exploring domain specific measures (e.g., physics self-concept). In addition to this, more research focusing on boys' verbal expectancies for success and task values needs to be done. The final research gap that was highlighted was the components of expectancy value that were included in the study. Some aspects of expectancy value were under-researched, leading to variables such as cost and attainment value being neglected in the study of gender differences in education. Further research should explore these constructs in greater depth to better understand student motivation.

In summary, this review has demonstrated the importance of exploring gender differences in relation to social and cultural context, as well as assessing differences across sub-disciplines within STEM. Hopefully this research encourages more studies to begin investigating gender in relation to other social categories, and to further explore and develop new research methods that address the limitations of a quantitative approach to intersectionality. Understanding the social and cultural contexts in which gender differences in self-beliefs and attitudes vary will ultimately help educators identify the particular groups of students who would benefit the most from interventions to decrease gender disparities in educational outcomes. Furthermore, by better understanding the role of social and cultural contexts, we can come a step closer to understanding the environments in which gender gaps in education become most problematic, and those in which gender gaps are at their smallest, thus, paving the way for future research that identifies the most beneficial environments for all students.

CHAPTER 6
THE INTERSECTION OF GENDER, SOCIAL CLASS, GEOGRAPHY AND CULTURAL CONTEXT

Study 2: The Intersection of Gender, Social Class and Cultural Context: A Replication and Extension of Meta-Analysis Findings

Study 1 explored the intersections between social and cultural context using meta-analytic strategies, however, meta-analytic tests of moderation have inherent weaknesses in that access to individual data is restricted, thereby limiting the analyses that can be performed (Marsh et al., 2009; and as discussed in Chapter 4). In contrast, large-scale quantitative databases are well suited to testing moderation effects and for allowing a more precise exploration of data beyond description of effect sizes (Marsh et al., 2009). For instance, do the effects of the meta-analysis replicate even when academic achievement levels are controlled for? Thus, the conclusions of meta-analytic studies can be bolstered by corroborating results with large-scale quantitative data (and vice versa). Study 2 capitalises on this approach by exploring the replicability of meta-analysis results from Study 1; thereby providing an opportunity to corroborate findings through the two complementary methods of meta-analysis and large-scale quantitative data analysis on a national representative sample.

A second aim of Study 2 is to extend upon existing findings in the extant literature and results from Study 1. As outlined in Chapter 3, most research on gender differences have failed to consider gender differences within the framework of social and cultural contexts. While a handful of studies have begun to investigate relations between self-beliefs and values with gender, ethnicity, and country level differences (e.g., Catsambis, 1994; 1995; Else-Quest et al., 2010; Else-Quest et al., 2013; Stoet et al., 2016; Zarret, Malanchuk, Davis-Kean, & Eccles, 2006), even less research has explored the intersection between gender and other social categories such as class and geography. In particular, there is almost no research that investigates the intersection between gender and geography. Although geographic location was decided upon as a key moderator for Study 1, there were simply not enough studies to test for moderation effects. This also holds true for Indigenous status, from which there were no studies that explored gender differences in EVT variables for Indigenous students. Finally, while a number of studies examine gender differences across ethnicities, there has been little research on this area outside of an American context, and in relation to immigrant youths. This is surprising given the wealth of evidence that has implicated social class, geography, Indigenous status, and immigrant background as critical factors in determining young people's educational outcomes (e.g., De Bortoli, & Thomson, 2009; Duong, Badaly, Lui, Schwartz, & McCarty, 2016; Parker, Jerrim, Anders, & Astell-Burt, 2016; Parker, Schoon, Tsai, Nagy, Trautwein, & Eccles, 2012).

Furthermore, as discussed in Chapter 3 most existing intersectional research has focused on achievement, but there has been less emphasis on self-beliefs and values, and even less focus on intersectional differences with regards to educational attainment. By utilising a nationally representative large-scale database of Australian Youth, Study 2 extends upon existing literature by analysing gender differences in self-beliefs, values, achievement and educational attainment with respect to the intersection of social class, Indigenous status, geographic location, and immigration. Thus, the central aim of Study 2 is to provide a test of replicability of the meta-analytic results in Study 1, and also to extend current research by exploring the intersection of gender across a diverse range of social and cultural contexts.

Research Questions and Hypotheses: Study 2

Hypothesis 1. Students from low SES backgrounds, rural/remote locations, and who identify as Indigenous or female will experience poorer outcomes in math and science achievement, math self-beliefs and values, as well as educational attainment in STEM.

Hypothesis 2. In line with meta-analysis findings, there will be a significant interaction effect between social class and gender for EVT-related factors, achievement, and attainment with the gender gap being largest amongst high SES students.

Research Question 1. Are there statistically significant interactions between Indigenous status, immigrant status, geography and gender for EVT-related factors, achievement, and attainment? What are the directions of significant interactions?

Research Question 2. Do the interactions mentioned above remain even after achievement is controlled for?

Method

Participants

The sample was taken from the 2003/Y03 cohort of Longitudinal Study of Australian Youth (LSAY; N = 10, 370; 50.82% male). LSAY follows large nationally representative samples of Australian youth from the age of 15 every year until their mid-twenties. The sample was largely comprised of Australian born students (77.9%), with 11.8% students as first-generation immigrants, and 10.3% second-generation Australian immigrants, 1.9% of the sample also identified as Indigenous Australians. 40% of the young people surveyed had at least one parent with a university level education. The average socio-economic index of participants on the International Socio-economic Index was 52.84 (SD = 15.93), which is substantially higher than the OECD average (OECD, 2011). Data for achievement, self-beliefs and attitudes came from Wave 1 of data collection whereby participant mean age was 15.69 years (*SD* = .29; *Range* = 15-16). Educational attainment data was based on senior high school course selection, where participants were in Grade 12 at age 18 (n = 6,658), and on university STEM course enrolment at age 19 (n = 2,235). Importantly, participants in the final data wave at age 19 were a subsample of the larger database, whereby only students who enrolled in senior high school STEM courses were asked follow-up questions regarding whether or not they continued their STEM studies at university (see Figure 1 for a visual representation). Importantly, readers should note that the Australian school system differs from the US educational system, whereby STEM courses are not compulsory for senior high school students.

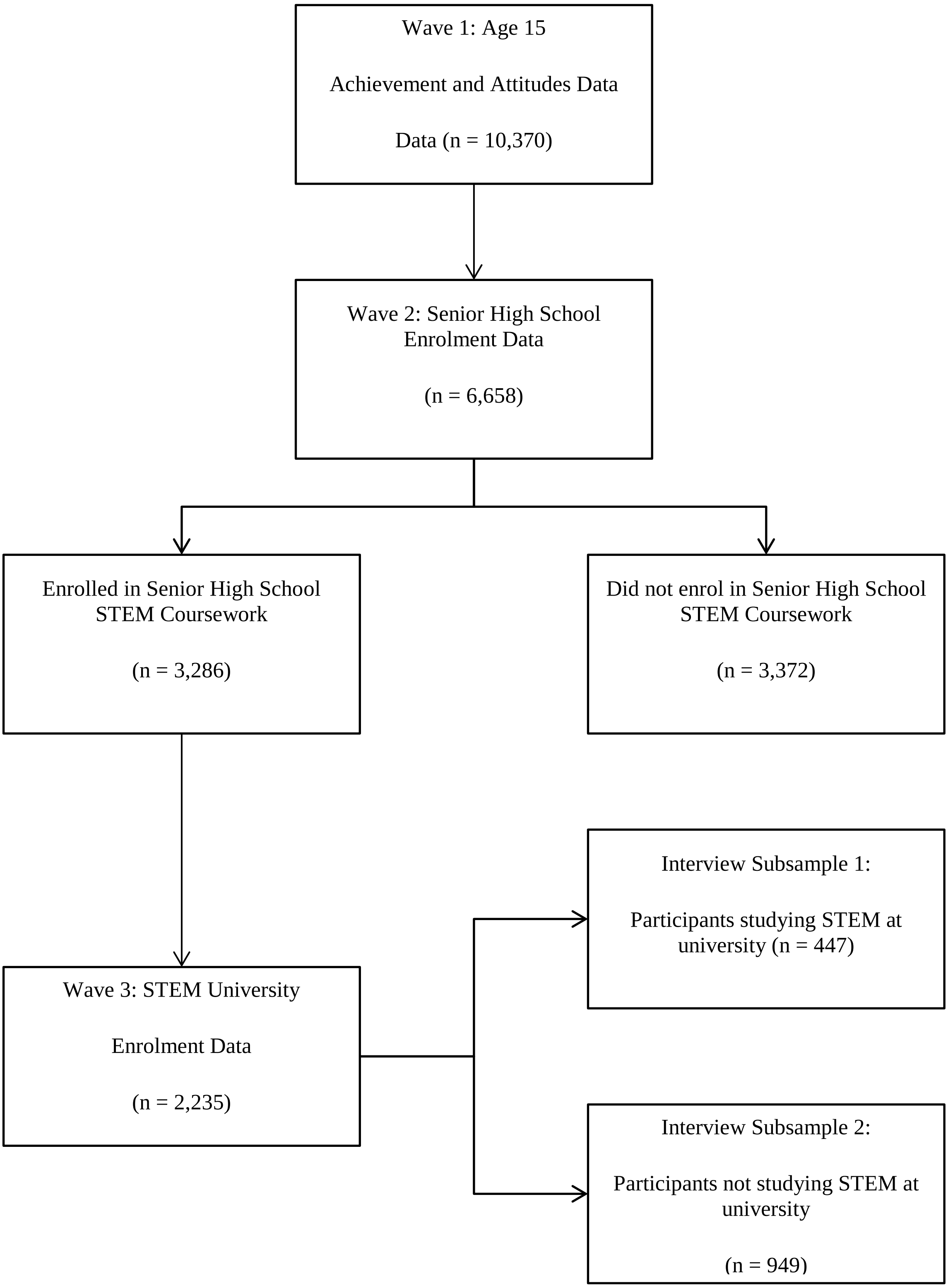

Figure 7. Flow diagram of data collection across time waves for Studies 2, 3 and 4.

Measurements

Math self-beliefs. Math self-beliefs incorporated in this study were the PISA indexes for math self-efficacy and math self-concept. Math self-efficacy was rated on a 4 point Likert scale, ranging from 0 (*very confident*) to 3 (*not at all confident*), and items were inversed so that higher values indicated higher levels of math self-efficacy. The PISA math self-efficacy index was based Bandura’s (1997) conceptualisation of self-efficacy, with items measuring an individual’s self-confidence in doing a number of applied mathematical tasks (e.g., “How confident do you feel about having to do the following mathematics tasks? – Calculating how much cheaper a TV would be after a 30% discount.). Math self-concept was rated on a 4 point Likert scale, ranging from 0 (*strongly agree*) to 3 (*strongly disagree*). Again, items were inversed so that higher values of indicated higher levels of math self-concept. Math self-concept focused on the ability component of subject specific self-concept beliefs as described by Eccles and Wigfield (1995) and as measured in Marsh’s measure for self-concept (e.g., “I get good marks in mathematics”; see Appendix for list of all items). Cronbach’s alphas for the Australian subset of PISA 2003 (the sample used in this study) were as follows: Math self-efficacy (α = 0.86); math self-concept (α = 0.89) (OECD, 2005).

Math attitudes and affect. Attitudes and affect towards math were measured by the PISA items of math interest, math instrumental motivation (hereafter referred to as utility value), and math anxiety. All scales were rated on a 4 point Likert scale, ranging from 0 (*strongly agree*) to 3 (*strongly disagree*). All scales were inverted so that positive values indicated stronger endorsement of each attitude. Math anxiety was based on Wigfield and Meece’s (1988) feelings of worry, stress and helplessness when doing mathematics E.g., “I get very nervous doing mathematics problems”. Math interest was measured on a 4 point Likert scale, and was based on what Wigfield et al. (1997) called the enjoyment aspect of task value (e.g., “I do mathematics because I enjoy it”). Math utility value measured the extent to which students were motivated to learn mathematics because of benefits for their future studies and career. E.g. “I will learn many things in mathematics that will help me get a job” (see Appendix for list of all index items). Cronbach’s alphas for the Australian subset of PISA 2003 (the sample used in this study) were as follows: Math interest (α = 0.90); math utility value (α = 0.89); and math anxiety (α = 0.82) (OECD, 2005).

Teacher support. Teacher support was measured in the PISA database as an indicator of perceived teacher support in math lessons, and was rated on a 4 point Likert scale ranging from 0 (*every lesson*), 1 (*most lessons*), 2 (*some lessons*), and 3 (*never or hardly ever*), with all scores reversed so that positive values reflected higher levels of perceived teacher support (see Appendix for list of all teacher support index items). Cronbach's alpha for the Australian subset of PISA 2003 (the sample used in this study) was ($\alpha = 0.87$) (OECD, 2005).

Academic achievement. Math, science, and reading achievement were measured by PISA achievement tests (see OECD, 2005 for detailed information on development and validation). Achievement tests featured multiple choice questions, as well as closed and open-ended responses. Math achievement measured ability in the areas of space and shape, quantity, change in relationships, and uncertainty. Reading achievement measured ability in retrieving information, interpreting, and reflecting. Science achievement measured ability in: describing, explaining and predicting; interpreting scientific evidence; and understanding scientific investigation. PISA assessments of ability utilise matrix sampling, and consequently use item response theory to create a set of 5 plausible values for each individual's underlying achievement in each domain (e.g., 5 scores for math, 5 scores for science). PISA survey organisers scaled these scores (across all OECD countries) to have a mean of 500 points and a standard deviation of 100. OECD (2005) reports for reliabilities of achievement scales were as follows: math achievement ($\alpha = 0.89$); science achievement ($\alpha = 0.84$); reading achievement ($\alpha = 0.85$) (OECD, 2005). Analyses for each plausible value was ran separately and results were combined using the formulas defined by Rubin (1987) (see Supplementary Materials for further details).

STEM course selection. During Wave 5, participants who had previously provided data that they had enrolled in a high school STEM course (1 = yes; 0 = no) were asked if they were currently studying in a STEM field at a tertiary level. Those who were studying in a math or science field at a tertiary level were coded as one; those that had not were coded as zero.

Social class. Social class was measured by the Economic, Social, and Cultural Status (ESCS) scale, a measure that is comprised on the household possessions, highest parental occupational status, and highest number of years of parental education indexes from PISA 2003. To create the ESCS scale missing values from these three variables were imputed to create a metric with an OECD average of 0, and an OECD standard deviation of 1. A principal components analysis was used to obtain ESCS scores with OECD population weights.

Geography. Geography was coded according to the location of school that students were attending age 15. Categories were grouped into metropolitan, provincial (referred to as regional centres in this thesis), and rural/remote locations. These categories were based on an existing school geographic classification framework utilised by the LSAY database (see the LSAY cohort report by Underwood & Rothman, 2007 for further information). In this study 71.7% of students lived in metropolitan areas, 27.7% of students came from regional towns and centres, and 0.6% students lived in remote locations (percentage reported as exists in the LSAY weighted database).

Indigenous status. Indigenous status of students was determined by the question, "*Are you of Aboriginal or Torres Strait Islander origin*?". A response of 0 indicated that students were non-Indigenous, and a response of 1 indicated that students identified as Indigenous Australians. Approximately 1.9% of participants identified as Indigenous.

Immigration status. Immigration status was determined from PISA 2003 questions relating to country of birth, and parental country of birth. Immigration status of students was coded as 1 (*students with at least both parents born in Australia),* 2 (*students with both parents born overseas*), and 3 (*students who were born overseas, and whose parents were also born overseas*).

Gender. Gender was coded as 1 (*male*) and 2 (*female*).

Data Analysis

LSAY is a large longitudinal database that utilises complex sampling procedures in order to capture the experiences of a diverse range of young Australians. In order to do this, LSAY employs oversampling of some demographic groups. Consequently, I used sampling weights provided by LSAY to correct for any sampling bias in the analyses. Another complexity of the LSAY database is the issue of attrition due to the passing of time between data collection points. I utilised attrition weights provided by LSAY to take into account any bias from participant attrition. LSAY also has a complex structure with schools as the primary sampling unit. To account for this complex design, I used the 80 balanced repeated replication weights provided by the survey organisers to ensure correct standard errors (see

Lumley, 2010). Finally, scales were standardised to a common metric (M = 0; SD = 1) to facilitate the comparison of parameter estimates (see Supplementary Materials for further details).

Analysis of the data was conducted using R (see Supplementary Materials for R scripts for each analysis). To investigate the relationship between gender and various social and cultural contexts, a number of interaction effects were run independently of one another. Significant interactions were plotted in graphical form to further identify the direction of effects. Finally, for each analysis a separate interaction that controlled for math and science achievement was run to identify the degree to which effects could be accounted for by differing levels of achievement.

Results

Gender and Social Class

Predicting attitudes and achievement. Results for gender by social class indicated that both gender and social class predicted math attitudes (see Table 3). Being male and being from a high SES background were both predictors of having positive attitudes towards math. Nevertheless, there was only one significant - but small - interaction effect, which was for math self-efficacy. This interaction showed that the gender gap in math self-efficacy became larger amongst high SES students (see Figure 8a below). Finally, despite the trend of large gender differences amongst higher SES students, low SES girls still had the lowest math self-efficacy out of all groups.

To test whether this interaction was significant regardless of achievement, a second analysis controlling for achievement was conducted. Results showed that once achievement was controlled for, gender had a stronger effect than social class on math attitudes. Overall, results showed that when comparing students of equal ability, female students had lower math self-efficacy regardless of their socioeconomic status. Nonetheless, the interaction effect between social class and gender remained significant, again showing that the gender gap for math self-efficacy became larger amongst high SES students (see Figure 8b below). The interaction was mainly due to boys benefiting from high SES more so than their female peers. There were no interaction effects between SES, gender and achievement; however, results showed that SES was a strong predictor of academic achievement when considered alone.

Table 3

SES and Gender Interaction Effects for Math Attitudes and Academic Achievement

	SES	SE	Gender	SE	SES by Gender	SE	Intercept
Attitudes without controls							
Math self-efficacy	0.27*	0.02	-0.38*	0.03	**-0.07***	0.03	0.11
Math self-concept	0.10*	0.02	-0.32*	0.03	-0.01	0.03	0.09
Math interest	0.04	0.03	-0.23*	0.03	0.02	0.03	0.09
Math utility value	0.07*	0.02	-0.23*	0.04	-0.01	0.03	0.07
Math anxiety	-0.12*	0.02	0.35*	0.03	0.02	0.03	-0.13
Math teacher support	0.09*	0.02	0.04	0.03	-0.03	0.02	-0.02
Achievement							
Math achievement	0.38*	0.02	-0.13*	0.04	-0.02	0.04	-0.11
Science achievement	0.41*	0.02	-0.06	0.03	-0.04	0.04	-0.13
Reading achievement	0.40*	0.03	0.36*	0.03	-0.02	0.03	-0.35
Attitudes with controls							
Math self-efficacy	0.10*	0.02	-0.32*	0.03	**-0.06***	0.03	0.16
Math self-concept	-0.05*	0.02	-0.27*	0.03	0.00	0.03	0.13
Math interest	-0.04	0.02	-0.20*	0.02	0.03	0.03	0.11
Math utility value	0.00	0.02	-0.21*	0.04	-0.01	0.03	0.09
Math anxiety	0.01	0.02	0.31*	0.03	0.02	0.03	-0.16
Math teacher support	0.06*	0.02	0.05	0.03	-0.03	0.02	-0.02

**Note.* Table 3 presents interaction effects for each EVT and achievement construct independently.
Attitude with controls are the interaction effects for attitudes once student achievement is controlled for.

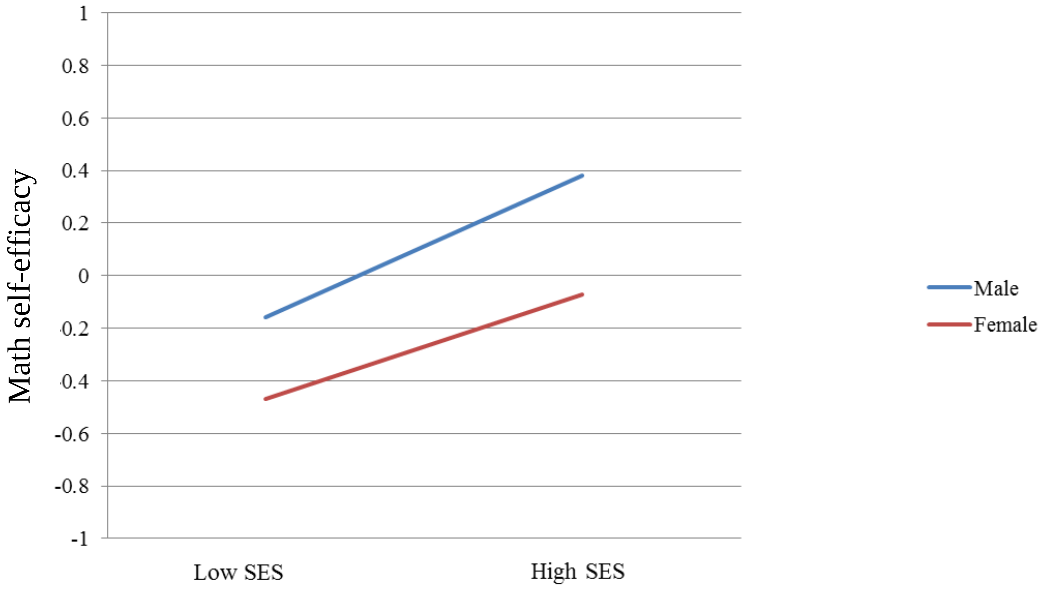

Figure 8a. Interaction effect of gender by social class for math self-efficacy (no controls for achievement).

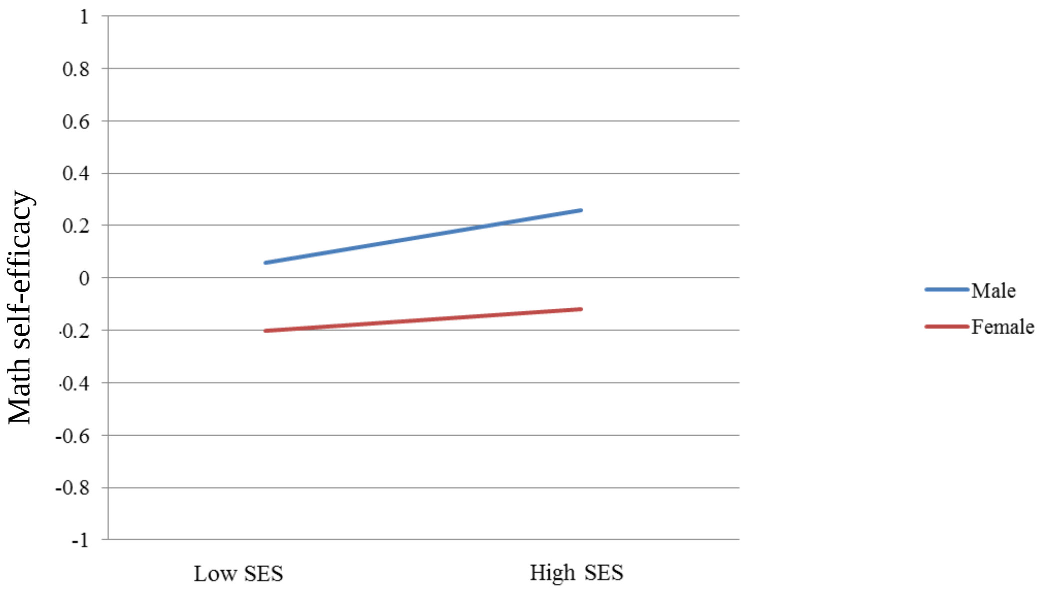

Figure 8b. Interaction effect of gender by social class for math self-efficacy (controlling for achievement).

Gender and Indigenous Status

Results for gender by Indigenous status indicated that being Indigenous and being female were independently related to poor attitudes towards math (see Table 4). Only one interaction effect (utility value) between Indigenous status and gender was statistically significant (see Figure 9a below). This interaction was relatively large, and showed that the effect of female gender was more negative for non-Indigenous girls than for their Indigenous peers. Overall, non-indigenous boys had the highest levels of math utility value, while Indigenous boys had the lowest levels of math utility value.

To test whether the aforementioned interaction occurred regardless of ability, another analysis was performed controlling for math achievement. These results showed a similar pattern: for non-indigenous students, being male had a positive effect on math utility value (see Figure 9b). In contrast, amongst Indigenous students being male had a negative effect on math utility value. Controlling for math achievement, Indigenous female students had the highest levels of math utility value while non-indigenous females had the lowest levels of math utility value. Results indicated that the negative effect of gender remained when controlling for achievement, whereas negative effects for math attitudes amongst Indigenous students did not remain once achievement was controlled for. These results reflected a similar trend seen for social class results.

Indigenous students experienced substantial disadvantage in relation to their scholastic achievement across all areas (see Figures 9c and 9d). Gender differences in achievement were larger for Indigenous students, with girls outperforming their male peers in math. Overall, Indigenous boys experienced the lowest levels of math achievement.

Table 4

Indigenous Status and Gender Interaction Effects for Math Attitudes and Academic Achievement

	Indigenous	SE	Gender	SE	Indigenous by Gender	SE	Intercept
Attitudes without controls							
Math self-efficacy	-0.30*	0.10	-0.37*	0.04	-0.05	0.15	0.07
Math self-concept	-0.24*	0.11	-0.32*	0.03	0.02	0.17	0.08
Math interest	-0.13	0.12	-0.24*	0.03	0.11	0.17	0.09
Math utility value	-0.32*	0.08	-0.24*	0.04	**0.44***	0.12	0.07
Math anxiety	0.33*	0.14	0.34*	0.03	-0.08	0.19	-0.11
Math teacher support	0.01	0.15	0.03	0.03	0.06	0.15	-0.03
Achievement							
Math achievement	-1.08*	0.12	-0.12*	0.04	**0.33***	0.15	-0.16
Science achievement	-1.03*	0.13	-0.05	0.04	0.33	0.16	-0.19
Reading achievement	-1.04*	0.15	0.37*	0.04	**0.42***	0.17	-0.41
Attitudes with controls							
Math self-efficacy	0.13	0.11	-0.31*	0.03	-0.10	0.14	0.15
Math self-concept	0.09	0.11	-0.27*	0.03	-0.03	0.16	0.14
Math interest	0.02	0.12	-0.22*	0.03	0.08	0.17	0.12
Math utility value	-0.18*	0.08	-0.22*	0.04	**0.42***	0.13	0.10
Math anxiety	0.05	0.14	0.30*	0.03	-0.04	0.19	-0.16
Math teacher support	0.10	0.12	0.05	0.03	0.05	0.16	-0.02

**Note.* Table 4 presents interaction effects for each EVT and achievement construct independently. Attitude with controls are the interaction effects for attitudes once student achievement is controlled for.

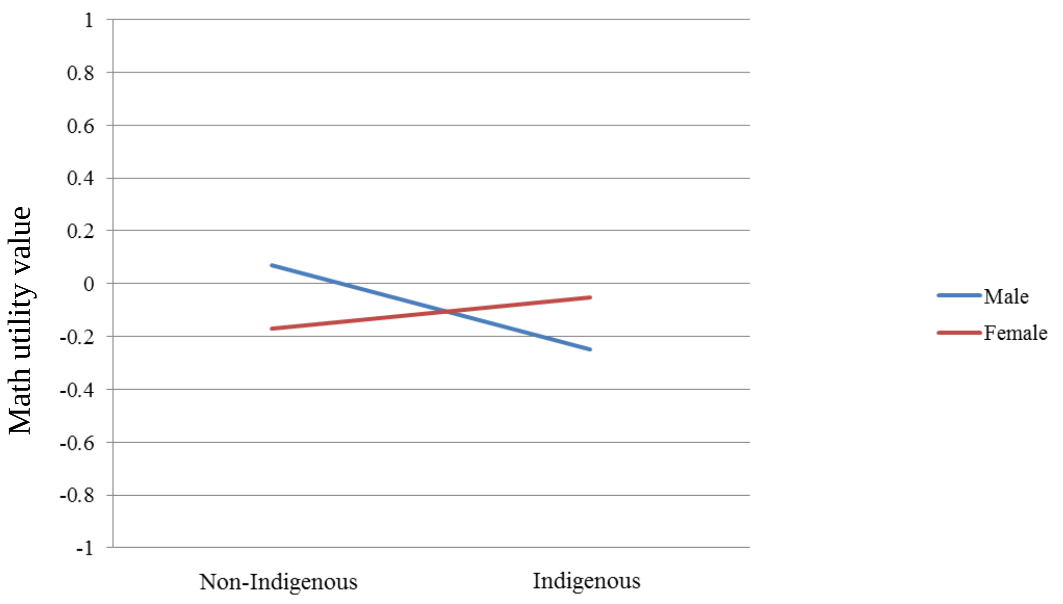

Figure 9a. Interaction effect of gender by Indigenous Status for Math Utility Value (no controls for achievement).

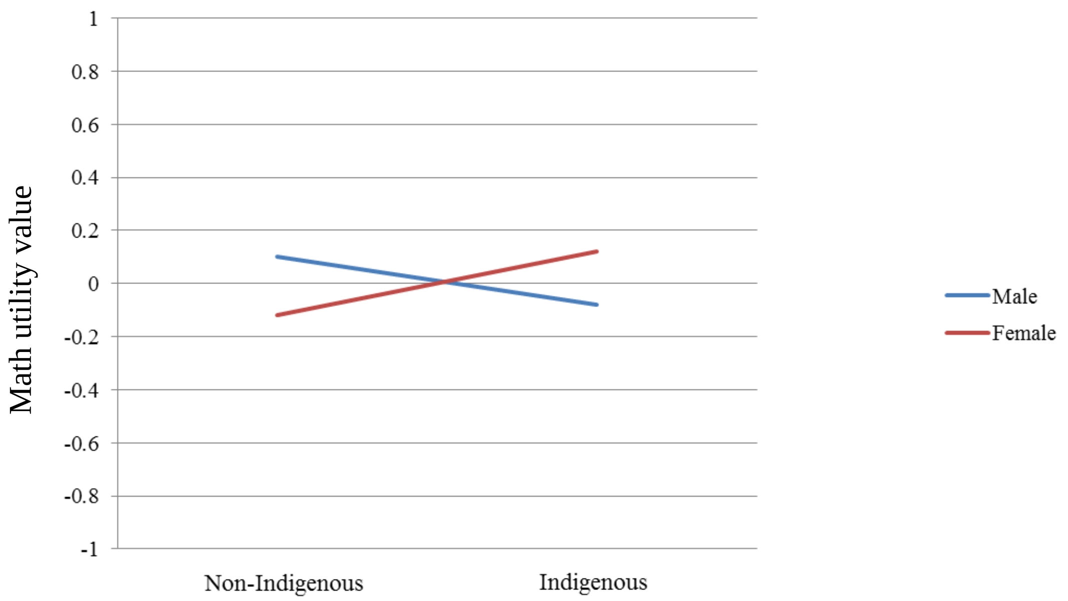

Figure 9b. Interaction effect of gender by Indigenous status for math utility (controlling for achievement).

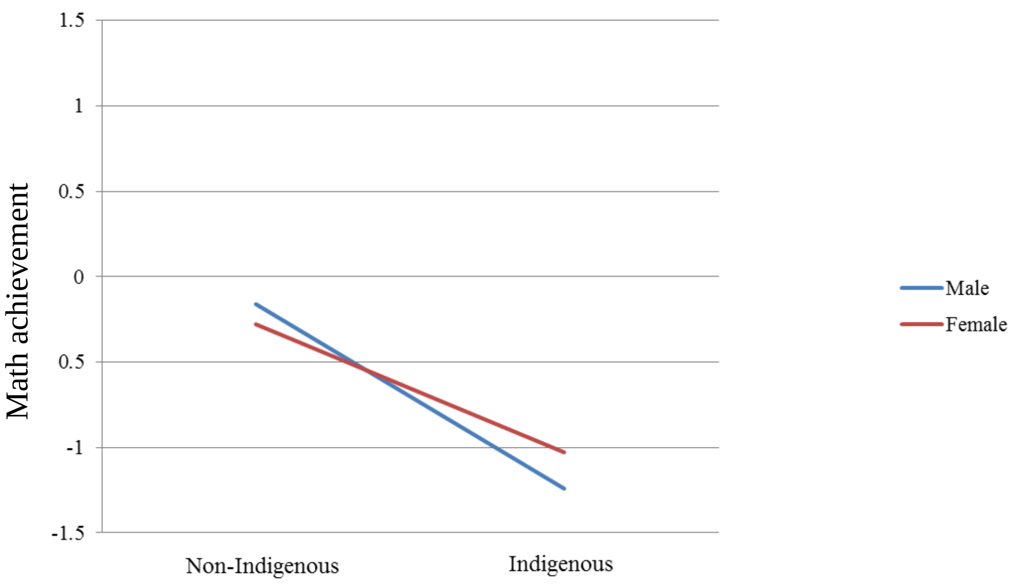

Figure 9c. Interaction effect of gender by Indigenous status for math achievement.

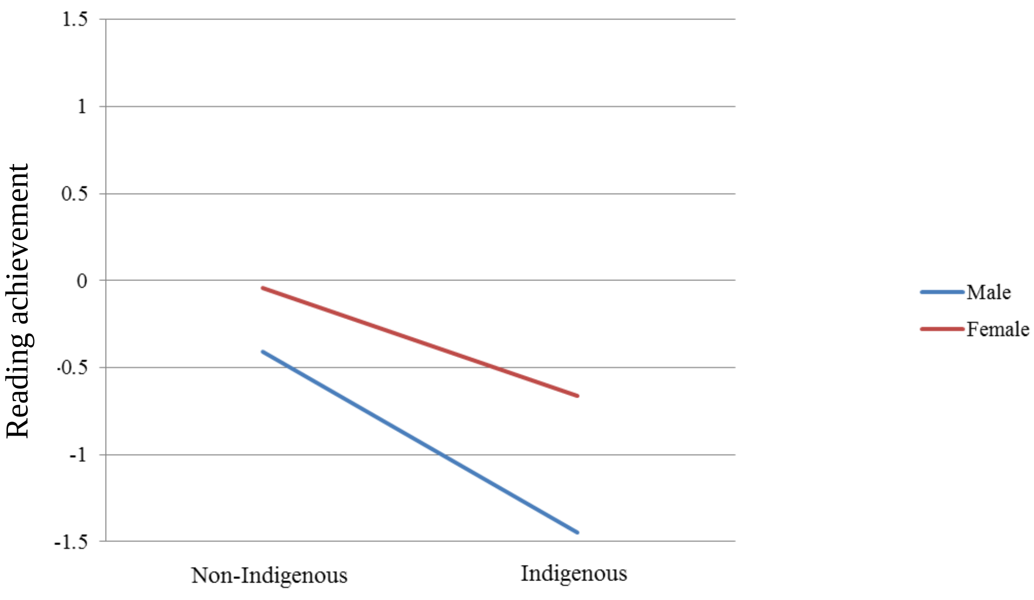

Figure 9d. Interaction effect of gender by Indigenous status for reading achievement

Gender and Geography

Geography was represented by three categories (metropolitan, regional, and remote). Remote students had significantly lower math anxiety scores than metropolitan students, but geography had no significant effect on other math attitudes. Gender differences in favour of girls were also significantly larger for math utility value, but gender did not interact with other math attitudes (see Table 5). Figure 10a shows that the gender gap for math anxiety is similar for students living in metropolitan and regional centres; however, the gender gap in math anxiety is larger amongst students from remote locations. Overall, girls from remote regions reported the highest levels of math anxiety, while boys from remote regions reported the lowest levels of math anxiety.

To test whether the interaction between gender and geography occurred regardless of ability, another analysis was performed controlling for math achievement. These results showed a similar pattern: gender was a much stronger predictor of math attitudes than the location a student lived in. The one exception to this was math anxiety, whereby living in a remote location (compared to urban) had a strong negative effect on math anxiety (see Figure 10b). In addition to this, the interaction effect for math anxiety was replicated, showing that even when comparing students of equal ability, the gender gap amongst remote students is much larger than the gender gap amongst students living in metropolitan and regional centres. After controlling for achievement, there was also a significant interaction effect for math utility value. This effect showed that the gender gap in math utility value became larger amongst students from remote locations, even when comparing students of equal ability.

Students from regional centres and remote communities had lower levels of achievement compared to their urban peers. There was only one significant interaction effect between gender and geography, which was for reading achievement (see Figure 10c).

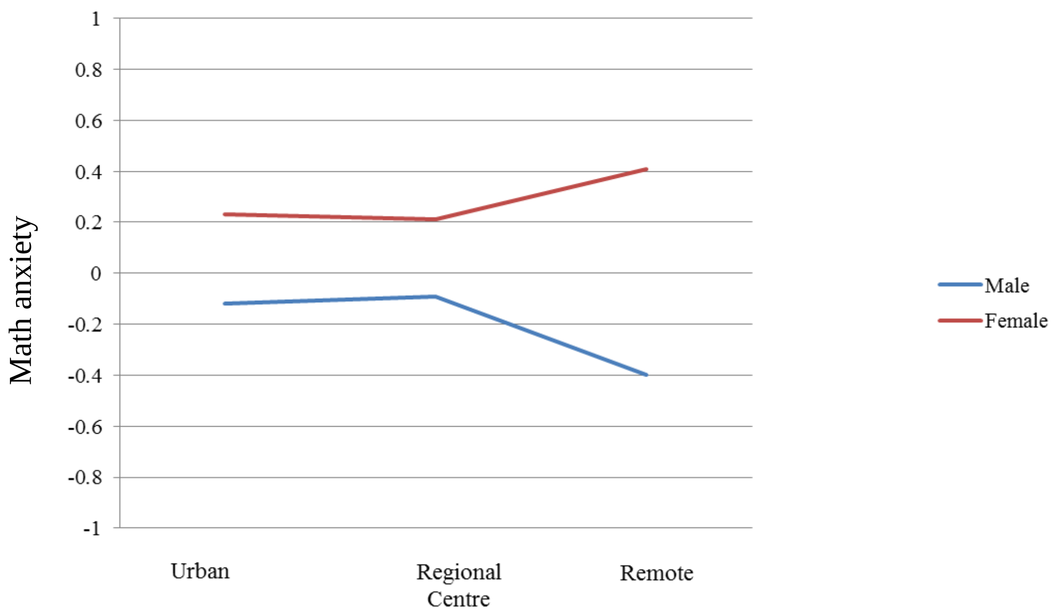

Figure 10a. Interaction effect of gender by geography for math anxiety (without controlling for achievement).

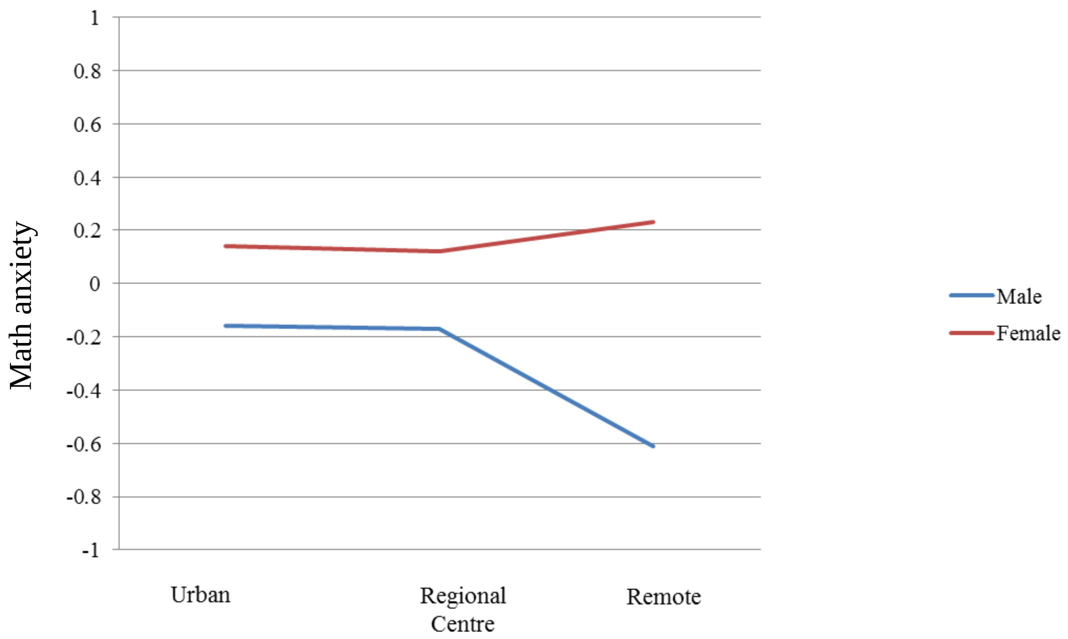

Figure 10b. Interaction effect of gender by geography for math anxiety (controlling for achievement).

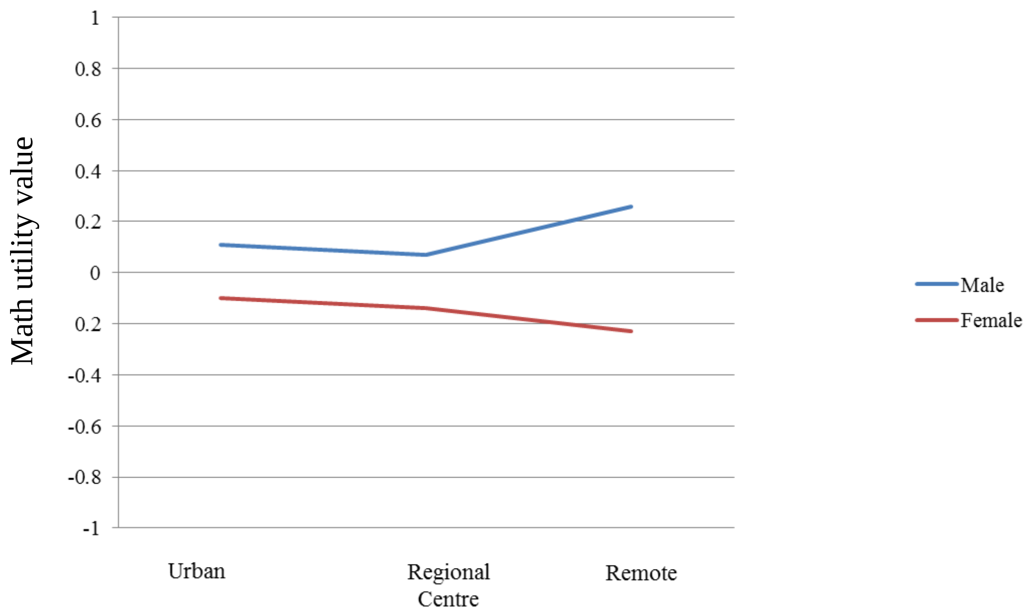

Figure 10c. Interaction effect of gender by geography for math utility value (controlling for achievement).

Table 5

Geography and Gender Interaction Effects for Math Attitudes and Academic Achievement

	Regional	SE	Remote	SE	Gender	SE	Regional by Gender	SE	Remote by Gender	SE	Intercept
Attitudes without controls											
Math self-efficacy	-0.16*	0.08	-0.42*	0.13	-0.39*	0.04	0.09	0.09	0.24	0.16	0.11
Math self-concept	-0.08	0.05	0.13	0.14	-0.33*	0.03	0.05	0.07	-0.11	0.15	0.10
Math interest	-0.11	0.06	-0.11	0.18	-0.24*	0.04	0.02	0.07	-0.07	0.11	0.12
Math utility value	-0.06	0.08	0.06	0.15	-0.24*	0.04	0.02	0.10	-0.24	0.15	0.08
Math anxiety	0.03	0.05	-0.28	0.20	0.35*	0.04	-0.05	0.08	**0.46***	0.13	-0.12
Math teacher support	-0.02	0.05	-0.23*	0.10	0.06	0.04	-0.09	0.06	0.06	0.10	-0.02
Achievement											
Math achievement	-0.14*	0.07	-0.51*	0.16	-0.15*	0.06	0.12	0.08	0.26	0.21	-0.14
Science achievement	-0.13*	0.06	-0.55*	0.20	-0.08	0.05	0.14	0.09	0.44	0.24	0.01
Reading achievement	-0.20*	0.06	-0.62*	0.20	0.31*	0.05	0.17*	0.08	0.48	0.25	-0.18
Attitudes with controls											
Math self-efficacy	-0.10	0.06	-0.16	0.17	-0.32*	0.03	0.05	0.07	0.12	0.21	0.18
Math self-concept	-0.04	0.04	0.33	0.18	-0.27*	0.03	0.02	0.06	-0.21	0.17	0.15
Math interest	-0.10	0.05	-0.02	0.20	-0.21*	0.03	0.00	0.07	-0.11	0.12	0.14
Math utility value	-0.04	0.08	0.15	0.17	-0.21*	0.04	0.00	0.09	**-0.28***	0.13	0.11
Math anxiety	-0.01	0.05	-0.45*	0.21	0.30*	0.03	-0.01	0.07	**0.54***	0.16	-0.16
Math teacher support	-0.01	0.05	-0.18	0.11	0.07*	0.04	-0.10	0.06	0.04	0.10	-0.01

**Note*. Table 5 presents interaction effects for each EVT and achievement construct independently. Statistically significant interactions are bolded. Attitude with controls are the interaction effects for attitudes once student achievement is controlled for.

Gender and Immigration Status

Results from Table 6 show that gender is a stronger predictor of math attitudes than whether students were immigrants to Australia or the child of an immigrant parent. Nonetheless, being an immigrant or a first-generation immigrant had a positive association with math self-beliefs, interest, and instrumental motivation. There were no significant interaction effects between gender and immigrant status.

When controlling for achievement, results were similar, showing that being from an immigrant family has a positive effect on math attitudes even when comparing students of equal ability. Again, there were no significant interaction effects between gender and immigrant status.

Predicting Educational Attainment in STEM

Social class. A logistic regression was run for STEM educational attainment at the senior high school level, and for university enrolment, and results were evaluated at the mean of math and science achievement (see Table 7a). Social class had a significant effect on senior high school enrolment, showing that students with higher SES were more likely to enrol in senior high school math or science classes. In contrast, there was almost no effect of social class predicting university enrolment in STEM courses. When comparing students of equal math and science ability, female students were more likely to enrol in senior high school STEM. However, being female had an extremely large negative effect on the likelihood of enrolling in a university STEM course. There were no statistically significant effects for STEM attainment.

Geography. A logistic regression was run for STEM educational attainment at the senior high school level, and for university enrolment. Results were evaluated at the mean of math and science achievement (see Table 7b). Geography was significant for senior high school enrolment in STEM, but not for university enrolment in STEM. Being a student in a regional location had a negative effect on senior high enrolment compared to students from urban areas. There was a significant interaction effect for gender and being from a regional centre in predicting senior high school enrolment. Amongst urban students, the gender gap for the likelihood of studying STEM at senior high school was very small, but this gap grew bigger amongst students living in regional areas, with girls being more likely to enrol in STEM at a high school level.

Table 6

Immigrant Status and Gender Interaction Effects for Math Attitudes and Academic Achievement

	First Generation Migrant	SE	Born Overseas	SE	Gender	SE	First Generation Migrant by Gender	SE	Born Overseas by Gender	SE	Intercept
Attitudes without controls											
Math self-efficacy	0.12	0.08	0.24*	0.09	-0.35*	0.04	-0.01	0.10	-0.22	0.12	0.04
Math self-concept	0.18*	0.06	0.20*	0.08	-0.32*	0.03	-0.07	0.07	0.01	0.11	0.05
Math interest	0.35*	0.06	0.41*	0.08	-0.23*	0.04	-0.09	0.13	-0.04	0.13	0.01
Math utility value	0.22*	0.08	0.17*	0.07	-0.25*	0.04	-0.04	0.10	0.06	0.11	0.03
Math anxiety	0.01	0.07	-0.03	0.10	0.31*	0.03	-0.10	0.10	-0.02	0.10	-0.16
Math teacher support	-0.08	0.08	0.00	0.09	0.06	0.03	0.03	0.10	-0.09	0.10	0.00
Achievement											
Math achievement	-0.03	0.09	-0.03	0.10	-0.10*	0.04	-0.03	0.11	-0.18	0.16	-0.15
Science achievement	-0.10	0.09	-0.19	0.11	-0.04	0.04	0.00	0.10	-0.13	0.16	-0.15
Reading achievement	-0.06	0.1	-0.13	0.11	0.39	0.04	-0.01	0.11	-0.21	0.16	-0.38
Attitudes with controls											
Math self-efficacy	0.13*	0.07	0.25*	0.07	-0.30*	0.03	0.01	0.09	-0.15	0.09	0.11
Math self-concept	0.19*	0.05	0.19*	0.09	-0.28*	0.03	-0.05	0.07	0.10	0.11	0.10
Math interest	0.36*	0.06	0.40*	0.08	-0.21*	0.03	-0.09	0.08	0.00	0.13	0.04
Math utility value	0.23*	0.08	0.17*	0.08	-0.23*	0.04	-0.04	0.10	0.09	0.11	0.06
Math anxiety	0.01	0.07	-0.03	0.10	0.31*	0.03	-0.10	0.10	-0.02	0.10	-0.16
Math teacher support	-0.08	0	0.00	0.08	0.06	0.03	0.03	0.10	-0.09	0.10	0.00

**Note*. Table 6 presents interaction effects for each EVT and achievement construct independently. Statistically significant interactions are bolded. Attitude with controls are the interaction effects for attitudes once student achievement is controlled for. First generation migrant refers to students whose parents were born overseas. Both first generation migrant and born overseas categories are compared to students whose parents were born in Australia.

Immigrant status. Table 7c shows the results for immigrant status and gender prediction of educational attainment in STEM, controlling for achievement scores. First and second-generation immigrant students were more likely to study STEM at a senior high school level, however, there was no effect for university enrolment. There was a significant interaction effect for senior high school enrolment. This effect showed that the gender effect for senior high school enrolment reversed amongst students who had immigrated to Australia, with male students who had immigrated from overseas being the most likely to enrol in a senior high school STEM course. In contrast male students without an immigrant background had the lowest odds of studying STEM in Year 11 or 12, followed closely by female students without an immigrant background.

Discussion

This study had two primary aims. The first aim was to explore the replicability of the results from Study 1. Thus, the current study investigated the degree to which moderation analyses from the meta-analysis could be corroborated in a large-scale nationally representative sample. The second aim of the Study 2 was to extend upon the current literature, by exploring whether social class, geography, Indigenous status and immigrant status were related to the magnitude of gender differences in expectancy value variables, as well as academic achievement and educational attainment in STEM. Overall, results painted a complex picture. There was some evidence that results from the current study mirrored existing literature and the meta-analysis of Study 1. However, this was tempered by the fact that a number of effects could not be replicated in the current study. Finally, social categories of Indigenous status, geography and immigrant status showed distinct patterns of effects that uniquely influenced experiences of gender. Research questions, hypotheses and directions for future research are discussed in further detail below.

Table 7a

SES and Gender Interaction Effects for STEM Attainment

	SES	SE	Gender	SE	SES by Gender	SE	Intercept
Senior high school enrolment	0.20*	0.06	0.20*	0.08	-0.14	0.07	-0.13
University enrolment	-0.11	0.09	-0.53*	0.11	-0.06	0.11	0.38

**Note.* Results control for STEM achievement and present the log odds of enrolment.

Table 7b

Indigenous and Gender Interaction Effects for STEM Attainment

	Indigenous status	SE	Gender	SE	Indigenous status by Gender	SE	Intercept
Senior high school enrolment	-0.19	0.31	0.18*	0.09	0.02	0.10	-0.11
University enrolment	-0.22	0.61	-0.59*	0.11	0.02	0.77	0.49

**Note.* Results control for STEM achievement and present the log odds of enrolment.

Table 7c

Geography by Gender Interaction Effects for STEM Attainment

	Regional	SE	Remote	SE	Gender	SE	Regional by Gender	SE	Remote by Gender	SE	Intercept
Senior high school enrolment	-0.33*	0.14	0.20	0.26	0.07	0.10	**0.41***	0.17	0.45	0.47	-0.03
University enrolment	0.08	0.19	0.60	0.69	-0.48*	0.12	-0.24	0.26	-0.48	1.71	0.01

**Note*. Results control for STEM achievement and present the log odds of enrolment. Predicted probabilities for senior high school STEM enrolment were as follows: Urban males = 0.49, urban females = 0.51, regional males = 0.41, regional females = 0.53, remote males = 0.54, and remote females = 0.66.

Table 7d

Immigrant and Gender Interaction Effects for STEM Attainment

	First Generation Migrant	SE	Born Overseas	SE	Gender	SE	First Generation Migrant by Gender	SE	Born Overseas by Gender	SE	Intercept
Senior high school enrolment	0.44*	0.14	1.18*	0.16	0.19*	0.10	0.17	0.20	**-0.48***	0.20	-0.28
University enrolment	0.03	0.27	-0.01	0.20	-0.67*	0.12	0.47	0.29	0.48	0.24	0.34

**Note*. Results control for STEM achievement and present the log odds of enrolment. Predicted probabilities for senior high school STEM enrolment were as follows: Australian born females = 0.48, Australian born males = 0.43, First generation females = 0.63, First generation males = 0.54, females born overseas = 0.65, and males born overseas = 0.71.

Educational Outcomes for Young Australians

In support of Hypothesis 1, the data showed that students from low SES backgrounds, Inidigenous status, and those who were female reported lower scores for math self-beliefs, attitudes, achievement, and STEM educational attainment. Students from lower SES backgrounds experienced more negative self-beliefs, attitudes, poorer achievement, and had lower odds of enrolling in senior high school STEM compared to their higher SES peers. Interestingly, the effect of social class all but disappeared once achievement was controlled for, indicating that the high SES advantage in self-beliefs and attitudes was not existent when comparing students of equal ability.

Indigenous students faced substantial disadvantage in all domains of academic achievement. Indigenous students also experienced more negative self-beliefs and attitudes towards math compared to their non-indigenous peers. This effect reflects previous research that has documented the substantial achievement gap that exists for young Indigenous Australians (e.g., De Bortoli & Thomson, 2009). Interaction effects for self-beliefs and attitudes diminished after controlling for achievement, suggesting that non-indigenous and Indigenous students of equal ability experience similar levels of self-beliefs and attitudes towards math. Indigenous status had a negative effect on educational attainment in STEM; however, these results did not reach statistical significance.

The effect of living in a regional or remote community was negative relative to urban students; however, this effect did not extend to self-beliefs and attitudes. Nonetheless, there were large differences in achievement (particularly for remote students) that showed that young people living in urban areas were more likely to perform well in school compared to their non-urban peers. Regional students were less likely to enrol in senior high school STEM courses compared to students from urban areas; however, this trend did not extend to students living in remote communities.

Being an immigrant, or belonging to an immigrant family, was positively related to higher math attitudes – even after controlling for achievement. This finding is in line with previous literature that has showed a positive effect of immigrant background on educational outcomes (e.g., Duong et al., 2016). Finally, students with immigrant backgrounds had a stronger likelihood of enrolling in high school STEM, but this effect did not extend towards university STEM coursework.

Replicability of Social Class Effects

A major finding of Study 1 was that for a number of EVT-related variables, there was a significant moderation effect for social class. In the current study, there was mixed evidence with regards to a social class by gender interaction effect (see Hypothesis 2). Although social class interacted with gender in a number of EVT-related variables in the meta-analysis (and in previous literature), many of these variables did not reach statistical significance in the current study. Nonetheless, the social class by gender effect was replicated for the variable of math self-efficacy, showing that the gender difference in math self-efficacy became larger for students from high SES backgrounds. This effect remained even when controlling for math and science achievement. Importantly, this effect was in the same direction as previous trends for social class effects seen in Study 1 and previous literature.

Intersections between Gender with Social and Cultural Contexts

Indigenous status. There was some evidence that Indigenous status influenced the size of gender effects for EVT-related variables. In particular, there was a large interaction effect for math utility value. While non-indigenous girls had lower levels of math utility compared to their male peers, the reverse was true for Indigenous students. In contrast, Indigenous boys reported lower math utility value compared to Indigenous girls. This interaction remained even when math and science achievement were controlled for. Notably, Indigenous girls had higher math utility value than any other group once achievement was held constant. A similar trend emerged for achievement variables, whereby female Indigenous students outperformed their male peers, while the reverse trend occurred for non-indigenous students.

Geography. Again, there was some evidence that indicates that geographical location influences the size of gender effects in EVT-related variables. The gender difference for math anxiety was significantly larger amongst students from remote locations. Interestingly, this effect became even larger when achievement levels were controlled for, showing that when comparing students of equal ability, girls from remote locations experienced substantially higher levels of math anxiety compared to their male counterparts. Furthermore, an interaction effect for math utility was uncovered once analyses controlled for achievement. This interaction revealed a similar trend to the math anxiety effect, whereby remote boys were advantaged in that they experienced higher levels of math utility compared to their female peers. In contrast, the math utility gender gap amongst urban and regional students was much smaller. Surprisingly, the reverse occurred for senior high school enrolment, where regional and remote girls were more likely to enrol in high school STEM compared to their male peers. Again, the gender difference amongst urban students was much smaller.

Immigrant status. There were no significant interaction effects for immigrant status by gender with regards to self-beliefs, attitudes and achievement. There was, however, an effect between gender and immigration status in predicting senior high school enrolment. This effect showed that the effect of gender reversed for students born overseas, relative to students who had been born in Australia. In other words, girls born overseas were less likely to enter STEM than their male counterparts. However, it should be noted that girls from overseas were still more likely to enrol in senior high school STEM compared to Australian born girls.

Conclusions and Future Directions for Research

This study was important as it was the first study (to the author's knowledge) to examine the intersection between gender and Indigenous status for math attitudes, achievement and educational attainment in STEM. Furthermore, this study extended on previous findings by exploring gender differences across geography, immigrant status, social class, and educational attainment. Implications of these results for policymakers, educators and researchers are outlined below.

This study showed that Indigenous, low SES, and female students were particularly disadvantaged with regards to STEM educational outcomes. Educators should be aware that these groups face difficulty in studying STEM. However, educators should note that female students were the only group whose low self-beliefs and attitudes remained even when comparing students of equal ability. Thus, enhancing the self-beliefs and attitudes of young female students should be a key goal of STEM educators.

Overall, findings from Study 1 painted a complex picture with regards to the circumstances in which gender differences were largest. For instance, gender differences for math self-efficacy became larger amongst high SES students; however, educators should note that low SES girls still experienced the lowest self-efficacy relative to low SES boys, and high SES girls and boys. The gender gap also was substantial for students from remote locations, whereby girls experienced significantly higher levels of math anxiety, and lower levels of math utility value, compared to students from more populated regions. In contrast, boys from remote locations had very low levels of math anxiety and high levels of math utility value compared to students from more populated areas. Finally, analyses revealed that Indigenous boys were at particular disadvantage with regards to achievement and math utility value. Thus, educators should be aware that Indigenous boys in particular struggle to see the value of math for study and career, and that they also achieve at lower levels compared to their female peers (particularly in literacy). These findings suggest that Indigenous boys could benefit from interventions that place emphasis on enhancing math utility value (see Harackiewicz, Canning, Tibbetts, Priniski, & Hyde, 2016 for an example of a math utility value intervention in the context of ethnicity and social class class). Finally, achievement results indicated that more resources are needed to further support the academic growth and development of young Indigenous Australians.

This research has identified that gender differences can vary in important ways across a number of social and cultural contexts. Thus, researchers should be encouraged to investigate the intersection of gender, social class, and cultural contexts, particularly in light of current research gaps in this area. Nonetheless, there were important similarities in the experience of gender across different social and cultural groups. For example, most variables did not exhibit significant interactions between gender and other social/cultural factors.

Finally, the negative effect of gender on STEM self-beliefs and attitudes was more pervasive than any other category once achievement was taken into account in analyses. This signifies that even when girls have equal ability to boys, female students are still experiencing low levels of self-beliefs and poor attitudes towards math. Thus, it is critical that further research investigates the role of self-beliefs and attitudes in predicting long-term outcomes such as university enrolment and career choice, and the degree to which EVT variables can account for gender disparities in STEM educational attainment.

CHAPTER 7

YOUNG WOMEN FACE DISADVANTAGE TO ENROLMENT IN STEM COURSES REGARDLESS OF PRIOR ACHIEVEMENT, SELF-BELIEFS AND ATTITUDES

Chapters 5 and 6 focused on investigating the degree to which the magnitude of gender gaps in expectancy for success and values was influenced by social and cultural contexts. In contrast, Chapter 7 focuses exclusively on the second major research question of this thesis: “To what degree can EVT-related variables account for the gender gap in educational attainment in STEM coursework at high school and the university level?” Furthermore, this study extends upon current research by adding a qualitative component to this question, exploring the degree to which student responses reflect existing theory, and whether responses highlight alternative mechanisms behind the STEM gender gap that may previously have been historically overlooked by quantitative research.

Thus, Chapter 7 presents the methods, results, and discussion for Studies 3-4: “Young Women Face Disadvantage to Enrolment in STEM Courses Regardless of Prior Achievement, Self-Beliefs and Attitudes”. Results from Study 3 show that while Expectancy Value Theory can account for some of the gender disparity in STEM enrolment, there is still a very large amount of difference that remains unexplained by current theory. Furthermore, results indicate that even when comparing male and female students of equal ability and attitudes, females still are significantly disadvantaged in terms of STEM university enrolment. A content analysis in Study 4 explores whether open-ended interview data can add to current theory. Results point to the role of dimensional comparison as critical to educational choices, but again, there were no major themes that arose that significantly deviate from current theory.

Study 3: Young Women Face Disadvantage to Enrolment in STEM Courses Regardless of Prior Achievement, Self-Beliefs and Attitudes

The overarching aim of this study is to address the research gaps discussed in Chapter 3, by utilising a multi-method approach to understanding STEM educational attainment. By using complementary methods we: 1) identify the unique contributions of the currently known EVT-related predictors of STEM educational attainment in predicting senior high school and university enrolment in STEM; 2) identify the degree to which the gender effect on STEM educational attainment might occur via these predictors; 3) determine the strength of gender as a predictor of educational attainment in the STEM areas when previous achievement, self-beliefs, interests, values, anxiety, and teacher support are controlled for; and 4) explore whether corresponding qualitative responses of young people support quantitative conclusions, and whether they offer alternative explanations for possible mechanisms behind gender gaps in educational attainment in the sciences.

Achievement as a Predictor of Educational Attainment

In research, as well as the popular imagination, many have argued that women disengage from STEM simply because they lack the aptitude or ability in the areas of math and science. Indeed, there is a vocal community of scholars who argue that men and women are innately different in terms of their skills and abilities due to evolutionary and biological impacts on development (e.g., Baron-Cohen, 2003; Brizendine, 2006; Buss, 1991). But is there evidence to support the claim that gender segregation in STEM educational attainment is driven primarily by differences in aptitude or achievement? Longitudinal research shows that prior achievement is predictive of enrolment in STEM and completing a STEM degree (Crisp, Nora & Taggart, 2009). Moreover, on average women tend to outperform men in verbal related tests of ability, whilst men tend to outperform women in tests of mathematical ability (Falch & Naper, 2013; Lietz, 2006; Machin & McNally, 2005; Machin & Pekkarinen, 2008; Voyer & Voyer, 2014). However, this finding is not clear, as the relationship between gender and math performance appears to vary as a function of how achievement is operationally defined (Falch & Naper, 2013; see also Ceci, Williams, & Barnett, 2009 for a review of gender differences in achievement across varying methods of assessment). Thus, it is questionable whether this difference is substantial enough to warrant differential levels of achievement as being the most important factor behind gendered outcomes in school and work.

In fact, Hyde and colleagues' gender similarities hypothesis (for a review of the literature see Hyde, 2005) shows that there are more similarities than differences between genders for most psychological variables, and when differences do arise they are usually small in size. This holds true for tests of ability, with Hyde demonstrating that differences in male and female math ability are almost non-existent in terms of effect size. Thus, Hyde argues that historically gender differences are consistently over-emphasised in the literature, while the overwhelming evidence of gender similarities across psychological variables has been ignored.

Recent research on achievement has indicated that achievement across different academic domains might provide a prediction of eventual enrolment. For instance, Wang, Eccles, and Kelly (2013) found that individuals who were highly capable in mathematics, but who also possessed high verbal skills were less likely to pursue STEM careers than individuals with high math skills, but moderate verbal skills. Authors noted that the group with high math and high verbal ability included more females, and thus the issue of competing competencies provides evidence that achievement in verbal domains may steer students away from the pursuit of STEM careers. A possible mechanism behind this process might lie in the processes described in Dimensional Comparison Theory (e.g., Moller & Marsh, 2013) and Marsh's (1986) Internal/External (I/E) Frames of Reference Model which has outlined how comparisons of performance in different domains of achievement can have negative consequences on self-concept in the opposing area of achievement (e.g., high verbal achievement resulting in lower math self-concept). Thus, in the current study verbal achievement was included as a predictor of eventual STEM school and university enrolment to test how achievement in an opposing domain impacts on enrolment in STEM educational attainment in senior high school and university study.

The Importance of Self-Beliefs and Attitudes: EVT Related Predictors

If math and science achievement and aptitude have limited reliability in explaining why STEM remains a gender stereotyped field, what other mechanisms might be behind the gendering of STEM? As described in Chapter 3, one of the most influential theories to explain gendered educational and occupational outcomes is EVT (Eccles, 1994).

Expectancies for success. As mentioned in Chapter 3, there is strong evidence that there are consistent gender differences in math self-concept/expectancy for success favouring boys (e.g., Skaalvik & Skaalvik, 2004), suggesting that self-beliefs in mathematical ability could be critical in explaining the gender gap in STEM attainment. However, despite this there has been limited research on the role of academic self-beliefs in predicting university enrolment (Parker et al., 2014). For instance, self-efficacy is linked to higher academic aspirations (Bandura et al., 1996), and self-concept can predict university enrolment, even when controlling for prior achievement (Marsh, 1991; Parker et al., 2012). Recent research has confirmed these findings within the context of STEM; showing that math self-efficacy at Grade 12 predicts intention to major in STEM at college (Wang, 2013). Moreover, Parker et al. (2014) showed that math self-concept remains a significant predictor of STEM university enrolment, even when other demographic and achievement factors are controlled for. Finally, recently published work by Priess-Groben and Hyde (2017) found that self-concept of math ability was the only predictor of high school course-taking behaviours and course-taking intentions after prior math achievement was controlled for – however, interestingly, there were no statistically significant differences between girls and boys on self-concept (or other EVT variables) in this study.

Values. Expectancy value built on previous literature by highlighting the importance of value judgements in shaping the education and career decisions of young people (Eccles, 1994). Students do not only perform self-assessments of abilities when they are making choices about educational engagement, but also rely heavily on their value judgments about the activity they are pursuing. For example, students are more likely to choose to enrol in a STEM course if they think that science and math are interesting or enjoyable – described as intrinsic value in expectancy value literature. Again, Chapter 3 highlights the strong evidence for gender differences in interest in math (e.g., Preckel, Goetz, Pekrun, & Kleine, 2008). Furthermore, there is emerging evidence to suggest that interest and liking of math are the strongest predictors of math course selection in senior high school for Australian adolescents (Watt, Eccles, & Durik, 2006), and the number of math courses taken during high school and the decision to study STEM at university (Guo, Parker, Marsh & Morin, 2015).

In addition to interest, other value beliefs have also been found to be integral in predicting student choice behaviour. One of the most researched of these is utility value, commonly defined as the usefulness of a task in the immediate or long-term future (Wigfield & Cambria, 2010). Utility value has been discussed somewhat less in the literature on STEM university entrance predictions. However, recent work has indicated that these values alongside interest are important and are more powerful predictors than achievement for prior enrolment in high school STEM (Maltese &

Tai, 2011). In contrast, work by Priess-Groben and Hyde (2017) showed that utility value did not have a significant direct effect on predicting high school and college enrolment. Nonetheless, latent-class modelling by Musu-Gillette, Wigfield, Harring, and Eccles (2015) has revealed that students who are able to sustain high self-concept, interest, and perceived importance/usefulness of math across their education are more likely to pursue a college major with a moderate to intense amount of math, compared to those whose expectancy value declined over time.

Drivers of STEM disengagement: math anxiety. Other possible mechanisms behind gender differences in educational attainment that have been highlighted include math anxiety and costs of entering STEM (e.g., Perez, Cromley, & Kapaln, 2014). In expectancy value literature, cost refers to what an individual has to give up to engage in a task (e.g., studying for an upcoming chemistry test means I can't go out with my friends), and the anticipated effort that will be required for a task (Eccles, 1994). However, although an important component of Eccles and colleagues' expectancy value theory, limited research has been conducted in the area of cost in predicting educational outcomes. Recent research has begun to explore other drivers of disengagement in STEM, particularly in terms of emotional cost, stress and anxiety. A longitudinal study charting college chemistry students' intent to leave a STEM course found that perceptions of effort cost (e.g., drawbacks associated with time and effort, lost opportunities, stress, and anxiety) predicted intent to leave a chemistry degree (Perez, Cromley, & Kaplan, 2014). Importantly, research has demonstrated that female students are more likely to report greater levels of math anxiety (e.g., Devine, Fawcett, Szűcs, and Dowker, 2012). Thus, although the Longitudinal Study of Australian Youth (LSAY) does not collect data on overall costs of entering STEM, math anxiety was incorporated into analyses to account for the potential negative effects of emotional cost that may occur alongside the positive effects of expectancy and value.

Teacher support. Finally, the key aspect of expectancy value that is overlooked is the importance of socialisers in influencing decisions. LSAY collects data about perceived teacher support in math, and thus in this paper the teacher support variable is utilised as a proxy for socialiser influence of educational outcomes. Teacher support may be an important motivator for students, but also a buffer against negative stereotypes for students. Indeed, qualitative research has found that female students who enrolled in STEM university courses reflected that previous learning experiences were critical to their decision to enter into STEM, alongside with role models in the field and parental support (Bieri Buschor, Berweger, Keck Frei, & Kappler, 2014; Fouad et al., 2010). Furthermore, in a review of the literature and in interviews with students, Fouad et al. (2010) showed that teachers had a strong influence on students' perceptions of supports and barriers for pursuing STEM.

Research Questions and Hypotheses: Study 3

Despite considerable theory and research in the field (as detailed in Chapters 2 and 3), there is still little consensus on exactly which of the above EVT-related factors are the most crucial in predicting educational attainment in STEM, but also in explaining the gender gap in STEM enrolments (e.g., see Ceci, Barnett & Williams, 2009 for an overview). Thus, the central aim for Study 3 was to establish what EVT-related factors during middle adolescence were predictive of: a) selecting to study STEM in senior high school, and b) deciding to continue STEM study via enrolment in a STEM university degree. Moreover, I sought to determine what EVT-related factors were the most critical predictors of STEM educational attainment, in terms of their unique contribution to increasing likelihood of STEM course selection, and also in terms of how they mediate the effect of gender on educational attainment. Essentially, I aimed to explore how much of the effect of gender could be explained by currently hypothesised mechanisms behind the gender gap as outlined in EVT (e.g., past achievement, self-beliefs, values and teacher support). Hypotheses and research questions are as follows:

Hypothesis 1. Female gender and math anxiety will negatively predict course selection for STEM senior high school and university study.

Hypothesis 2. Math self-efficacy, self-concept, interest, utility value, math and science achievement, and teacher support at age 15 will positively predict course selection for STEM senior high school and university study.

Research Question 1. To what degree do EVT predictors explain the effect of gender on educational attainment in STEM senior high school and university study?

Research Question 2. Are there differences between the EVT-related factors that predict senior high school STEM study and enrolment in university STEM degrees?

Research Question 3. When previous achievement, attitudes, math-anxiety, interest, self-beliefs, and teacher support are controlled for, what is the remaining effect of gender in predicting educational attainment in STEM at high school and university study?

Research Question 4. In line with dimensional comparison theory, does verbal achievement negatively predict outcomes in STEM educational attainment, and EVT-related STEM predictors of educational attainment, in contrast to the positive predictions based on math achievement?

Method Study 3

Participants

The sample was taken from the 2003/Y03 cohort of Longitudinal Study of Australian Youth (demographics described in Study 2). Again, data for achievement, self-beliefs and attitudes came from Wave 1 of data collection whereby participant mean age was 15.69 years (*SD* = .29; *Range* = 15-16). Educational attainment data was based on senior high school course selection, where participants were in Grade 12 at age 18 (n = 6,658), and on university STEM course enrolment at age 19 (n = 2,235). Importantly, participants in the final data wave at age 19 were a subsample of the larger database, whereby only students who enrolled in senior high school STEM courses were asked follow-up questions regarding whether or not they continued their STEM studies at university (see Figure 11 for a visual representation). Importantly, readers should note that the Australian school system differs from the US educational system, whereby STEM courses are not compulsory for senior high school students.

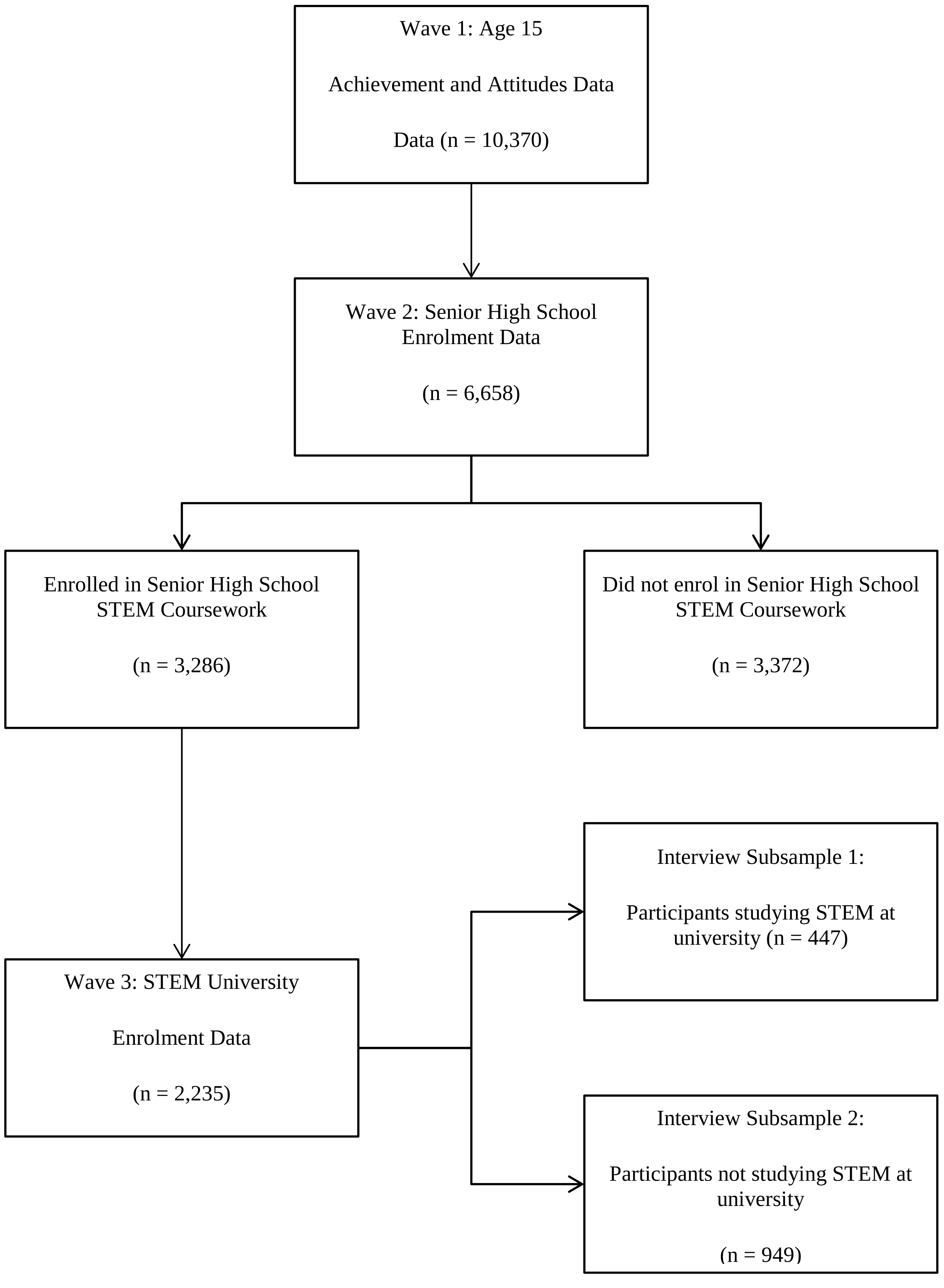

Figure 11. Flow diagram of data collection across time waves for Studies 3 and 4.

Measurements

Predictors of STEM enrolment. Measures in this study included the PISA indexes for math self-efficacy, math self-concept, math interest, math instrumental motivation (hereafter referred to as math utility value), math anxiety, perceived teacher support in math lessons, and gender. Each of these variables is described in detail in the previous chapter (Study 2), alongside reliabilities for each scale measuring self-beliefs and student attitudes (all $\alpha > 0.82$). A list of scale items is available in Appendix D. Finally, measures for math, science, and reading achievement are described in detail in the previous chapter, and information regarding the use of plausible values in analyses (as recommended by PISA) are described in Supplementary Materials.

STEM course selection. During Wave 5, participants who had previously provided data that they had enrolled in a high school STEM course (1 = yes; 0 = no) were asked if they were currently studying in a STEM field at a tertiary level. Those that were studying in a math or science field at a tertiary level were coded as one; those that had not were coded as zero.

Data Analysis

Analysis of the data was conducted using R (see Appendix for R scripts for each analysis). Analyses for prediction of senior high school STEM course selection and university STEM course selection were conducted independently of one another. In each analysis, the current known factors predicting STEM engagement were entered into the regression alongside gender in four steps: 1) each variable independently of one another; 2) gender independently predicting educational attainment; 3) gender and achievement predicting educational attainment; and 4) gender, achievement, and EVT-related variables predicting educational attainment. The aim of this was to identify the unique contribution of gender in predicting STEM educational attainment once other critical known achievement and EVT-related predictors had been controlled for. In other words, “how much of the effect of gender can be explained by currently known mechanisms of achievement, attitudes and teacher support?” In addition to the above, I ran a series of regressions to form a path analysis on the factors that mediated the effect of gender in predicting STEM senior high school enrolment and STEM university enrolment.

LSAY is a large longitudinal database that utilises complex sampling procedures in order to capture the experiences of a diverse range of young Australians. In order to do this, LSAY employs oversampling of some demographic groups. Consequently, I used sampling weights provided by LSAY to correct for any sampling bias in the analyses. Another complexity of the LSAY database is the issue of attrition due to the passing of time between data collection points. I utilised attrition weights provided by LSAY to take into account any bias from participant attrition. LSAY also has a complex structure with schools as the primary sampling unit. To account for this complex design, I used the 80 balanced repeated replication weights provided by the survey organisers to ensure correct standard errors (see Lumley, 2010). Finally, scales were standardised to a common metric (M = 0; SD = 1) to facilitate the comparison of parameter estimates (see Supplementary Materials for further details).

Results Study 3

Table 8 shows descriptive statistics for the variables used in this study. Male students scored higher on all measures except for math anxiety, and teacher support. The largest gender differences were in math self-efficacy (with boys reporting higher scores) and math anxiety (with girls reporting higher scores). Out of the 6,658 students who provided data for high school STEM course selection, 49.4% (n = 3,286) of students surveyed had enrolled in a STEM course during senior high school. At the fifth wave of data collection 2,235 of these students provided responses to whether they continued their STEM education into university. From this subsample of participants 54.6% (n = 1,221) of the young people surveyed reported that they were currently studying a science, engineering, math, or IT course at university.

Effect of gender on achievement and EVT related predictors of STEM attainment

Before considering the effect of gender on STEM attainment outcomes and on whether this effect could be partially explained by prior achievement and EVT related variables, I first considered whether gender predicted these achievement and EVT-related variables in expected directions. Results (see Table 9) showed that gender significantly predicted all variables except for teacher support in math class and science achievement. Being female had a positive effect on reading achievement, but also increased math anxiety. However, being female had a negative effect on math self-efficacy, math self-concept, math interest, math utility value, and to a lesser extent math achievement. Thus, these constructs are potentially useful in explaining any resulting gender differences in STEM attainment.

Table 8

Means, Standard Deviation, and Gender Difference for all Variables

	Mean Male	SD Male	Mean Female	SD Female	Cohen's d
Math achievement	529.22	99.49	518.56	92.82	-0.11
Science achievement	526.99	106.05	522.51	98.12	-0.04
Reading achievement	508.30	100.41	542.56	92.03	0.36
Math self-concept	0.28	0.87	-0.01	0.91	-0.33
Math self-efficacy	0.26	1.05	-0.11	0.87	-0.38
Math interest	0.14	0.93	-0.09	0.94	-0.25
Math utility value	0.34	0.95	0.12	0.96	-0.23
Math anxiety	-0.20	0.87	0.10	0.86	0.35
Teacher support in math	0.25	0.98	0.29	0.99	0.04

Notes. Positive values indicate female advantage. All data was collected at time wave 1.

Predicting Senior High School STEM Enrolment: Are Girls Really Less Likely to Enrol in High School STEM?

Univariate. Results from Table 10 outline the univariate and multivariate results for each variable predicting senior high school STEM course selection. As a general rule, past achievement in all domains (math, science, and reading) were the strongest predictors of senior high school STEM course selection. EVT related variables such as self-efficacy, self-concept, interest, and utility value also showed strong predictive power in predicting senior high school STEM course selection. Math anxiety had a negative effect on a student's likelihood of educational attainment in high school STEM. However, a critical finding was that gender was the *only* non-significant univariate predictor of senior high school STEM course selection. This finding was in direct contrast to the hypothesis that predicted female gender would be a negative predictor of senior high school STEM enrolment.

Table 9

Effect of Gender on Achievement and EVT Related Predictors of STEM Attainment

	β	SE
Math achievement	-0.11*	0.04
Science achievement	-0.04	0.04
Reading achievement	0.36*	0.04
Math self-concept	-0.32*	0.03
Math self-efficacy	-0.37*	0.04
Math interest	-0.24*	0.03
Math utility value	-0.23*	0.03
Math anxiety	0.34*	0.03
Teacher support in math	0.03	0.03

Notes. All scales are z-scored, and results indicate what could be predicted by each variable when considered alone. * $p = < .05$ Note that this table provides equivalent information to Cohen's d presented in Table 8.

Multivariate. In line with the first hypothesis, STEM achievement and positive math attitudes were associated with a greater likelihood of choosing STEM at senior high school. In particular, previous math achievement exerted the largest influence on STEM course selection once other factors had been controlled for. Attitudes were also significant predictors - in the final stage of the regression, math self-efficacy and utility value, and reading achievement uniquely predicted STEM high school course enrolment over and above other factors such as teacher support, math anxiety, and science achievement. In contrast to hypothesis I, results showed that once other variables were controlled for math anxiety no longer had a strong negative effect on choosing to study STEM.

Table 10

Predictors of Senior High School STEM Course Selection

	Univariate	SE	Step 1	SE	Step 2a	SE	Step 2b	SE	Step 3	SE
Gender	0.04	0.08	0.04	0.08	0.12	0.10	0.47*	0.08	0.29*	0.10
Math achievement	1.18*	0.05			0.83*	0.11			0.59*	0.12
Science achievement	1.11*	0.06			0.31*	0.13			0.21	0.13
Reading achievement	1.00*	0.05			0.15	0.11			0.30*	0.11
Math self-concept	0.73*	0.04					0.25*	0.05	0.12*	0.06
Math self-efficacy	0.90*	0.05					0.69*	0.04	0.26*	0.05
Math interest	0.54*	0.04					0.01	0.06	0.18*	0.06
Math utility value	0.57*	0.04					0.30*	0.04	0.35*	0.05
Math anxiety	-0.52*	0.04					-0.08	0.06	0.02	0.06
Teacher support in math	0.26*	0.03					0.02	0.03	0.02	0.04
Pseudo R-squared			0.00		0.19		0.14		0.23	
Area under the curve			0.50		0.71		0.68		0.73	
Predicted Probabilities for Gender			-0.01		-0.03		-0.12		-0.07	

Notes. Variables from age 15 predicting senior high school course selection in Grade 12. All scales are z-scored. Univariate effects are presented first, followed by the multiple regression. Multiple regression results are presented in steps: 1) gender; 2a) gender controlling for cognitive factors, 3) gender controlling for cognitive factors and non-cognitive factors. Each log-odd represents the unique contribution of a variable controlling for other variables in each step. * p = <.05.

Analysing the unique contributions of each variable on high school enrolment revealed that the effect of gender on high school STEM enrolment was suppressed by other variables. That is, as achievement and EVT-related variables were controlled for, female gender became a significant predictor of enrolment in senior high school STEM courses. In other words, when boys and girls of equal ability and attitudes were compared, girls were *more* likely, by seven-percentage points, to select a senior high school STEM course than their male peers. Again, this finding was in direct contrast to the hypothesis that girls would be disadvantaged at all levels of STEM attainment. Instead, two mechanisms seemed to be at play. One, whereby female students are negatively affected by attitudes and achievement in terms of their enrolment choices, and another unaccounted positive residual effect of gender that remained unexplained by expectancy value variables.

Effect of Reading Achievement on EVT-Related Predictors of STEM Attainment

There was evidence of suppression in relation to reading achievement that was not significant in Step 2 of the analysis, but once the EVT-related variables had been controlled for, became significant. It was hypothesised that this was due to factors associated with the internal/external frame of reference discussed above. Namely, that reading achievement has a positive effect on STEM outcomes in as far as it reveals underlying academic ability. However, as predicted by IE theory, reading achievement will have a negative effect on self-beliefs and values in math. If the individual is relatively better at reading than they are at math (i.e., a counteracting negative dimensional comparison effect). Thus, we would expect that controlling for STEM ability, reading achievement would have a negative effect on self-beliefs and values.

Table 11 showed that when science and math achievement were controlled for, reading achievement had significant negative effects on math self-concept, math self-efficacy, math interest, and math utility value. On the other hand, reading achievement positively predicted math anxiety. There was no significant effect of reading achievement for teacher support in math classes. Nonetheless, reading achievement may be significant to STEM entry through the effect it exerts on self-beliefs and attitudes towards math.

Table 11

Effect of Reading Achievement on Non-Cognitive Variables Controlling for Math and Science Achievement

	β	SE
Math self-concept	-0.33*	0.04
Math self-efficacy	-0.28*	0.04
Math interest	-0.26*	0.03
Math utility value	-0.19*	0.04
Math anxiety	0.27*	0.03
Teacher support in math	-0.01	0.03

Notes. All scales are z-scored, and results indicate what could be predicted by each variable when considered alone. * $p = < .05$.

University STEM Enrolment Among Students Studying Senior High School STEM

Univariate. The second set of longitudinal analyses focused on understanding what factors predicted university enrolment in STEM for high school students already enrolled in a STEM senior high school course (see Figure 11). Results from Table 12 show the effect of the predictor variables at age 15 predicting the decision to continue studying STEM at a university level at age 19 among those students who had already enrolled in senior high school science courses. Results showed a starkly different pattern of results for gender. In contrast to the high school model, the model predicting university STEM enrolment showed that gender was the strongest predictor, indicating that being female was negatively associated with enrolling in a STEM university course for students already engaged in STEM study at a high school level. Furthermore, results suggested that EVT-related variables at age 15 were more powerful at predicting university study choice than prior achievement. Thus, for the subset of students who had engaged in STEM study at high school, the decision to disengage from further STEM study after senior high school is influenced more by young people’s self-beliefs and attitudes, rather than their achievement levels. This finding makes sense given the subsample of students studying senior high school STEM. Indeed, amongst the group of students who were already relatively high achieving, achievement might become less of a critical factor in determining university entry. In contrast to the previous model, reading achievement did not have a significant effect on university STEM enrolment. Again, teacher support in math class had little effect on young people's enrolment decisions.

Multivariate. Multivariate results showed that even when controlling for previous achievement, attitudes, interest, anxiety, and teacher support in math lessons, young men were 1.52 times more likely to enrol in a STEM university course than young women. In other words, being female was associated with lower odds of enrolling in a STEM tertiary degree regardless of past achievement and beliefs. In contrast to the first hypothesis and the prior high school model, only a few factors remained significant predictors of STEM enrolment in the final model that controlled for all variables. As mentioned above, gender was the strongest predictor of STEM enrolment. Utility value and interest in math at age 15 were the other two significant predictors of whether or not a student will eventually enrol in STEM at university once other variables had been controlled for.

How Much Variance Does Current Theory Explain for Gender Differences in University STEM Enrolment?

Our final research question focused on to what extent can current theory explain gender differences in university STEM enrolment for participants who undertook senior high-school STEM subjects (i.e., those who already had positive achievement and attitudes related to STEM fields that led to STEM enrolment in high school). Taken together, the results suggest that traditional variables focusing on achievement and expectancy value theory do explain STEM enrolment to some degree for this population (see Figure 11). In the first model, males have a 0.15 greater probability on entering tertiary STEM than females. However, accounting for achievement and EVT-related variables, this was reduced by a third to 0.10. While this suggests these variables accounted for a considerable amount of the gender effect, the vast majority of the gender difference for students already enrolled in high school STEM remained unexplained. Thus, research needs to consider additional reasons that could explain this gender gap. A powerful way of doing this is by asking participants of different genders directly why they did or did not chose to study STEM at a tertiary level.

Table 12

Predictors of University STEM Enrolment

	Univariate	SE	Step 1	SE	Step 2a	SE	Step 2b	SE	Step 3	SE
Gender	-0.59*	0.11	-0.59*	0.12	-0.50*	0.12	-0.43*	0.11	-0.42*	0.12
Math achievement	0.25*	0.06			0.23	0.12			0.10	0.12
Science achievement	0.19*	0.06			0.06	0.11			0.04	0.12
Reading achievement	0.07	0.09			-0.11	0.16			-0.01	0.15
Math self-concept	0.37*	0.05					0.11	0.10	0.09	0.10
Math self-efficacy	0.28*	0.05					0.04	0.07	0.00	0.08
Math interest	0.40*	0.05					0.18	0.10	0.19*	0.09
Math utility value	0.39*	0.05					0.24*	0.06	0.25*	0.06
Math anxiety	-0.27*	0.05					-0.02	0.08	-0.02	0.08
Teacher support in math	0.08	0.05					-0.05	0.06	-0.04	0.06
Pseudo R-squared			0.02		0.02		0.05		0.05	
Area under the curve			0.56		0.56		0.60		0.60	
Predicted Probabilities for Gender			0.15		0.12		0.11		0.10	

Notes. Table 12 shows variables from age 15 predicting university STEM enrolment at age 19. All scales are z-scored. Univariate results are presented first, followed by multiple regression. Multivariate results are presented in steps: 1) gender; 2a) gender controlling for cognitive factors, 3) gender controlling for cognitive factors and non-cognitive factors. Each log-odd represents the unique contribution of a variable controlling for other variables in each step.
* p = <.05.

Study 4: Interview Data

Results from Study 3 showed that gender was a critical factor in young people's likelihood of continuing STEM study from senior high school to university level. Furthermore, EVT-related factors were the only variables to independently predict STEM enrolment over and above achievement, showing that for students already engaged in STEM education at senior high school, expectancies for success and value are critical in the decision to continue with STEM study at a tertiary level. Likewise, these variables differed by gender and may thus explain some proportion of the gender effect. However, results from Study 3 highlight that much of the variance in choosing STEM at university is not currently explained by the currently hypothesised factors in the literature. For this reason, this study takes a more detailed qualitative look at gender differences in the reasons young people engage or disengage from STEM university study to explore this unexplained variance. Finally, I hoped to take a more in-depth look at gender differences for expectancy value constructs as it relates to decision making about what university course to pursue.

Research Questions and Hypotheses: Study 4

Study 4 used a mixed method approach to extend upon and refine existing theory. More specifically, I hoped to identify new and novel themes in the self-reported attitudes of young people regarding their choices to engage or disengage from STEM study, that are not adequately represented in the current body of literature. Furthermore, I sought to explore whether there were substantial differences between the responses of young women and young men in terms of frequency and also content.

The research aims and questions guiding the current study were as follows:

1. Identifying the reasons for engagement and disengagement from STEM study:

1a. What are the most common reasons given by young people for enrolling in tertiary STEM education? Does this differ by gender?

1b. What are the most common perceived barriers, as reported by young people, which are perceived to discourage enrolment in tertiary STEM education? Does this differ by gender?

1c. What do students feel would have to change for them to reconsider enrolling in a tertiary STEM course? Does this differ by gender?

2. Replication and extension of current theory:

2a. Do the responses of young people reflect what is known from the existing body of literature? (E.g., Do students spontaneously refer to expectancies for success, and subjective task value in their reasons for choosing science when unprompted?) Are there any gender differences in how well responses from male and female students fit with previous theory?

2b. Do the responses of young people indicate that there are new and/or under-recognised facets to student motivation, that have not been fully covered by current theory? Are the alternative explanations similar for both male and female students?

2c. Do the responses of young people reveal that some factors are more commonly cited by young people as factors determining educational choice? (E.g., Do students frequently cite one aspect of task value as their perceived reason for choosing science, or fail to mention other areas?) Does this differ by gender?

2d. Is there evidence that young people engage in dimensional comparative processes when making career choices? (E.g., Choosing humanities over STEM, because of higher perceived interest of competence in comparison to STEM subjects). Does this differ by gender?

3. Exploration of social and cultural influences on STEM engagement:

3a. Are there instances in which young people spontaneously refer to gender role issues as determinants of their engagement or disengagement from STEM?

Method Study 4

Participants

Participants for Study 4 came from a subset of the above sample (see Figure 1). Data was taken from Wave 5 (2007) of the LSAY data collection, meaning that respondents were approximately age 19 at the time of being surveyed. This subset was made up of young people who in a previous time wave had indicated that they had studied a STEM course in high school (consisting of 13.46% of participants from data collected at age 15). Young people who had indicated ‘yes’ to this question were then directed through one of two series of questions depending on whether they had enrolled in a STEM course at tertiary level. Thus, the data for Study 4 can be divided into two subsets of eligible participants: a) young people (45.90% female) who indicated enrolment in a STEM course *n* = 447; and b) young people (60.50% female) who indicated they had not enrolled in a STEM course, *n* = 949 (see Table 13).

Procedure

Data were collected from participants through a structured telephone interview as part of the LSAY survey used in the prior study. Depending on whether or not the young person indicated they were enrolled in a tertiary science course, they were asked one of two series of closed and open-ended questions (see interview schedule in appendix). Close-ended questions (see Supplementary Materials for responses) required participants to assign a numerical value from 1-6 (*1 = very important, 2 = important, 3 = neither important or unimportant, 4 = not important, 5 = not at all important, 6 = don't know/can't say*), indicating the degree to which they endorsed particular statements in relation to choosing to study STEM or choosing to not study STEM *(e.g., You were influenced by your parents*). Responses to open-ended questions (e.g., *What would need to change for you to consider choosing to study science, engineering, maths or IT?*) were transcribed by interviewers, and responses were kept in written-form alongside a participant ID number that corresponded to the rest of the quantitative LSAY data. In order to obtain participant demographic information, demographic data from the quantitative dataset was merged with qualitative data and matched according to participant ID.

Table 13

Demographic Data for Young people Studying STEM University Courses (n = 447) and Young People not Studying STEM (n = 949)

	% Sample 1	% Sample 2
Female	45.90	60.50
Indigenous	0.90	2.20
Rural/remote	23.30	23.40
Blue collar highest skilled parent	8.30	7.80
First generation immigrant students	27.30	26.20

Note. Percentages displayed represent the percentage of participants from each subsample (e.g., 45.90% of Sample 1 identified as female, while 60.50% of Sample 2 identified as female).

Approach to Analysis

A content analysis methodology was used to analyse the open-ended responses of the young people participating in the study. Content analysis enables researchers to systematically describe and quantify phenomena (Downe-Wamboldt, 1992). The analysis involves coding participant responses into categories that summarise the content of the data and evaluating the frequency of these categories (Wilkinson, 2000). Content analysis has been criticised for being overly simplistic or reductionist in nature, however modern content analysis goes beyond a simple frequency count (Downe-Wamboldt, 1992). Instead, content analyses are concerned with the context of the data, and a major goal of analysis is to enhance the interpretation of results by relating categories to the context of the data, and the environment in which the data was produced (Downe-Wamboldt, 1992). Furthermore, content analysis is a flexible and complex data analysis strategy that can used on a diverse range of data, alone or in conjunction with other methods, and can be modified to suit both theory and data driven approaches to analysis (Elo & Kyngas, 2008).

Data Analysis

Data were collected from all participants during structured telephone interviews, and participant responses were transcribed word-for-word by the research staff administering the survey. Tests of mean differences were conducted across groups to establish the degree to which certain participant groups endorsed the statements provided to them about factors affecting their decision to enrol or not enrol in a STEM course. The qualitative analysis process that followed can be clearly divided into two separate strategies. The first focused on a deductive approach to analysing the data, in which coding was guided by the expectancy value model (see coding manual in supplementary materials for coding framework). The second strategy required a more exploratory approach, and thus, an inductive data-driven coding process was implemented.

Both inductive and deductive analyses have similar preparation phases (Elo & Kyngas, 2008). Thus, the first stage of the analysis for both studies was to decide on the unit of analysis (as recommended by Cavanagh, 1997; Guthrie et al., 2004). In this case, the units of analyses were the sentences or short responses that each participant gave in response to their respective questions. Coding focused on recognising themes and patterns within these responses. Themes were coded each time they were raised in response (for example, if a young person said they chose science because they found it interesting *and* they wanted a job with good income, the response was coded once for intrinsic value and once for utility value). Thus, the frequency counts for this study reflect the number of instances a particular category (e.g., intrinsic value) was referred to throughout the dataset. This approach was chosen because I was interested in the different themes raised by students, as opposed to the frequency of responses at the student level.

The next stage of analysis involved a repeated reading of the data in order to achieve immersion and a sense of familiarity with the data as a whole (Tesch, 1990). The key aim of this process is to gain an overall sense of 'what is going on' in the data before beginning the coding process (Elo & Kyngas, 2008). Once immersion had been achieved, I proceeded to analyse the two datasets separately (i.e., those who had enrolled in STEM courses, and those who had not).

Results Study 4

Analysis 1: Factors Motivating Enrolment in STEM

Common themes. Responses from young people who had studied STEM at university show that, in line with the results of Study 3, students overwhelmingly report that they were motivated to enter STEM primarily because of EVT-related variables, in particular intrinsic and utility value of STEM. Intrinsic value included responses that focused on enjoyment, interest, and curiosity about STEM study or careers. Responses coded for utility value were subdivided between three major themes. Students often cited the belief that STEM would provide them with study and career opportunities, and that they were motivated to pursue STEM for financial gain. A smaller number of students reported that they saw STEM study as a pathway to a desirable lifestyle. Reflecting the results of Study 3, self-belief, or students' expectancy for success in STEM was not a frequent theme in young people's responses. Other themes in student responses were attainment value (i.e., describing STEM as a long-held life goal or passion); influence of family and friends; previous exposure to STEM in either education, work experience or outside of school; concern for society and the environment; and entering STEM due to a lack of other options.

Gender differences and similarities across responses. Overall, both male and female participants show similar patterns of responses, with the exception of a few themes. There were similar percentages of male and female students endorsing intrinsic value, utility value (study/career opportunities and pathway to a good lifestyle), expectancy for success, previous exposure to STEM, family/friend/mentor influence, and attainment value as the reason behind enrolling in a STEM course at university. A chi-square significance test showed that male students were more likely to report financial gain as being a motivator of their decision to study STEM at university (χ^2 *(1, 447) = 11.83, p = .00)*. In contrast, young women were statistically more likely to say that they entered STEM studies because of a concern for society and/or the environment (χ^2 *(1, 447) = 8.40, p = .00)*, or that they felt they had no other options (χ^2 *(1, 447) = 5.69, p = .02)*. There were slight but non-significant gender differences in other themes (e.g., family, friends, and mentors; attainment value). Thus, overall results showed that there were more gender similarities than differences across the responses of students in relation to why they had chosen to study STEM at university.

Relation to previous literature and alternative mechanisms. Responses indicate that young people's personal reflections of their own decisions are congruent with longitudinal models of STEM university enrolment in Study 3, and to some degree with current theoretical models (e.g., expectancy value theory). A key difference, however, in these responses was that there were only a very small number of young people who cited high self-beliefs as a motivating factor in entering STEM study at university. One noteworthy finding was the importance of family and friends in influencing young people's decisions to choose to study STEM at university. Furthermore, it appeared that previous exposure to STEM was also an important factor for students in choosing STEM. These findings provide alternative explanations that are often not included in longitudinal models predicting educational attainment (particularly previous out of school exposure to STEM), however, are integral to the formation of self-beliefs and values according to expectancy value theory.

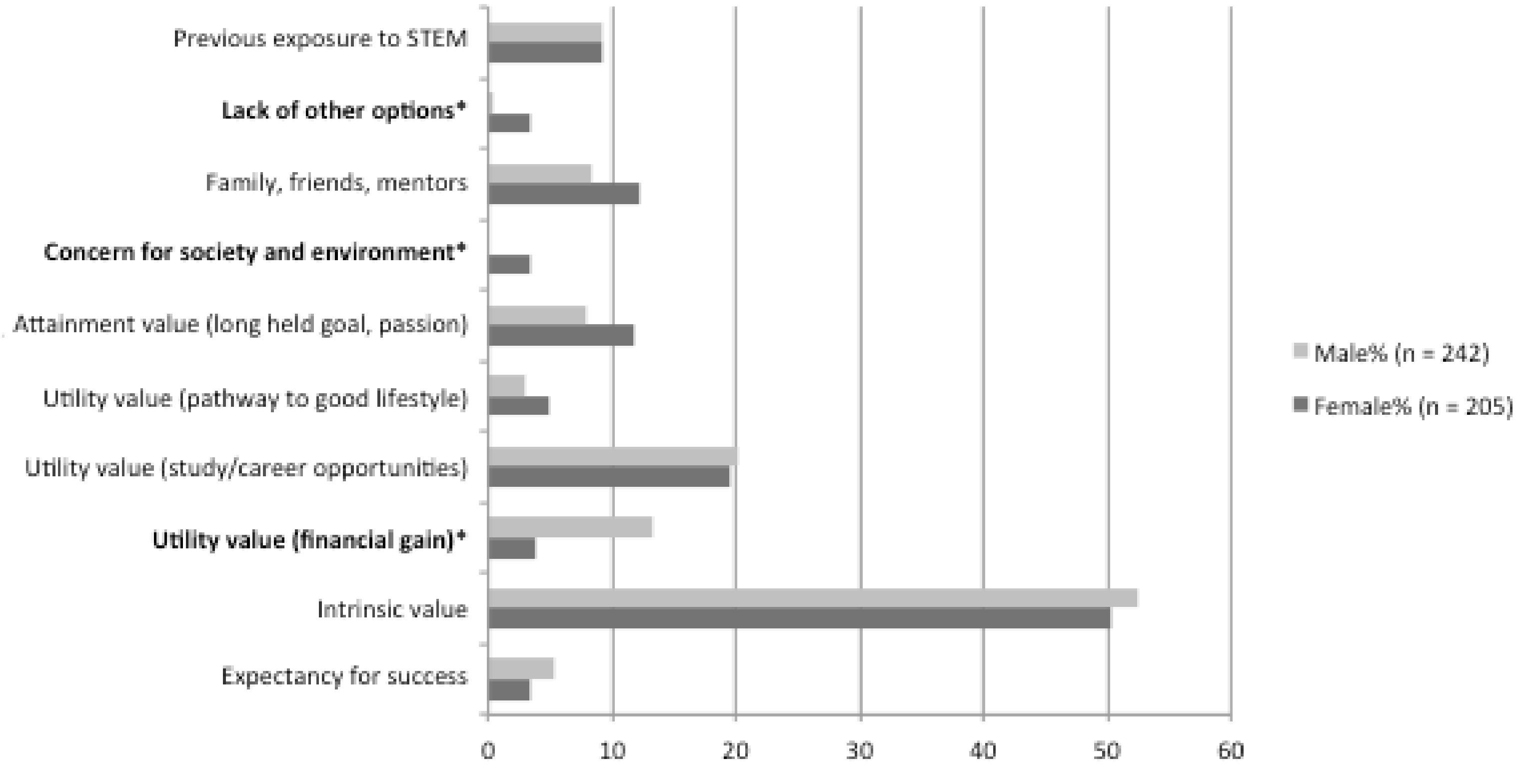

Figure 12a. Self-reported factors that encouraged STEM enrolment at university. * p = < .05.

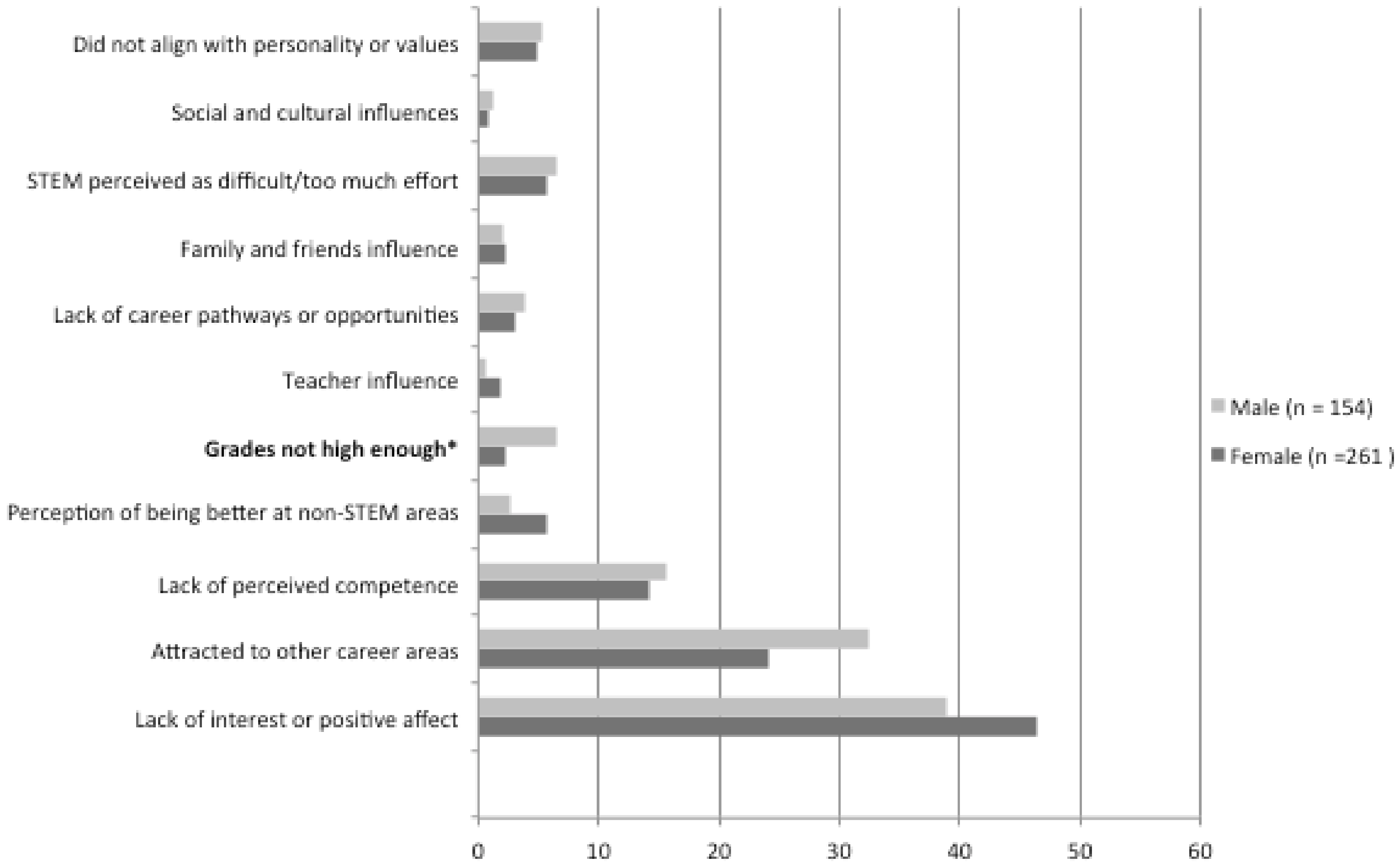

Figure 12b. Self-reported barriers to STEM. * p = < .05.

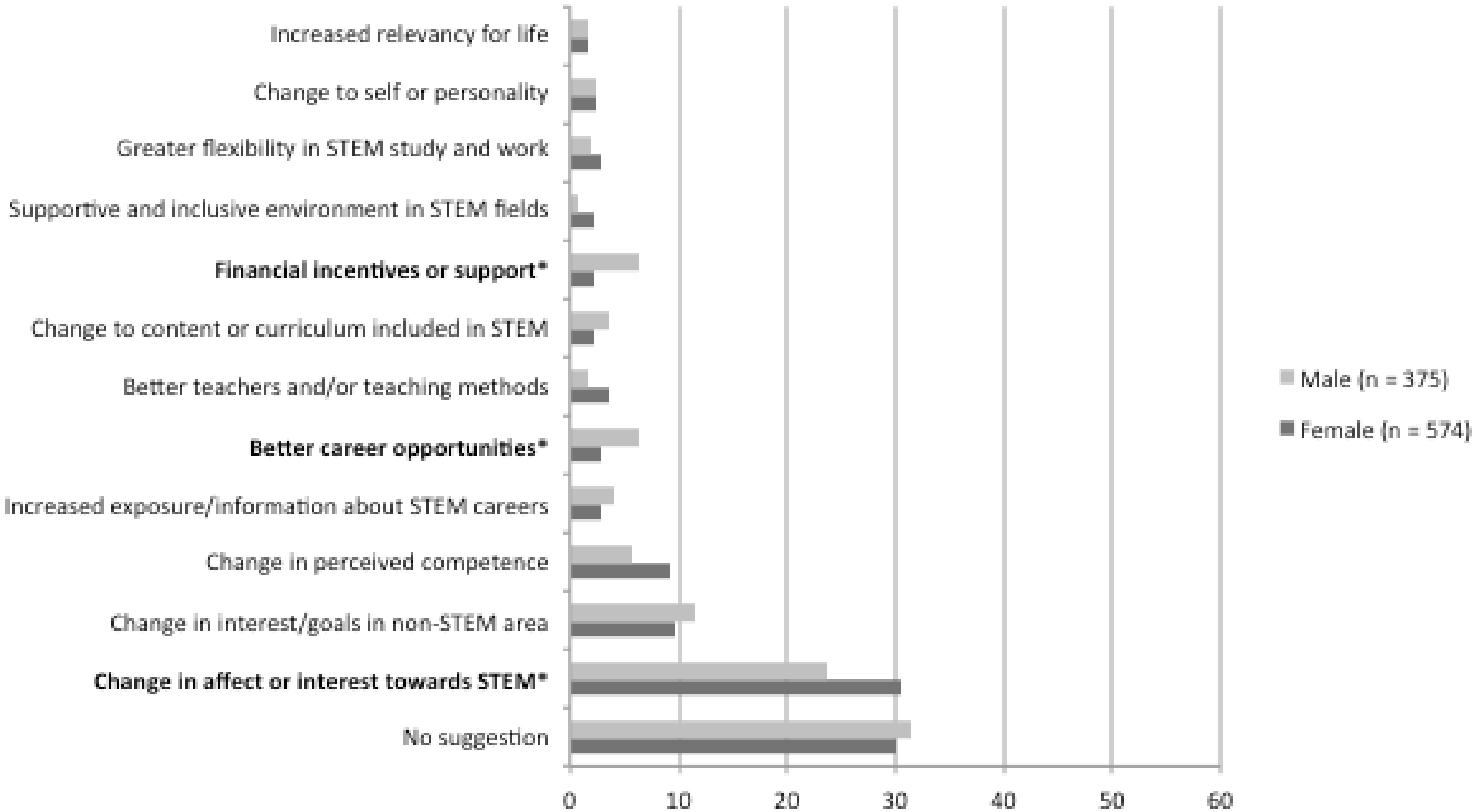

Figure 12c. Significant gender difference for what needs to change for students to consider STEM enrolment * $p = < .05$.

Analysis 2: Perceived Barriers to Further STEM Study

Common themes and alternative explanations for STEM disengagement. In keeping with the above findings, a lack of interest or positive affect towards STEM was the most frequently cited reason for choosing not to study STEM at a university level. A key finding though, was the substantial number of young people who responded that they were drawn away from STEM due to competing interests in other non-STEM areas. For example, one young woman reflected, *"I was interested in those areas, but as a career, I wanted to teach special needs children instead."* Importantly, like many other current quantitative studies in the area, there was no measure of this in Study 3, suggesting there is a need to develop measures around these explicit trade-offs. Interestingly, there was little mention of a lack of utility value as being a perceived barrier to entering STEM. This was in contrast to Study 3's longitudinal findings on utility value as being central to predicting STEM educational attainment.

In line with previous literature, a lack of perceived competence in STEM was the third most commonly cited theme. Again, a pattern of dimensional comparison influence seemed to arise, whereby students reported that they perceived themselves to have better academic strengths in non-STEM areas. The following two responses by young women reflected these overall themes:

> *"I just chose to continue with my strongest subjects and what I enjoy most. I love literature and reading"*

> *"…. Because I performed better in other subjects e.g., English and History than I did in science and maths subjects."*

Other key themes that emerged in the data were that STEM was perceived as a difficult and demanding discipline to study at university, and many students expressed concerns that the effort required or the study load would be too much. Participants also commented that STEM did not align with their personality or values.

Gender differences and similarities. Again, overall there were more similarities than differences in the responses of male and female participants. The only statistically significant difference was for the theme "grades not high enough"

(χ^2 *(1, 415) = 4.60, p = 0.03)*, whereby male participants were more likely to say they did not enter into STEM due to low scores on assessments (e.g., tertiary entrance ranks, class grades. *Note:* in Australia entry into university courses is subject specific, whereby different degrees have different entry score requirements). There were other differences across gender for some themes (e.g., attracted to other career fields, and lack of interest/positive affect towards STEM), however, none of these were statistically significant. Again, the overwhelming picture was one of gender similarities as opposed to gender differences.

Analysis 3 Results: What Would Need to Change?

Common themes and alternative explanations for STEM engagement. Most young people in the subsample were unsure or unable to provide a suggestion for what would need to change in order for the students surveyed to consider studying STEM at university. Of those who did provide a suggestion the most common themes centred around changes in interest. Over 30% of young women and 24% of young men said they would need their personal affect or interest levels to change in order for them to consider studying STEM. These findings are consistent with findings in Study 3 that showed interest as a crucial factor in predicting university entrance to STEM. Again, dimensional comparison type responses emerged, with a number of young people recounting that a change of interest or career goals in a non-STEM area would be required for them to consider studying STEM. Alongside the responses for perceived barriers to STEM, these findings suggest that dimensional comparison processes may be an underemphasised factor, critical to the formation of young people's career and study career plans.

Other themes that were coded in responses included: a change in perceived competence in STEM; increased information or exposure to STEM; provision of better career opportunities; financial incentives or support to study STEM; improved teaching; changes to STEM curriculum; supportive and inclusive STEM environments; greater flexibility in studying and working in STEM; increased relevancy for daily life; and a change to self or one's personality.

Gender differences and similarities. Overall, there were more similarities than differences in young people's suggestions for change. The only exceptions were female students being statistically more likely to report a change of interest or affect towards STEM. Male students were statistically more likely to report that financial incentives (χ^2 *(1, 949) = 11.54, p = .00)* and better career opportunities (χ^2 *(1, 949) = 6.49, p = 0.01)* would be useful in encouraging more students to study STEM at university. A larger share of female participants was also more likely to report a change in perceived competence as being required to consider enrolment in a STEM course, however, this difference did not reach statistical significance (χ^2 *(1, 949) = 4.17, p = .04).*

Overall Summary of Themes and Alternative Mechanisms

Overall, responses indicated that the strongest motivators for engagement and disengagement from STEM study were in keeping with EVT constructs. The strongest themes were intrinsic value and utility value. One noteworthy finding was that expectancy for success (or lack of) did not feature heavily throughout the responses. However, perhaps a more pertinent question, given the amount of unexplained variance in the previous study, is whether the interview data revealed any alternative mechanisms that could explain both the effect of gender, and also STEM university course selection.

The most convincing alternative explanation for explaining STEM course selection was the presence of dimensional comparison processes and competing interests. In other words, a number of young people stated that they engaged in an internal comparison process whereby they compared their interests or competence in STEM fields to their interests or competence in the humanities. Hence, it was not so much negative perceptions in relation to STEM, but rather positive perceptions in relation to non-STEM alternatives that motivated students to disengage from university STEM.

Only a small number of students raised concerns over gender issues when recounting their decisions to either enter STEM or their perceived barriers to entry. However, although analyses have focused on recording frequencies of responses, most research on interview data aims to go beyond merely counting numbers - themes with lower frequency counts can still be practically significant. Thus, although gender did not emerge as an independent theme in and of itself, it is pertinent to note the experiences of young women who, without prompting or questioning, spontaneously recalled explicit accounts of gender being a barrier to their entry in STEM at university. For example, one interviewee's reply to her being asked about the reasons she did not enter STEM was, *"Male dominated, and lots of facts and figures"*. There was evidence that some young women are aware of gender imbalances, but used this as motivation to enter STEM, *"There are not many women in that area so we need to boost the numbers in that area"*. However, other young women indicated that gender issues needed to change for them to consider STEM study. When asked what would need to change in order for them to consider a STEM degree one young woman said, "*Making it easier for women to work in those fields,*" and another, *"I think you'd need it to be a bit more girly – it's a very man oriented job"*. Finally, one young woman reflected, *"Maybe more support from the industry, there is not much support for people in those fields, for females such as engineering for men, like work experience in this field – I've tried and been knocked back"*. Again, I am cautious to not over-emphasise these responses in relation to the overwhelming responses referring to interest and utility value factors as crucial to influencing their study decisions. However, the role of gender roles and how they influence the degree to which young women feel comfortable in STEM remains a possible factor in accounting for the effects of gender that are unexplained by individual level factors such as achievement and attitudes.

Discussion Study 3 and 4

Overall, the results of Studies 3 and 4 illustrate the practical significance of EVT-related variables like values, interests and self-beliefs in determining young people's engagement and disengagement with STEM at senior high school and university levels. Achievement across all domains and positive attitudes toward math were associated with a greater likelihood of choosing STEM coursework at senior high school. However, results highlighted the complexities underlying the relationships between gender and STEM educational attainment. While Study 3 shows that traditional individual-level variables (e.g., achievement, self-beliefs, interests and values) can explain at least some of the gender gap university STEM enrolment, there still remains a substantial amount of variance that is unexplained by current theory. Moreover, results for high-school STEM enrolment reveal that gender does not just have a singular consistent effect but that competing mechanisms may be in operation for at least some STEM related choices. Furthermore, considering qualitative data about individual self-reported reasons for choosing or not choosing STEM university courses provided critical insights but only incrementally added to our understanding of gender differences in this area. In this discussion, I summarise the key findings from Studies 3 and 4, and discuss their significance to theory and practice in STEM education.

Girls of Equal Ability and Attitudes are More Likely to Enter Senior High School STEM

Perhaps the most unexpected finding was that gender was the *only* non-significant univariate predictor of senior high school STEM enrolment. This finding is in contrast to a wealth of literature that has speculated female representation in STEM decreases during high school as part of the ‘leaky pipeline’ (e.g., Sells, 1976). Furthermore, even more surprising was that once prior achievement and EVT-related variables were controlled for, female gender was positively related to senior high school STEM course selection. These results suggest that for boys and girls of equal abilities and attitudes, girls are more likely to enrol in high school STEM, despite assertions that the gender gap in STEM careers begins with a lack of engagement in the senior levels of high school math and science.

What factors are implicated in the relationship between gender and senior high school STEM enrolment? The significant effect favouring girls once achievement and attitudes have been controlled for suggests competing mechanisms.

Mechanism favouring boys. Further analyses of indirect effects of gender show that traditional predictors of educational attainment (e.g., achievement and attitudes) can explain at least some of the effect in gender. A strong pattern of gender stereotypical effects was evident (males stronger in math self-concept, self-efficacy, interest, utility value, and to a lesser degree math achievement; females higher in math anxiety and reading achievement). This pattern was reflected in the multiple regression whereby attitudes and achievement accounted for a notable change in predicted probabilities for entering high school STEM.

Mechanism favouring girls. However, it is also important to note that results suggest that there was also a counter-mechanism at play that lead to girls being more likely to enrol in high-school STEM once attitudes and achievement had been taken into account. What can explain the residual positive effect of being female and enrolling in STEM that is unaccounted for by prior achievement and attitudes? I speculate that it may be due to the utility value that advanced courses in STEM may have a pay-off in terms of boosting students’ university entrance marks (i.e., that preferential weighting of good performance in STEM subjects compared to other options make choosing these subjects attractive). The fact that female students have greater aspirations and motivation for university attainment compared to that of their male peers is well supported (e.g. OECD, 2009; Schoon, 2010). As such, female students may be more motivated to take advantage of strategies that boost their chance of entering university given the propensity for young women to be more likely to aspire towards university than young men. Clearly this remains speculative until the strategic aspect of course selection is better understood.

Predicting University Enrolment: Explaining the Emergence of Gender Gaps in Attainment

Our longitudinal analyses showed that there were critical differences in the factors that prompted continued engagement from senior high school science/math to STEM at a university level. For students already enrolled in senior high school STEM, university STEM enrolment was not significantly related to prior achievement in middle adolescence – a finding that was unsurprising given that achievement was a key predictor of senior high STEM enrolment. EVT-related variables remained strong predictors of STEM enrolment, despite the 5 year gap between predictor and outcome variable data collection. In particular, math utility value and interest in math at age 15 were key predictors of choosing to study STEM at university. However, the relation of gender to university STEM enrolment was perhaps the most noteworthy finding. In contrast to high school STEM selection, young women who had enrolled in high school STEM classes were far less likely to enrol into STEM at university than young men. I emphasise again that this analysis was a sub-sample of those who chose to do STEM related courses in year 11 and 12 (however, I also emphasise that undertaking STEM in senior high-school is often a requirement for undertaking STEM at university). As I controlled for prior achievement and attitudes, this negative relationship remained strong declining from a 15-percentage point gap to a 10-percentage point gap. Thus, young women had substantially lower odds of entering a STEM university course, even when I compared young men and women of equal abilities, similar self-beliefs, interests, and attitudes towards math. This suggests that current theoretical models of gender and educational choice are lacking and that innovation in theory is required. I suspected that qualitative interviews with participants could shed some light on where theory could be developed. These results are discussed below:

Why did Results Differ for High School Enrolment and University Enrolment in STEM?

This study has tracked students from high school to university STEM enrolment amongst a sample of students who had already enrolled in senior high school STEM courses. Thus, for the first dependent variable (choosing STEM at high school) I looked at students choosing to study STEM outright, while the second sample (choosing STEM at university) followed only the students who had previously indicated they had enrolled in STEM at high school. Thus, it is no surprise that variables such as achievement and attitudes become less important in further continuation of STEM study, as most students enrolling in senior high school STEM were already relatively high in achievement and attitudes towards STEM. Thus, it seems that after drawing from a pool of relatively strong achievers with positive attitudes towards math, there would be different factors that become more important in terms of who continues to engage in STEM at a university level. Put simply EVT-related variables may have less predictive power when considering STEM transitions for already high achieving and engaged students.

Furthermore, the choice of choosing what to study at high school, and what to study at university are essentially very different decisions. In Australia students often choose subjects to study in senior high-school in order to maximise their chances for a high university entrance rank (this score is used to assign university places in Australia and consists of a combination of school assessment and standardized testing). In contrast to the United States, Australian university course selection requires a student to make a forced choice, whereby they must choose to focus on a specific discipline. I contend that this may explain why there is such a stark change in the relationship between gender and STEM enrolment as tested across high school and university course selection. Essentially, young women may be more likely to enter into STEM in high school when they are able to study STEM alongside other subjects, however, when forced to choose one discipline of study, young women on average are more likely to gravitate towards the humanities.

Young People's Perspectives on STEM Engagement and Disengagement

Our longitudinal findings showed that the most critical factors in determining whether young people continue to remain engaged in STEM study were utility value, interest and gender. However, changes in predicted probabilities indicated that there still remains a great deal of unexplained variance in attempts to model STEM university enrolment with prior attitudes and achievement. Analysis of interview data aimed to identify potentially overlooked factors important to young people's engagement with STEM study.

Overall, interview data supported the initial conclusions drawn from longitudinal analyses in Study 3. Namely, young people responded that interest and utility value (e.g., financial gain, and career opportunities) were critical to their decisions to continue studying STEM at a tertiary level. This trend continued to young people's discussions of perceived barriers to STEM study, whereby lack of interest was a key theme.

However, perhaps a more critical question was whether the interview data provided alternative explanations for STEM engagement and disengagement. After analysing the responses of young people, several new lines of enquiry for further research have been identified. Namely, responses of young people illustrated the significance of dimensional comparisons in not only self-assessments of ability, but also interests and career goals in determining choice behaviour. In line with previous literature, it seems that young people are deterred from STEM, not simply because they don't like or enjoy STEM, but because they have competing interests elsewhere. To some degree there was a similar pattern in Study 3, whereby reading achievement negatively impacted self-beliefs and attitudes towards math, but also had negative (but not significant) effect on STEM university entry. Thus, quantitative measurements of STEM perhaps need to be further developed, in order to directly tap into dimensional comparison processes. Furthermore, other factors such as the importance of friends and family members in choosing to study STEM illustrate the need for further research to fully investigate this relationship in more depth.

Did Interview Data Explain the Unexplained Gender Effect of Study 3?

Overall, there were more gender similarities than differences amongst the responses of young people. The only exceptions were that young men were more likely to report than they were motivated by financial gain associated with STEM study, while young women were more likely to be motivated by concern for society and the environment, or that they entered into STEM because they felt they lacked other options. In terms of barriers and factors that would need to change to encourage STEM enrolment, males were more likely to report low grades as the reason they did not enter STEM. Finally, males were more likely to report that financial incentive could have changed their decision to disengage from STEM study, while more young women were likely to report that a change in interest towards STEM would be required for them to consider studying STEM at a tertiary level. Even these differences, however, were small and the overwhelming story from interview data was one of gender similarities, as opposed to differences.

Thus, an important question remains unanswered. If gender differences cannot be explained by EVT-related variables in longitudinal analyses (Study 3) or interview data (Study 4), then what factors can account for the effect of gender that is unexplained by prior achievement, self-beliefs and attitudes towards STEM? Indeed, most of the arguments for biological differences as drivers of gendered choice focus on ability and interests. Likewise, psychological research coming from a gender socialisation standpoint has focused on self-beliefs, interests, and values as key mechanisms behind gendered choice behaviours. However, the present investigation shows that the substantial gender effect in university coursework selection cannot be fully explained by individual-level characteristics such as ability and attitudes, and a large proportion of the gender effect for STEM university entry is unexplained. The two-third residual unexplained gender effect for university level STEM found in this research then might have its roots in structural inequality. For example, more abstract concepts such as gender roles, STEM stereotypes, and overt and covert discrimination may play a more powerful role in influencing decisions independently of indirect effects through changes in interest and self-beliefs. Even if a female student is both interested and good at STEM, she still might choose to avoid that career path if she perceives STEM careers as having hostile or unfriendly environments, or if she sees STEM as incongruent with femininity.

Interviewees in this study were not asked directly about structural inequality or gender roles (given the often insidious effect of gender roles and structural inequalities, it is unclear that this would have been helpful to do so in any case). Despite this, there were some interview responses that alluded to these issues. Overall, responses showed that although the vast majority of young women did not mention gender issues in their decisions, there were some participants who explicitly stated concerns with gender – even though they were not prompted to do so. This finding illustrates that the influence of gender seems to have an impact on at least some young women, in terms of major decisions such as whether or not to enrol in a STEM course. However, due to the abstract nature of gender roles and societal norms, it is most likely difficult for participants (especially younger participants) to identify how societal influences and structural inequality impacts their decisions without being specifically prompted to reflect on these processes. These findings shine a light on the need for educational institutions and industries to create more inclusive and supportive environments that welcome students and workers who do not fit into the traditional stereotypes of STEM.

Conclusions and Directions for Future Research

This study provided a unique contribution to the literature by testing what current known factors predict high school and university STEM enrolment independently of one another. Perhaps the most important contribution is the finding that gender still exerts a significant and large influence even when past achievements, self-beliefs, values, interest, anxiety, and teacher support are controlled for. This indicates that there is a need for new research to explore potential other mechanisms that may further explain the gender gap in STEM educational attainment, particularly the role and influence of structural gender inequalities on STEM enrolment in university.

Another finding of interest was that gender did not have a significant impact on senior high school selection, but negatively predicts entry into university STEM study. While I am cautious to avoid over-interpretation of these findings, it seems that interventions to attract young women into STEM may be better placed to focus on highlighting STEM university study, as opposed to the benefits of senior high school study. Findings suggest that utility value and interest in math are key factors that need to be addressed in interventions to encourage retention of students in STEM areas at a university level.

Finally, longitudinal results demonstrated that attitudes, especially utility value, self-concept, self-efficacy and interest assessments made at ages as young as 15 remain powerful predictors of not only course selection in senior high school, but decisions made during the early adulthood post-school transition period. Thus, educators and policymakers aiming to attract young people to the sciences should aim to focus interventions not only in the immediate time frame before university, but also during earlier adolescence.

Readers should be aware of several limitations within this study. Firstly, analyses are based within an Australian context, and thus, patterns of STEM enrolment may differ across different educational systems. Secondly, this research on university entrance applies to students already enrolled in STEM at senior high school, rather than those who are not. Thus, readers should not discount the role of math and science achievement in predicting STEM entry from the general population of students. Nonetheless, I maintain that following students who are already enrolled in STEM at high school is important, because these relatively high achieving students are arguably the group of students with the greatest potential to remain engaged in STEM at a tertiary level. Finally, attitudinal and achievement data stems from age 15, to predict outcomes at age 19. Despite a four-year difference between attitudinal and achievement data, and the outcome variable of university entrance, there were still significant effects from attitudes at age 15, showing the powerful role of attitudes about math from an early age. Nonetheless, readers are reminded that some of the effect sizes may be diminished because of the four year gap between time waves at age 15 and university study.

Further research is needed to establish whether attitudes formed at even younger ages can predict similar outcomes in adulthood. Finally, interventions aimed at attracting more young women in STEM should focus on increasing interest, levels of usefulness, and self-beliefs such as self-efficacy and self-concept. Further research on the role of dimensional comparisons in self-beliefs, interest, and career goals looks to be a promising line of enquiry that may help to explain why young women disengage from STEM. Indeed, these results combined with recent research indicate that women disengage, not simply due to deficits in self-beliefs or interests, but rather competing interests, and values that are incongruent with enrolment in STEM.

CHAPTER 8
DISCUSSION OF RESEARCH FINDINGS

This thesis explored two critical and unresolved questions in relation to EVT. The first major focus of the thesis was to address current research gaps in relation to the investigation of how social and cultural contexts influence the magnitude of gender differences typically observed in EVT-related variables. This question was addressed through meta-analysis in Study 1, and large-scale quantitative data analysis in Study 2. The second key focus of this thesis was to explore the degree to which traditional EVT variables can account for observed gender differences in tertiary educational attainment in STEM. Furthermore, this thesis investigated whether alternative mechanisms outside of EVT were revealed in the interview data of young Australians reflecting on their decision to engage or disengage from STEM study at a university level. These research questions were explored using longitudinal data analysis in Study 3, and a content analysis of open-ended interview data in Study 4.

In this chapter, I provide a brief recap of the critical findings from Studies 1-4 of this thesis (discussed in greater detail in Chapters 5-7). This is followed by a more general discussion that synthesises the above findings, and an exploration of how the results of this thesis contribute to the broader context of gender differences in STEM self-beliefs, motivation and educational attainment. Finally, this chapter concludes by highlighting promising new directions for research in this field.

Summary of Findings of Studies 1-4:

Study 1: Evidence from meta-analyses. Study 1 (see Chapter 5) explored whether social class and cultural context influenced the size of gender gaps in expectancy value variables in math, science, and verbal domains. Results showed there were two key findings from the meta-analysis. Firstly, gender differences in EVT variables followed a gender stereotypical pattern across academic domains. In other words, gender differences favouring boys tended to be strongest in domains that are traditionally associated with men (e.g., computing, physical sciences), while gender differences in EVT variables were the smallest in the biological sciences. In contrast, gender differences in EVT variables tended to favour girls in verbal domains.

There was, however, considerable heterogeneity in effect sizes. Thus, some of the most interesting findings came from the moderation analyses that revealed evidence of important differences and similarities with regards to social and cultural moderators of the gender gap in EVT constructs. Specifically, moderation effects showed a trend whereby gender differences became larger amongst studies characterised by a majority of high SES students, and high national level gender equality. Importantly, these two effects often explained a considerable degree of heterogeneity observed for pooled effects. There was some evidence that this trend also occurred for countries with high levels of gender segregation in university educational attainment, and amongst samples taken from advanced and elective courses. In contrast, there was little evidence that age and publication date moderated the size of gender differences in EVT variables. Overall, findings from Study 1 provided promising evidence that there are important differences in how gender effects for expectancy and value variables may vary across social and cultural context.

Study 2: Evidence from replication and extension analyses. Study 2 (see Chapter 6) investigated the degree to which the meta-analytic findings could be replicated in a large-scale nationally representative database of young Australians. Furthermore, this study extended upon previous research by exploring whether geography and Indigenous status moderated the size of gender differences for EVT-related variables. Results painted a complex picture. Main effects showed that students from low SES backgrounds, and those who identified as Indigenous, or female, reported lower scores for math self-beliefs, attitudes and STEM educational attainment. Indigenous students faced particular disadvantage, reflecting the need for better educational support to bridge educational disparities that are experienced by young Indigenous Australians (De Bortoli & Thomson, 2009). Interestingly, female students were the only group of students who experienced substantially lower self-beliefs and values towards math after achievement differences were controlled for. There was some evidence of replicating of the moderating effects from Study 1. For example, social class moderated the size of gender effects for math self-efficacy, where gender differences favouring boys were larger amongst students from high SES backgrounds. This effect remained even when controlling for achievement. However, this was tempered by the fact that a number of meta-analysis SES moderation effects for other EVT-related variables could not be replicated in Study 2.

Indigenous status, geography and immigrant status showed distinct interaction effects - revealing a complex interplay between various social and cultural factors with gender. These effects included a large interaction between Indigenous status and gender for math utility value. This revealed that amongst young Indigenous Australians, the typically observed gender effect was reversed, resulting in Indigenous girls reporting much higher levels of math utility value compared to their male peers.

Additionally, interaction effects were observed between gender and geography whereby the gender difference favouring boys in math anxiety and math utility grew significantly larger amongst students from remote locations. Surprisingly, the reverse occurred for senior high school enrolment, with regional and remote girls being more likely to enrol in high school STEM compared to their male peers, whereas the difference was much smaller among urban students. Importantly, this difference occurred with and without controls.

Finally, there were no significant interactions for self-beliefs and values when looking at immigrant status by gender. However, there was an interaction between gender and immigration status predicting high school STEM enrolment. This effect showed that girls born overseas were less likely to enter STEM than their male counterparts who were born overseas. However, it was noted that girls born overseas were still more likely to enrol in senior high school STEM compared to Australian girls.

Overall, results from Study 2 painted a complex picture of the role between social and cultural contexts in predicting gender differences in EVT-related variables. Conditions in which gender differences in EVT-related variables became larger included high socioeconomic status and living in a remote location. However, interestingly these effects did not necessarily translate into behavioural outcomes (e.g., STEM enrolment in high school and university). Gender differences in EVT-related variables became smaller amongst low SES and urban students. In contrast, typical patterns of gender differences showed some tendency to reverse amongst Indigenous students. Finally, there was little evidence that gender differences significantly differed according to immigrant status.

Study 3: Evidence from longitudinal analyses. Study 3 (see Chapter 7) used longitudinal data to explore the degree to which EVT could account for the gender differences in senior high school and university STEM enrolment. Achievement across all domains and positive attitudes toward math were associated with a greater likelihood of choosing STEM coursework. However, unexpectedly, gender was the only non-significant univariate predictor of senior high school STEM enrolment. Even more surprising, was that being female was positively related to senior high school STEM course selection when self-beliefs and attitudes were controlled for. These results suggest that for boys and girls of equal abilities and attitudes, girls are more likely to enrol in high school STEM, despite assertions that the gender gap in STEM careers begins with disengagement in high school math and science courses.

The effect of gender was very different when predicting university STEM enrolment amongst students who had already enrolled in senior high school STEM coursework. Young women were far less likely to enrol in university STEM than men even after achievement and attitudes were controlled for. Overall, results revealed that although traditional EVT-variables explained at least some of the gender gap in university STEM enrolment, there still remained a substantial amount of variance that is unexplained by EVT.

Study 4: Evidence from interview data. Study 4 (see Chapter 7) used interview data to explore whether open-ended interview data could add to current theory. Although interview data provided some insight into students' disengagement and engagement in STEM, it only incrementally added to current understandings of gender differences in STEM. Nonetheless, interview data highlighted the importance of dimensional comparisons in self-assessments of ability, but also interests and career goals in determining choice behaviour (see Chapter 7 for a more detailed discussion of this issue). In other words, the issue of competing interests may be critical in determining STEM engagement and disengagement amongst high school students. However, for this process, and others, there were more gender similarities than differences in the responses of students. The exception to this general pattern of gender similarities was that young men were more likely to report being motivated towards STEM because of perceived financial gain, while young women who chose STEM were more likely to have done so because of concern for society and the environment. As for barriers and factors needing change, young men were more likely to report that financial incentive could have changed their decision to disengage from STEM study, while more young women were likely to report that a change in interest towards STEM was critical. Overall, results indicated a need to identify alternative mechanisms outside of traditional EVT variables.

General Discussion of Findings and Future Directions for Research

Gender differences are largest in affluent and high-performing contexts.

A key theme that emerged from Studies 1-2 was a pattern whereby gender differences tended to become larger in resource-rich contexts. For example, both the meta-analysis and LSAY dataset revealed that the gender gap for math self-efficacy becomes largest amongst high SES students. This trend was also exhibited amongst countries with levels of high national level gender equality, and to a lesser degree amongst countries with high levels of gender segregation in university educational attainment, and samples from advanced and elective courses (see Chapters 5 and 6).

What do all these moderating factors have in common? It could be argued that the uniting feature of these factors is wealth and affluence. Pointedly, countries with high gender equality, and high gender segregation in university educational attainment, are often regarded as some of the most highly developed and wealthy countries – a pattern observed by Charles and Bradley (2009). Moreover, students in advanced and elective STEM courses (as included in the meta-analysis of Study 1) are more likely to come from higher SES backgrounds where scholastic achievement is highly prized. For example, in the LSAY database analysed in this thesis, SES significantly and positively predicted senior high school enrolment in STEM (see Chapter 6). Importantly, recent research has found that gender gaps (favouring boys) in STEM attitudes also tend to grow larger in high performance environments (Mann, Legewie, & DiPrete, 2015).

What can explain the large gender gaps seen in affluent and high-performing contexts? As discussed in further detail in Study 1, there are several possible explanations. One explanation is that affluent families provide greater opportunity for gender socialisation of interests to occur, as there is some evidence showing that high SES parents expose their children to more gender stereotypical extra-curricular activities and also exhibit higher levels of gender specific parenting patterns compared to parents of lower SES families (Lareau, 2003). Essentially, as children gain more experience in gender congruent activities, and less experience with gender incongruent activities – their self-beliefs and attitudes become more and more confined to gender stereotypical patterns. This pattern is consistent with the data from Studies 1 and 2, showing greater differentiation between the self-beliefs of girls and boys in math.

Another explanation is that the high achieving and competitive environments typically observed in affluent settings are responsible for larger gender differences in self-beliefs and attitude in STEM. As discussed in previous chapters, there is a growing body of evidence that shows that on average women respond less favourably to competitive environments compared to men (e.g., Bönte 2015; Gneezy et al., 2003; Niederle & Vesterlund 2007, 2010). These findings have been applied to university STEM admission, with Alon and Di Prete (2015) showing that the intensity of competition as signaled by admission standards has a larger deterring effect on female applicants compared to males. This further supports the work of Mann and DiPrete (2016) who showed that the unexpected relationships between national level gender equality and STEM attitudes disappeared after controlling for the national performance levels in STEM. Thus, the trend towards larger gender differences in attitudes towards STEM may not be caused by liberal and individualistic values that are held in post-industrial nations but rather could in part be the result of competitive performance cultures that are less enticing for young women.[6]

Finally, it is important to note that the data from Studies 1 and 2 could fit a pattern of sexual dimorphism, a pattern whereby sex differences become larger in resource rich environments (i.e., high SES contexts, highly developed nations, advanced coursework). According to this evolutionary psychology perspective, sexual dimorphism is enhanced in resource rich conditions, whereas poorer environments stunt the development of sex characteristics, particularly for males. Importantly, this pattern has been observed in both animals and humans (Stinson, 1985). Nonetheless, it should be noted that most studies examining sexual dimorphism and environment in humans tend to focus on physical sex characteristics – traits that are much more clearly tied to male-female dichotomies than inclination towards STEM (Carothers & Reis, 2013). Furthermore, in line with my earlier criticisms of evolutionary psychology (see Chapter 2), it can be problematic relying on post-hoc theorising to explain effects – an issue that is rife within evolutionary psychology (e.g., Buller, 2005). There is little hard evidence to support an evolutionary perspective in this instance, other than speculations observed for physical sex traits. Thus, I emphasise that an evolutionary perspective of this data should be interpreted with caution, as it

[6] Interestingly, medicine is a highly competitive field, yet generally has better representation of women, at least in junior roles, compared to other STEM fields (e.g., physics). I propose that the Goal Congruity Model (to be explained in coming paragraphs) can account for why this highly competitive field remains desirable for women.

relies on heavy speculation and assumptions that attitudes towards STEM are biologically impacted.

Overall, it is difficult to know exactly what might be behind the observed effects of gender gaps increasing in affluent, high achieving samples in this thesis and the previous literature. Increased gender socialisation and sexual dimorphism arguments provide possible alternatives, but there is little evidence to support these conclusions in this particular context. In contrast, there is promising new research that supports the conclusion that highly competitive environments are often associated with poorer outcomes for women in the STEM field (e.g., Alon & DiPrete, 2015; Mann et al., 2015). Thus, future research should further explore the complex interplay between competitive performance environments and gender differences in STEM.

It's Not Only the Leaky Pipeline

Results from this thesis (see Chapter 7) revealed a surprising finding – that is, there is no significant difference between boys and girls in relation to the likelihood of enrolling in a senior high school STEM course. Even more surprising, girls were actually *more* likely than boys to enrol in STEM courses when achievement and attitudes were controlled to be equal. This result questions the long-held assumption that girls' disengagement from STEM coursework during the high school years is the critical point at which gender disparities in STEM participation begin to emerge. Instead, the findings from this PhD suggest that the transition between high school and university might be the period where substantial gender disparities in STEM participation first emerge in the Australian school system.

However, it is critical to highlight that even though gender differences in STEM participation did not emerge until university in this dataset, results point towards EVT-related attitudes being a critical precursor that at least partially explain one third of the gender gap in university STEM enrolment amongst young Australians studying senior high school STEM. Importantly, Study 1 (see Chapter 5) revealed that gender disparities in STEM are unlikely to dramatically change over the course of young people's schooling. That is, there was no significant difference between 7 year old elementary students and students in older age groups (e.g., middle school, high school, and university age). Thus, educators should be aware that self-beliefs and attitudes towards STEM play a powerful role in driving course-taking behaviour amongst young people, and that these differences are likely to emerge during early childhood.

Explaining the unexplained: What can account for the residual gender effect? A major finding from Study 3 of this PhD (see Chapter 7) was that gender differences in achievement, self-beliefs, and values could not explain two thirds of the residual gender effect in university STEM enrolment. Data taken from open-ended interview responses contributed significantly to understanding the decision-making process young people undertake when selecting university majors, showing that young people spontaneously refer to EVT constructs when reflecting on their own decision-making processes. Importantly, this data revealed that young women and men were overwhelmingly more similar than different in the reasons they gave for their choices, or their reported thought processes. Consequently, interview responses in Study 4 only marginally added to current understandings of gender disparities in STEM educational attainment.

Thus, a critical question for researchers, educators and STEM professionals remains. If EVT can only account for a small amount of the gender difference seen in STEM university enrolments, then what can explain the residual gender effect observed in this thesis? The answer to this question is complex and multifaceted. Nonetheless, in this discussion I will attempt to highlight several possible explanations discussed in light of this PhD and current literature in the field.

I can do it, I want to do it, and I think it's important: The role of EVT. Although this PhD shows that EVT may not be able to account for *all* of the gender difference in course selection across high school and university enrolment, it certainly accounts for a sizeable portion of the gender gap. Even amongst a subsample of students who had already enrolled in senior high school STEM coursework, EVT-related variables were still strong predictors of enrolment in a university STEM course. In addition to this, results showed that EVT-related variables could explain a third of the gender gap in STEM university course selection (decreasing from a 15% point gap to a 10% point gap). These results are corroborated by the findings of Study 1, where effect sizes of each domain tended to correspond to the size of the study and work gender gap within the field. For example, the meta-analysis showed larger differences in EVT-related variables in computing and "hard sciences" compared to biological sciences, hinting that EVT-related variables are critical precursors to career choices. Furthermore, interview data from Study 4 lent further credibility to the external validity of EVT, with student responses organically reflecting the importance of interest and other EVT constructs when reflecting on the reasons behind their university course selection choices. Thus, these results support the conclusions of Eccles and colleagues' EVT framework (Eccles, 1994; Eccles, 2005; Eccles & Wigfield, 2002), finding strong empirical evidence that students' expectancies for success, interests, and utility value can explain a reasonable portion of the gender gap in university STEM entry.

Competing interests: An overlooked factor in STEM and gender. Despite the fact that EVT could explain a reasonable portion of the gender gap in STEM university entry, two-thirds of the gender effect remained unexplained in Study 3. Interview responses in Study 4 revealed the importance of competing interests, an issue that was not fully captured through the quantitative measures in the longitudinal analyses of Study 3. In particular, results revealed that young people undertook dimensional-type comparisons for self-assessments of ability, but also interests and career goals (see Chapter 7 for a description of dimensional comparison theory). In line with dimensional comparison theory, young people were deterred from STEM, not only because they didn't like or enjoy STEM, but because of their competing interests and abilities elsewhere. This conclusion makes sense in the context of Study 1, where gender differences were reversed for expectancy for success and values in the verbal domain, indicating that girls and women are far more likely to feel confident, to value, and to be interested in academic pursuits focusing on verbal skills like reading and writing. Finally, the importance of dimensional comparisons is further highlighted by Study 3, which showed that reading achievement negatively impacted on self-beliefs and attitudes towards math, and also negatively predicted university STEM enrolment (however this effect was not significant).

What is the underlying mechanism of dimensional comparisons? Marsh's (1986) work on the I/E model (the underlying framework for dimensional comparison theory) shows that even when differences in a students' achievement in math and verbal domains are small in size, they still can result in large differences in self-concept due to internal/dimensional comparisons. In other words, differences in students' self-concept and attitudes seem to be a much larger reflection of differences in a students' actual achievement. Importantly, this pattern is likely to be amplified in discreet choice situations (e.g., choosing a STEM major or not at university), as students are forced to choose between competing interests.

The I/E model and dimensional comparison theory provide a strong account of possible cognitive factors that could contribute to highly differentiated patterns of self-concept and attitudes in the face of trivial individual differences in academic achievement across math (d = -0.11) and science (d = -0.04). However, while the issue of competing interests and dimensional type comparisons in this PhD highlights a consistent pattern in gender differences, it only addresses one possible mechanism behind gender gaps in STEM.

Is STEM incompatible with communal goals?. In many ways, the idea of ‘competing interests’ as a mechanism behind the gender gap is more descriptive than explanatory. If researchers are truly interested in understanding the gender disparities in STEM then they need to investigate what motivational processes outside of math and science attitudes can explain gender stereotyped choices. Diekman and colleagues’ (Diekman, Brown, Johnston, & Clark, 2010; Diekman, Clark, Johnston, Brown, & Steinberg, 2011) Goal-Congruity Model highlights a promising alternative motivational model to explain gender disparities in STEM. The Goal-Congruity Model stipulates that people are motivated to act in accordance with social roles. Work by Diekman and colleagues (e.g., 2010; 2011) have shown that women are more likely to hold communal goals that focus on helping others (mirroring social roles that associate femininity with caregiving and nurturing). Importantly, despite the fact that much of the work within science contributes towards the betterment of society, many women and girls often perceive that STEM careers are incompatible with their communal goals. Thus, highlighting the prosocial and communal aspects of STEM may be critical in reducing gender disparities in STEM participation.

Indeed, Diekman and colleagues have found that by emphasising the communal and prosocial aspects of engineering careers through a service learning project, young women were likely to hold higher levels of interest and participation in engineering (Belanger, Diekman, & Steinberg, 2016). This model might help to explain why STEM fields such as medicine have better female representation, at least in junior roles, despite a highly competitive performance environment. Medicine is naturally more oriented towards prosocial and communal goals, given the discipline’s focus on helping others through illness. This theory may also explain why gender differences in confidence, liking, and interest in the biological sciences are more favourable for women and girls, compared to other STEM disciplines with less focus on helping others. Interestingly, results from this thesis show that one of the only gender differences in why young people enrolled in a STEM university course was that young women were more likely to enrol because of concern for society and the environment. In summary, the Goal Congruity Model provides a promising new line of inquiry for researchers in the field, as well as policymakers who are interested in encouraging greater female participation in the STEM fields.

Chilly climates and implicit biases: Does STEM have a problem with women?. So far, this discussion has focused on the role of self-confidence, attitudes and values in determining gender disparities in the STEM workforce. Indeed, this reflects current debates about gender and work. But is solving the gender gap in STEM university as simple as ‘*leaning in’*, as Sheryl Sandberg (2013) would say? Do women simply need to get some confidence and like science and math more?

Results from this thesis suggest that confidence and attitude change is likely to solve only a small portion of the STEM gender gap. The substantial residual gender effect that remains unexplained from Studies 3 and 4, shows that even when math achievement and attitudes are controlled for, young women are still far less likely to enrol in STEM university degrees compared to their male peers. Even when adding the idea of goal congruity and competing interests to the mix, it seems highly unlikely that these factors alone would explain such a large residual gender effect. Furthermore, explanations focusing exclusively on confidence and attitudes, while neglecting to measure why negative confidence and attitudes develop in the first place, potentially fail to address the original sources of the issue.

For example, many commentators have documented both overt and subtle hostility directed towards women in the tech industry that results in a toxic work culture (e.g., Bates, 2013). Essentially, women who choose to study STEM at university or work within the field tend to find hostile work environments, which leads them to drop out. For instance, one young woman in Study 4 mentioned that she tried to find work experience in engineering but was rejected and reported feeling as though the field was unsupportive of women. This issue has been labelled by some commentators as the ‘trapdoor’ problem (Evans, 2014). These findings might explain why female representation in STEM becomes lower and lower amongst senior positions. However, there has been little effort to investigate the degree to which younger women are aware of these experiences in the STEM industry, and the degree to which awareness of these experiences may discourage future young women from enrolling in STEM university courses.

Although the issue of hostile work environments may be important, recent work by Ceci et al. (2009) argues that overt discrimination and prejudice is now unlikely to be the main culprit behind gender disparities in STEM. Instead, the more insidious, and harder to detect, issue of implicit bias may be critical. For example, the well-known 'name' study by Moss-Racusin, Dovidio, Brescoll, Graham, and Handelsman (2012) showed that after assigning identical job applications with the name 'Jennifer' or 'John', managers were more likely to rate male applicants as more competent and hireable. Other studies have shown that employers tend to overestimate mathematical competencies of men, and underestimate the same competencies for women (Reuben, Sapienza, & Zingales, 2014). Interestingly, implicit bias favouring men in math have been documented as early as seven years of age (Cvencek, Meltzoff, & Greenwald, 2011).

Furthermore, there is emerging evidence to suggest a link between internalised implicit bias and university study choice. For example, Smeding (2012) showed that female engineering university students had much weaker implicit gender math and gender reasoning stereotypes, compared to female students who had enrolled in a humanities course (and male students in both engineering and humanities courses). Moreover, recent work by Lane, Goh, and Driven-Linn (2012) has shown that implicit gender science stereotypes accounted for the gender gap in students' intentions to study science. Thus, future research would benefit by assessing how implicit biases earlier in life impact on later behaviours amongst children and adolescents (e.g., course selection, extracurricular activities, career aspirations, college choice).

General Limitations

Chapter 4 of this thesis discussed the various strengths and limitations of the methods used in this PhD. The complementary approach of combining meta-analytic techniques with large-scale secondary data analysis has to some extent reduced the limitations of each particular method. For example, meta-analytic evidence in this thesis meant that conclusions from Study 1 were highly generalisable and reliable, given the large international sample made up of multiple published and unpublished studies. Furthermore, large scale secondary data analysis in Studies 2-3 provided the opportunity to replicate meta-analytic moderation effects with important controls, and to explore other interaction effects that had yet to be fully tested in previous literature. Nonetheless, methods to explore the research questions of this thesis have relied exclusively on existing and secondary data. Therefore, there are several general limitations that readers should be aware of as they interpret results (limitations specific to each study are discussed in chapters 5-7).

A major restriction of secondary data analysis is researchers are limited by the scales, items, and demographics that are used in existing datasets. For instance, moderation analyses in Study 1 were weakened by the fact that many studies simply did not report enough information about categories such as social class and ethnicity, to provide sufficient power for moderation to be tested across all variables. Thus, conclusions from such tests are less reliable, and also subject to biases (e.g., possibility of systematic biases in what demographics are mentioned and not mentioned in participant descriptions, as well as what groups of people are studied or not studied by researchers).

Other limitations of secondary data analysis include the requirement to formulate hypotheses and questions to fit the data where possible (Elder, Pavalko, Clipp, 1993). In this thesis, there were many questions I would have liked to explore in relation to university STEM entry, as well as many variables (e.g., cost) that would have been valuable predictors of educational attainment. Thus, to some degree research questions and hypotheses were restricted by the existing data available for analysis. However, although this approach can constrain what questions are asked, it also provided a chance for important questions to be answered. For example, using existing data I was able to conduct a number of extensive, large-scale longitudinal analyses that simply would not have been possible given the time frame and scope of a PhD. By meta-analysing existing publications and theses, I was able to synthesise existing data on gender differences across a wide array of international social and cultural contexts. Furthermore, by utilising the existing LSAY database I was able to explore the role of self-beliefs and attitudes on actual behaviour over a very long period of time. These are all tasks that would typically be untenable within the constraints of a PhD.

Implications for Policy and Practice

How do the findings of this thesis translate into policy and practice? Based on the above discussion, this discussion concludes by providing a number of recommendations for educators and researchers interested in policy and practice targeting young women studying STEM in schools.

Are girls disadvantaged in competitive STEM environments?. Based on the results of this thesis, and also previous literature, it is reasonable to speculate that young women may be disadvantaged in highly competitive STEM environments, especially when compared to young men. Although, this is an area still needing further research, educators and policymakers should be aware that there is mounting evidence that gender differences in self-beliefs within high-achieving, affluent, and competitive environments are likely to be larger compared to other environments. If future research supports the causal hypotheses put forth in this discussion (i.e., competitive environments increase the gender gap in STEM), then students might benefit from efforts to reduce levels of perceived competitiveness in order to downplay potential negative effects on female students. Importantly, this recommendation is in line with landmark motivation studies showing that pressure from competition can be detrimental to intrinsic motivation when experienced as controlling (e.g., Deci, Betley, Kahle, Abrams, & Porac, 1981; Reeve & Deci, 1996).

Boosting confidence and bolstering values: the role of attitudes in gender differences. The thesis provided further support to the importance of expectancies for success and value in STEM education for young women and girls. Even though the gender difference in math and science achievement is small in size, girls often have more negative self-beliefs and attitudes towards STEM subjects, compared to their male peers. This means that even when girls are achieving at similar levels to boys, they are far more likely to experience a lack of confidence and negative attitudes towards STEM. Thus, teachers should be mindful that even when girls are performing well, they still might struggle to feel as confident as their male peers. Likewise, boys are likely to experience lower levels of self-belief and attitudes in reading and writing (however, the gap between actual achievement and self-beliefs is smaller). Nonetheless, teachers should be aware of how self-beliefs and attitudes follow gender stereotypical patterns regardless of actual achievement.

Importantly, this thesis also showed that EVT related beliefs and attitudes could explain one third of the gender effect observed in Study 3, and were often spontaneously referred to by students in interviews in Study 4. Thus, educators should not underestimate the power of self-beliefs and attitudes in determining major educational choices such as course selection – even years before these decisions are made. These findings indicate that policymakers and educators aiming to reduce the gender gap in university STEM enrolment should support interventions and teaching strategies designed to foster self-confidence, intrinsic motivation, and attitudes towards science and math. Given the past success of self-concept and attitude interventions (see Gaspard's 2016; O'Mara et al., 2006), this recommendation provides a realistic and effective option for educators looking for evidence-based solutions to decrease gender gaps in STEM education.

Sometimes confidence isn't enough?. Finally, it is important to note that self-beliefs and attitudes at age 15 cannot account for a large portion of the gender effect for university course selection amongst students who have already enrolled in senior high school STEM. Educators and policymakers should be aware that even when students are of similar ability level and hold similar attitudes towards STEM, young women are *still* less likely to enrol in university STEM coursework. While the exact reason why there is such a substantial residual gender effect remains unknown, educators should be mindful of the possible role of implicit biases and stereotypes in deterring young women from science.

Conclusion

In conclusion, this thesis has presented four studies that explored two unresolved questions in relation to EVT. Results from the meta-analysis and also the analysis of a large-scale database showed that there are critical differences in how social and cultural contexts influence the size of gender differences typically observed in EVT-type variables. In particular, this thesis highlighted common patterns in data from this PhD and also existing literature, showing that gender differences in attitudes towards STEM favouring boys are larger in affluent and high-achieving contexts. Thus, this research highlights the need for researchers to further explore psychological gender differences across social and cultural contexts. Finally, this thesis explored the degree to which traditional EVT variables can account for observed gender differences in tertiary educational attainment in STEM. EVT variables at age 15 can explain a reasonable portion of the gender differences in university enrolment in STEM amongst students already engaged in high school STEM, however, results showed that a substantial portion of the gender effect could not be explained. Thus, it is of critical importance that future research seeks to explore other predictors of university STEM attainment in addition to self-beliefs and attitudes. In summary, this thesis has addressed several major unresolved issues in the EVT literature, and has also highlighted a number of promising avenues for future research – providing valuable insight into how educators, researchers, and policymakers can work together to make the study of STEM accessible to all.

References

*Abarbanel, T. (2008). *Predicting change across and after the transition to high school: A longitudinal path analytic examination of math related beliefs and values* (Unpublished doctoral dissertation). University of Ottawa, Ottawa, Canada.

Ader, D. N., & Johnson, S. B. (1994). Sample description, reporting, an analysis of sex in psychological research: A look at APA and APA division journals in 1990. *American Psychologist, 49*(3), 216. doi:10.1037/0003-066X.49.3.216

*Adeyemi, A. A. and O. M. Adeniyi (2010). A study of gender differences in the attitude of mathematically gifted and non-gifted senior secondary school students in Nigeria. *Gender & Behaviour, 8*(2), 3102-3116. doi:10.4314/gab.v8i2.61936

* Adkinson, J. E. (2007). *Does cooperative learning affect girls' and boys' learning and attitudes toward mathematic transformation skills in single-sex and mixed-sex classrooms?* (Unpublished doctoral dissertation). University of South Alabama, Mobile.

Ahmed, W., Minnaert, A., van der Werf, G., & Kuyper, H. (2010). Perceived social support and early adolescents' achievement: The mediational roles of motivational beliefs and emotions. *Journal of Youth and Adolescence, 39*(1), 36. doi:10.1007/s10964-008-9367-7

*Alliman-Brissett, A. E. (2006). *Factors Associated with African American Adolescents' Math Career Self-efficacy*. (Unpublished doctoral dissertation). University of Minnesota, Minneapolis.

Alon, S., & DiPrete, T. A. (2015). Gender differences in the formation of field of study choice set. *Sociological Science, 2,* 50-81. doi:10.15195/v2.a5

*Anderman, E. M. (1992). *Motivation and Cognitive Strategy Use in Reading and Writing*. Paper presented at the Annual Meeting of the National Reading Conference, 42nd San Antonio, TX, December 2-5, 1992.

Anker, R. (1998). *Gender and jobs: Sex segregation of occupations in the world.* Geneva: International Labour Office.

Ankney. C. D. (1992). Sex differences in relative brain size: The mismeasure of woman, too? *Intelligence,* 16, 329-336. doi:10.1016/0160-2896(92)90013-h

Apfelbaum, E. P., Phillips, K. W., & Richeson, J. A. (2014). Rethinking the baseline in diversity research: Should we be explaining the effects of homogeneity?

Perspectives on Psychological Science, 9(3), 235-244. doi: 10.1177/1745691614527466

Arnot, M. (2002). *Reproducing gender: Essays on educational theory and feminist politics.* London: Routledge Falmer

*Asonye, E. I. (2002). *Psychological Differences in Attitude Toward and Academic Achievement in Mathematics Between Male and Female Students in Imo Secondary Schools, Nigeria.* (Unpublished doctoral dissertation). Walden University, Minneapolis.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*(6p1), 359. doi:10.1037/14156-004

Atkinson, J. W. (1964). *An introduction to motivation.* Oxford, England: Van Nostrand.

Auwarter, A. E., & Aruguete, M. S. (2008a). Effects of student gender and socioeconomic status on teacher perceptions. *The Journal of Educational Research, 101*(4), 242-246. doi:10.3200/joer.101.4.243-246

Auwarter, A. E., & Aruguete, M. S. (2008b). Counselor perceptions of students who vary in gender and socioeconomic status. *Social Psychology of Education, 11*(4), 389-395. doi:10.1007/s11218-008-9056-0

Baker D. P., Jones D. P. (1993) Creating gender equality: Cross-national gender stratification and mathematical performance. *Sociology of Education*. 66:91–103. doi:10.2307/2112795

*Baker, L. and A. Wigfield (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly, 34*(4), 452-477. doi:10.1598/rrq.34.4.4

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. doi:10.1037/0033-295X.84.2.191

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: Freeman.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development, 67*(3), 1206–1222. doi:10.2307/1131888

Barman, C. R. (1997). Students' views of scientists and science: Results from a national study. *Science and Children, 35*(1), 18.

Baron-Cohen, S. (2003) *The Essential Difference: Male and Female Brains and the Truth About Autism.* Basic Books: New York, NY.

Baron-Cohen, S., Lutchmaya, S., & Knickmeyer, R. C. (2004). *Prenatal testosterone in mind: Amniotic fluid studies*. MIT Press.

Bates, L. (2013). Women in science: 'Whoa, what are you doing here?' *The Guardian.* https://www.theguardian.com/lifeandstyle/womens-blog/2013/oct/17/women-in-science-ada-lovelace-gender Retrieved Monday, 7th August, 2017.

*Battle, E. S. (1966). Motivational Determinants of Academic Competence. *Journal of Personality and Social Psychology 4*(6), 634-642. doi:10.1037/h0024028

Belanger, A. L., Diekman, A. B., & Steinberg, M. (2016). Leveraging communal experiences in the curriculum: Increasing interest in pursuing engineering by changing stereotypic expectations. *Journal of Applied Social Psychology, 47,* 305-319. doi:10.1111/jasp.12438

Berman, E., & Machin, S. (2000). Skill-biased technology transfer around the world. *Oxford Review of Economic Policy, 16*(3), 12-22. doi:10.1093/oxrep/16.3.12

*Beyer, S. (2014). Why Are Women Underrepresented in Computer Science? Gender Differences in Stereotypes, Self-Efficacy, Values, and Interests and Predictors of Future CS Course-Taking and Grades. *Computer Science Education*, *24*(2), 40. doi:10.1080/08993408.2014.963363

*Bhanot, R. T. and J. Jovanovic (2009). The Links Between Parent Behaviors and Boys' and Girls' Science Achievement Beliefs. *Applied Developmental Science 13*(1), 42-59. doi:10.1080/10888690802606784

*Black, J. M. (2008). *Academic self-concept, subjective task value, and beliefs about intelligence in dual-language and English-only elementary school students.* (Unpublished doctoral dissertation). Stanford University, Stanford.

Blackburn, R. M., Brooks, B., & Jarman, J. (2001) *The gendering of work around the world: Occupational gender segregation and inequality* (Report to the United Nations on International Occupational Gender Segregation, 1999), Cambridge Studies in Social Research 9, Cambridge: Cambridge University, SRG Publications

Blackburn, R., Browne, J., Brooks, B., & Jarman, J. (2002). Explaining gender segregation. *The British Journal of Sociology, 53*(4), 513-536. doi:10.1080/0007131022000021461

Blackburn, R. M., Jarman, J., & Brooks, B. (2000). The puzzle of gender segregation and inequality: A cross-national analysis. *European Sociological Review*. doi:10.1093/esr/16.2.119

*Boe, M. V. & E. K. Henriksen (2013). Love it or leave it: Norwegian students' motivations and expectations for postcompulsory physics. *Science Education, 97*(4), 550-573. doi:10.1002/sce.21068

*Bonitz, V. S., Larson, L. M., & Armstrong, P. I. (2010). Interests, self-efficacy, and choice goals: An experimental manipulation. *Journal of Vocational Behavior, 76*(2), 223-233. doi:10.1016/j.jvb.2009.09.003

*Bonnot, V., & J. C. Croizet (2007). Stereotype internalization, math perceptions, and occupational choices of women with counter-stereotypical university majors. *Swiss Journal of Psychology*, *66*(3). 169-178. doi:10.1024/1421-0185.66.3.169

Bönte, W. (2015). Gender differences in competitive preferences: New cross-country empirical evidence. *Applied Economics Letters*, *22*, 71-75. doi:10.1080/13504851.2014.927560

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis.* Chichester, West Sussex: John Wiley & Sons, Ltd.

Bowleg, L. (2008). When Black + lesbian + woman ≠ Black lesbian woman: The methodological challenges of qualitative and quantitative intersectionality research. *Sex Roles, 59,* 312–325. doi:10.1007/s11199- 008-9400-z

Bouchey, H. A., & Harter, S. (2005). Reflected appraisals, academic self-perceptions, and math/science performance during early adolescence. *Journal of Educational Psychology, 97*(4), 673. doi:10.1037/0022-0663.97.4.673

*Bouffard, T., Marcoux, M. F., Vezeau, C., &, Bordelau, L. (2003). Changes in self-perceptions of competence and intrinsic motivation among elementary schoolchildren. *British Journal of Educational Psychology, 73*(2), 171-186. doi:10.1348/00070990360626921

Bowen, G. L., Hopson, L. M., Rose, R. A., & Glennie, E. J. (2012). Students' perceived parental school behavior expectations and their academic performance: A longitudinal analysis. *Family Relations, 61*(2), 175-191. doi:10.1111/j.1741-3729.2011.00695.x

*Branom, C. M. (2014). *The school context of gender disparities in math motivation* (Unpublished doctoral dissertation). University of California, Berkley.

Brewster, K. L., & Rindfuss, R. R. (2000). Fertility and women's employment in industrialized nations. *Annual Review of Sociology, 26*(1), 271-296. doi: 10.1146/annurev.soc.26.1.271

*Britner, S. L. (2002). *Science self-efficacy of African American middle school students: Relationship to motivation self-beliefs, achievement, gender, and gender orientation* (Unpublished doctoral dissertation). Emory University, Atlanta.

Brizendine, L. (2006). *The female brain.* New York: Morgan Road/Broadway Books.

*Brockman, G. (2007). *What factors influence achievement in remedial mathematics classes?* (Unpublished doctoral dissertation). University of Southern California, Los Angeles.

Brofenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press: Cambridge, Massachusetts, London, England.

*Broome, P. (2001). The gender-related influence of implicit self-theories of one's intelligence with regard to academic performance in introductory physics classes. *Psychologische Beitrage, 43*(1), 100-128.

Brosnan, M. (2006). Digit ratio and faculty membership: Implications for the relationship between prenatal testosterone and academia. *British Journal of Psychology*, 97, 455–466. doi:10.1348/000712605x85808

Brotman, J. S., & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching, 45*(9), 971-1002. doi:10.1002/tea.20241

Brown, B. B. (1990). *Peer groups and peer cultures.*

Brown, E. R., Smith, J. L., Thoman, D. B., Allen, J. M., & Muragishi, G. (2015). From bench to bedside: A communal utility value intervention to enhance students' biomedical science motivation. *Journal of Educational Psychology, 107*(4), 1116. doi:10.1037/edu0000033

Buller, D. J. (2005). Evolutionary psychology: the emperor's new paradigm. *Trends in Cognitive Science, 9*(6), 277-283. doi:10.1016/j.tics.2005.04.003

Buss, D. M. (1991). Evolutionary personality psychology. *Annual Review of Psychology, 42*, 459-491. doi:10.1146/annurev.ps.42.020191.002331

Buss, D. M. (1995). Psychological sex differences: Origins through sexual selection. *American Psychologist, 50*(3), 164-168. doi:10.1037/0003-066X.50.3.164

Cahill, L. (2003). Sex and hemisphere related influences on the neurobiology of emotionally influenced memory. *Progress in NeuroPsychopharmacology & Biological Psychiatry, 27*(8), 1235-1241. doi:10.1016/j.pnpbp.2003.09.019

*Campbell, N. K., & G. Hackett (1986). The effects of mathematics task performance on math self-efficacy and task interest. *Journal of Vocational Behavior, 28*(2), 149-162. doi:10.1016/0001-8791(86)90048-5

*Campos-Sanchez, A., Lopez-Nunez, J. A., Carriel, V., Martin-Piedra, M. A., Sola, T., & Alaminos, M. (2014). Motivational component profiles in university students learning histology: a comparative study between genders and different health science curricula. *BMC Medical Education*, *14*(13). doi:10.1186/1472-6920-14-46

Caplan, P. J., & Caplan, J. B. (2016). *Thinking critically about research on sex and gender* (3rd ed.). New York: Routledge

Caplan, P. J., McPherson, G. M., & Tobin, P. (1985). Do sex-related differences in spatial ability exist? *American Psychologist, 40*, 786–799. doi: 10.1037//0003-066x.41.9.1012

Card, N. A. (2012). *Applied meta-analysis for social science research.* New York: The Guilford Press.

Carnevale, A. P., N. Smith, and M. Melton. (2011). *STEM*. Washington, DC: Georgetown University Center on Education and the Workforce.

Carothers, B. J., & Reis, H. T. (2013). Men and women are from Earth: examining the latent structure of gender. *Journal of Personality and Social Psychology, 104*(2), 385. doi: 10.1037/a0030437

Catsambis, S. (1994). The path to math: Gender and racial-ethnic differences in mathematics participation from middle school to high school. *Sociology of Education*, 199-215. doi:10.2307/2112791

Catsambis, S. (1995). Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching, 32*(3), 243-257. doi: 10.1002/tea.3660320305

*Cavas, P. (2011). Factors affecting the motivation of Turkish primary students for science learning, *Science Education International*, *22*(1), 31-42.

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. *Psychological Bulletin, 135*(2), 218. doi:10.1037/a0014412

*Chan, J. Y. (2015). *Effect of students' affective characteristics on learning and achievement in first-year general chemistry.* (Unpublished doctoral dissertation). University of New Hampshire, New Hampshire.

Charles, M. (1990). *Occupational sex segregation: a log-linear analysis of patterns in 25 industrial countries* (Doctoral dissertation, to the Department of Sociology. Stanford University)

Charles, M. (1992). Cross-national variation in occupational sex segregation. *American Sociological Review*, 483-502. doi:10.2307/2096096

Charles, M., & Bradley, K. (2009). Indulging our gendered selves? Sex segregation by field of study in 44 countries. *American Journal of Sociology*. 114, 924-976. doi:10.1086/595942

Charles, M., & Grusky, D. B. (2004). *Occupational ghettos: The worldwide segregation of women and men* (Vol. 200). Stanford, CA: Stanford University Press.

Charles, M., Harr, B., Cech, E., & Hendley, A. (2014). Who likes math where? Gender differences in eighth-graders' attitudes around the world. *International Studies in Sociology of Education*, *24*(1), 85-112. doi:10.1080/09620214.2014.895140

*Cheong, Y. F., Pajares, F., & Oberman, P. S. (2004). Motivation and academic help-seeking in high school computer science, *Computer Science Education*, *14*(1). 3-19. doi:10.1076/csed.14.1.3.23501

Cheung, M. (2011). *metaSEM: meta-analysis: A structural equation modelling approach*. R package version 0.5-3. Retrieved from http://courses.nus.edu.sg/course/psycwlm/Internet/metaSEM

Cheung, M. (2014). Modeling dependent effect sizes with three-level metaanalyses: A structural equation modeling approach. *Psychological Methods, 19,* 211-229. doi:10.1037/a0032968

*Chiu, M. S. (2011). Effects of a women-in-sciences/men-in-mumanities intervention on Taiwanese adolescents' attitudes towards learning science. *Asia-Pacific Education Researcher, 20*(2), 322-335. doi:10.1111/j.1442-2018.2007.00312.x

Choo, H. Y., & Ferree, M. M. (2010). Practicing intersectionality in sociological research: A critical analysis of inclusions, interactions, and institutions in the study of inequalities. *Sociological Theory, 28,* 129–149. doi:10.1111/j.1467-9558.2010.01370.x

*Chouinard, R., & N. Roy (2008). Changes in high-school students' competence beliefs, utility value and achievement goals in mathematics. *British Journal of Educational Psychology, 78*(1), 31-50. doi:10.1348/000709907x197993

*Chouinard, R., Vezeau, C., Bouffard, T. Jenkins, B. (1999). Gender differences in the development of mathematics attitudes. *Journal of Research & Development in Education, 32*(3), 184-192.

*Chow, S. J., & B. C. S. Yong (2013). Secondary School Students' Motivation and Achievement in Combined Science. *US-China Education Review B, 3*(4), 213-228.

Chrisler, J. C., & McCreary, D. R. (2010). *Handbook of gender research in psychology (Vol. 1).* New York NY: Springer.

Christiansen, K., & Knussman, R. (1987). Sex hormones and cognitive functioning in men. *Neuropsychobiology, 18,* 27–36. doi:10.1159/000118389

Clarke, Edward H. (1873/2006). *Sex in Education: Or, a fair chance for girls.* Boston: James R. Osgood & Company. Ebook retrieved from The Project Gutenberg.

*Coffin, R. J. and P. D. MacIntyre (1999). Motivational influences on computer-related affective states. *Computers in Human Behavior, 15*(5), 549-569.

Cohen, J. (1977). *Statistical power analysis for behavioral sciences*. New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist, 64,* 170–180. doi:10.1037/a0014564

Coley, R. J. (2001). *Differences in the gender gap: Comparisons across racial/ethnic groups in education and work*. Policy Information Report.

Coltrane, S., & Adams, M. (2008). Gender and families (Vol. 5). Lanham, Maryland: Rowman & Littlefield.

Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A., & Ahluwalia, J. (2000). Sex differences in human neonatal social perception. *Infant Behavior and Development*, *23*, 113–118. doi:10.1016/s0163-6383(00)00032-1

*Copping, K. E. (2012). *High school students' career aspirations: Influences of gender stereotypes, parents, and the school environment* (Unpublished doctoral dissertation). University of North Carolina, Chapel Hill.

Cosmacini G. (1997). *The long art: The history of medicine from antiquity to the present*. Rome: Oxford University Press.

* Crane, L. R., Gustafson, J. L., & Poziemski, C. (2000). *Motivational Aspects of Reading and Its Measurement in Community College Students.* AIR 2000 Annual Forum Paper**:** 45.

Crenshaw, K. (1989). *Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics.* University of Chicago Legal Forum, 140, 139–167.

*Cribbs, J. D. (2013). *The development of freshman college calculus students' mathematics identity and how it predicts students' career choice.* (Unpublished doctoral dissertation). Clemson University, Clemson.

*Crombie, G., Abarbanel, T., &, Trinneer, A. (2002). All-female classes in high school computer science: Positive effects in three years of data. *Journal of Educational Computing Research, 27*(4), 385-409.

*Crombie, G., & P. I. Armstrong (1999). Effects of classroom gender composition on adolescents' computer-related attitudes and future intentions. *Journal of Educational Computing Research, 20*(4), 317-327. doi:10.2190/vrd4-69af-wpq6-p734

Crosnoe, R., Riegle‐Crumb, C., Field, S., Frank, K., & Muller, C. (2008). Peer group contexts of girls’ and boys’ academic experiences. *Child Development, 79*(1), 139-155. doi:10.1111/j.1467-8624.2007.01116.x

*Cupani, M., & R. M. Pautassi (2013). Predictive contribution of personality traits in a sociocognitive model of academic performance in mathematics. *Journal of Career Assessment, 21*(3), 395-413. doi:10.1177/1069072712475177

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

Darwin, C. (1871/2016). *The Descent of Man.* Diversion Books: New York, NY.

Davison, K., & Susman, E. (2001). Are hormone levels and cognitive ability related during early adolescence? *International Journal of Behavioral Development, 25,* 416–428. doi:10.1080/016502501316934842

*Deacon, M. M. (2012). *Classroom learning environment and gender: Do they explain math self-efficacy, math outcome expectations, and math interest during early adolescence?* (Unpublished doctoral dissertation). University of Virginia, Charlottesville.

*DeBacker, T. K. and R. Nelson (1999). Variations on an expectancy-value model of motivation in science. *Contemporary Educational Psychology, 24*(2), 71-94. doi:10.1006/ceps.1998.0984

Deci, E. L., Betley, G., Kahlc,.J., Abrams, L., & Porac, J. (1981). When trying to win: Competition and intrinsic motivation. *Personality and Social Psychology Bulletin, 7*, 79-83. doi:10.1177/014616728171012

Deci, E. L., & Ryan, A. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York, NY: Plenum Publishing Co.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. doi: 10.1207/s15327965pli1104_01

*Deemer, E. D., Smith, J. L., Thorman, D. B., & Chase, J. P. (2014). Precision in career motivation assessment: Testing the subjective science attitude change measures. *Journal of Career Assessment*, *22*(3), 489-504. doi:10.1177/1069072713498683

Denmark, F. L., & Fernandez, L. C. (1993). Historical development of the psychology of women. *Psychology of Women: A Handbook of Issues and Theories*, 4-22.

*Denner, J., Werner, L., O'Connor, L., &, Glassman, J. (2014). Community college men and women: A test of three widely held beliefs about who pursues computer science. *Community College Review, 42*(4), 342-362. doi:10.1177/0091552114535624

Department of Innovation, Industry, Science and Research (2011). *2011 strategic roadmap for Australian research infrastructure.* Commonwealth of Australia. Retrieved from https://docs.education.gov.au/system/files/doc/other/national_collaborative_research_infrastructure_strategic_roadmap_2011.pdf

Desy, E., Peterson, S., & Brockman, V. (2009). Attitudes and interests among university students in introductory nonmajor science courses: Does gender matter? *Journal of College Science Teaching, 39*(2), 16-23.

Devine, A., Fawcett, K., Szűcs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and Brain Functions, 8*(1), 33. doi:10.1186/1744-9081-8-33

*DeWitt, J., Archer, L., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2011). High aspirations but low progression: The science aspirations-careers paradox amongst minority ethnic students. *International Journal of Science and Mathematics Education, 9*(2), 243-271. doi:10.1007/s10763-010-9245-0

*Dickhauser, O., & Stiensmeier-Pelster , J. (2002). Gender differences in computer work: Evidence for the model of achievement-related choices. *Contemporary Educational Psychology, 27*(3), 486-496. doi:10.1006/ceps.2001.1106

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics courses. *Psychological Science, 21*(8), 1051-1057. doi:10.1177/0956797610377342

Diekman, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., & Steinberg, M. (2011).Malleability in communal goals and beliefs influences attraction to STEM careers: Evidence for a goal congruity perspective. *Journal of Personality and Social Psychology, 101*(5), 902. doi:10.1037/a0025199

DiPrete, T. A., & Buchmann, C. (2013). *The rise of women: The growing gender gap in education and what it means for American schools*. New York: Russell Sage Foundation.

Dinella, L. M., Fulcher, M., & Weisgram, E. S. (2014). Sex-typed personality traits and gender identity as predictors of young adults’ career interests. *Archives of Sexual Behavior, 43*(3), 493-504. doi:10.1007/s10508-013-0234-6

*Dlamini, M. S. (1998). *The relationship between students' attitude toward mathematics and achievement in mathematics in Swaziland* (Unpublished doctoral dissertation). Ohio State University, Columbus.

*Dogbey, G. Y. (2011). *Attitudes of community college developmental students toward mathematics and their perception of mathematically intensive careers* (Unpublished doctoral dissertation). Ohio University, Columbus.

*Doube, W. and C. Lang (2012). Gender and Stereotypes in Motivation to Study Computer Programming for Careers in Multimedia. *Computer Science Education*, *22* (1), 63-78. doi:10.1080/08993408.2012.666038

Duong, M. T., Badaly, D., Liu, F. F., Schwartz, D., & McCarty, C. A. (2016). Generational differences in academic achievement among immigrant youths: A meta-analytic review. *Review of Educational Research*, *86*(1), 3-41. doi: 10.3102/0034654315577680

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation.* Hillsdale, NJ: Erlbaum.

Eagly, A. H., Eaton, A., Rose, S. M., Riger, S., & McHugh, M. C. (2012). Feminism and psychology: analysis of a half-century of research on women and gender. *American Psychologist, 67*(3), 211. doi:10.1037/a0027260

Eccles, J. S. (1994). Understanding women’s educational and occupational choices. *Psychology of Women Quarterly, 18*(4), 585-609. doi: 10.1111/j.1471-6402.1994.tb01049.x

Eccles, J. S. (2005). Subjective task value and the Eccles et al. model of achievement-related choices. *Handbook of competence and motivation*, 105-121.

Eccles (Parsons), J. Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C., Meece, J. L., & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75-146). San Francisco, CA: W. H. Freeman.

Eccles, J. S., & Hoffman, L. W. (1984). Socialization and the maintenance of a sex-segregated labor market. In H. W. Stevenson & A. E. Siegel (Eds.), *Research in Child Development and Social Policy* (Vol. 1, pp. 367-420). Chicago: University of Chicago Press.

Eccles, J. S., & Jacobs, J. E. (1986). Social forces shape math attitudes and performance. *Signs, 11*, 367-380. doi:10.1086/494229

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues, 46*(2), 183-201. doi:10.1111/j.1540-4560.1990.tb01929.x

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109-132. doi: 0.1146/annurev.psych.53.100901.135153

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

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin, 94(1),* 100. doi:10.1037/0033-2909.94.1.100

Elder, G. H. (1998). The life course as developmental theory. *Child Development, 69*(1), 1-12. doi:10.2307/1132065

Elder, G. H., Pavalko, E. K., & Clipp, E. C. (1993). *Working with archival data: Studying lives* (Vol. 88). Sage.

*Elisha-Primo, I., Sandler, S., Goldfrad, K., Ferenz, O., & Perpignan, H (2010). Listening to students' voices: A curriculum renewal project for an EFL graduate academic program. *System, 38*(3), 457-466. doi:10.1016/j.system.2010.02.002

*Elmore, P. B. and E. S. Vasu (1980). Relationship between selected variables and statistics achievement: Building a theoretical model. *Journal of Educational Psychology, 72(*4), 457-467. doi:10.1037/0022-0663.72.4.457

Else-Quest, N. M., & Grabe, S. (2012). The political is personal: Measurement and application of nation-level indicators of gender equity in psychological research. *Psychology of Women Quarterly*, *36*(2), 131-144. doi:10.1177/0361684312441592

Else-Quest, N. M., & Hyde, J. S. (2016a). Intersectionality in quantitative psychological research I. Theoretical and epistemological issues. *Psychology of Women Quarterly*, *40*(2), 155-170. doi:10.1177/0361684316629797

Else-Quest, N. M., & Hyde, J. S. (2016b). Intersectionality in quantitative psychological research II. Methods and techniques. *Psychology of Women Quarterly*, *40*(3), 319-336. doi:10.1177/0361684316647953

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

*Else-Quest, N. M., Mineo, C. C., & Higgins, A. (2013). Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychology of Women Quarterly*, *37*(3), 293-309. doi:10.1177/0361684313480694

*Ethington, C. A. (1991). A test of a model of achievement behaviors. *American Educational Research Journal, 28*(1), 155-172. doi:10.3102/00028312028001155

Evans, J. (2014). *Women in tech: It's not just a pipeline problem.* http://social.techcrunch.com/2014/08/23/just-another-white-dude-writing-about-diversity/. [Online; accessed 26th July, 2017].

*Falco, L. D. (2008). *"Skill-builders": Enhancing middle school students' self-efficacy and adaptive learning strategies in mathematics* (Unpublished doctoral dissertation). University of Arizona, Tucson.

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

Fausto-Sterling, A. (1985). *Myths of gender: Biological theories about women and men.* Basic Books: New York, New York

Fausto-Sterling, A. (2000). *Sexing the body: Gender politics and the construction of sexuality.* New York, NY: Basic Books.

Feather, N. T. (1982). Expectancy–value approaches: Present status and future directions. In N. T. Feather (Ed.), *Expectations and actions: Expectancy–value models in psychology* (pp. 395–420). Hillsdale, NJ: Erlbaum.

*Feather, N. (1988). Values, valences, and course enrollment: Testing the role of personal values within an expectancy-valence framework. *Journal of Educational Psychology, 80*(3), 381-391. doi:10.1037/0022-0663.80.3.381

Felson, R. B. (1989). Parents and the reflected appraisal process: A longitudinal analysis. *Journal of Personality and Social Psychology, 56*, 965–971. doi:10.1037//0022-3514.56.6.965

*Fennema, E., & J. Sherman (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. *American Educational Research Journal 14*(1), 51-71. doi:10.3102/00028312014001051

*Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences, 19*(4), 499-505. doi:10.1016/j.lindif.2009.05.004

Ferry, T. R., Fouad, N. A., & Smith, P. L. (2000). The role of family context in a social cognitive model for career-related choice behavior: A math and science perspective. *Journal of Vocational Behavior, 57*(3), 348-364. doi:10.1006/jvbe.1999.1743

Few-Demo, A. L. (2014). Intersectionality as the ‘‘new’’ critical approach in feminist family studies: Evolving racial/ethnic feminisms and critical race theories. *Journal of Family Theory & Review, 6*, 169–183. doi:10.1111/jftr.12039

Field, A. P. (2003). The problems in using fixed-effects models of metaanalysis on real-world data. *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences, 2*, 105-124. doi: 10.1207/S15328031US0202_02

Fine, C. (2010). *Delusions of gender: The real science behind sex differences*. Icon Books: London

Finegan, J. K., Niccols, G. A., & Sitarenios, G. (1992). Relations between prenatal testosterone levels and cognitive abilities at 4 years. *Developmental Psychology, 28*, 1075–1089. doi:10.1037//0012-1649.28.6.1075

Fink, B., Brookes, H., Neave, N., Manning, J. T., & Geary, D. C. (2006). Second to fourth digit ratio and numerical competence in children. *Brain and Cognition, 61*, 211–218. doi:10.1016/j.bandc.2006.01.001

Fox, D., Prilleltensky, I., & Austin, S. (Eds.). (2009). *Critical psychology: An introduction.* London: Sage.

Frank, K. A., Muller, C., Schiller, K. S., Riegle-Crumb, C., Mueller, A. S., Crosnoe, R., & Pearson, J. (2008). The social dynamics of mathematics coursetaking in high school 1. *American Journal of Sociology, 113*(6), 1645-1696. doi: 10.1086/587153

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: growth trajectories in two male-sex-typed domains. *Developmental Psychology, 38*(4), 519. doi:10.1037//0012-1649.38.4.519

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

*Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics -A "hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education, 22*(4), 497-514. doi:10.1007/bf03173468

Friedman, L. (1989). Mathematics and the gender gap: A meta-analysis of recent studies on sex differences in mathematic tasks. *Review of Educational Research, 59*, 185-213. doi:10.3102/00346543059002185

Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology, 74*(2), 435. doi: 10.1037//0022-3514.74.2.435

*Fryer, L. K. (2015). Predicting self-concept, interest and achievement for first-year students: The seeds of lifelong learning. *Learning and Individual Differences, 38*. 107-114. doi:10.1016/j.lindif.2015.01.007

Fryer, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics, 2*(2), 210-240. doi:10.3386/w15430

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., Ismaila, A. (2011). Conducting quantitative synthesis when comparingmedical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology, 64*, 1187–1197. doi: 10.1016/j.jclinepi.2010.08.010

*Fullerton, J. A., & Umphrey, D. (2001). *An analysis of attitudes toward statistics: Gender differences among advertising majors.* Paper presented at the Annual Meeting of the Association for Education in Journalism and Mass Communication (84th, Washington DC, August 5-8, 2001).

Galton, F. (1907). *Inquiries into the human faculty and its development.* London: Dent.

Gannon, L. (2002). A critique of evolutionary psychology. *Psychology, Evolution & Gender, 4*(2), 173-218. doi:10.1080/1461666031000063665

*Gasco, J., Goni, A., & Villarroel, J. D. (2014). *Sex differences in mathematics motivation in 8th and 9th grade.* 5th World Conference on Educational Sciences, 116: 1026-1031.

Geary, D. C. (1998). *Male, female: The evolution of human sex differences.* Washington, DC: American Psychological Association.

Gharibyan, H., & Gunsaulus, S. (2006). Gender gap in computer science does not exist in one former soviet republic: results of a study. *ACM SIGCSE Bulletin, 38*(3), 222-226. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.559.4987&rep=rep1&type=pdf

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development.* Cambridge, MA: Harvard University Press.

*Githua, B. N. and J. G. Mwangi (2003). Students' mathematics self-concept and motivation to learn mathematics: Relationship and gender differences among Kenya's secondary-school students in Nairobi and Rift Valley Provinces. *International Journal of Educational Development, 23*(5), 487-499. doi: 10.1016/s0738-0593(03)00025-7

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3-8. doi:10.2307/1174772

Gleitman, H. (1981). *Psychology.* New York: Norton.

Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist, 56*, 109–118. doi:10.1037/0003-066X.56.2.109

*Glynn, S. M., et al. (2011). Science Motivation Questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, *48*(10), 1159-1176. doi:10.1002/tea.20442

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *Quarterly Journal of Economics, 118*, 1049-1074. doi:10.1162/00335530360698496

Goldstein, H. (1995). *Multilevel statistical models.* London, UK: Edward Arnold.

Greene, J. C. (2007). *Mixed methods in social inquiry.* San Francisco, CA: Jossey-Bass.

*Greene, B. A., DeBacker, T. K., Ravindran, B., & Krows, J. A. (1999). Goals, values, and beliefs as predictors of achievement and effort in high school mathematics classes. *Sex Roles: A Journal of Research*, *40*(5).

Griscom, J. L. (1992). Women and power: Definition, dualism, and difference. *Psychology of Women Quarterly, 16*(4), 389-414. doi:10.1111/j.1471-6402.1992.tb00264.x

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 105–117). London: Sage.

Guimond, S., Branscombe, N., Brunot, S., Buunk, A., Chatard, A., De´sert, M., . . . Yzerbyt, V. (2007). Culture, gender, and the self: Variations and impact of social comparison processes. *Journal of Personality and Social Psychology*, 92, 1118-1134. doi:10.1037/0022-3514.92.6.1118

Guimond, S., Chatard, A., Martinot, D., Crisp, R., & Redersdorff, S. (2006). Social comparison, self-stereotyping, and gender differences in self-construals. *Journal of Personality and Social Psychology*, 90, 221. doi:10.1037/0022-3514.90.2.221

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science*, 320, 1164-1165. doi:10.1126/science.1154094

Guo, J., Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations, and attainment: A longitudinal study. *American Educational Research Journal*, *52*(2), 371-402. doi:10.3102/0002831214565786

*Guo, J., Marsh, H. W., Parker, P. D., Morin, A. J., & Yeung, A. S. (2015). Expectancy-value in mathematics, gender and socioeconomic background as predictors of achievement and aspirations: A multi-cohort study. *Learning and Individual Differences, 37*, 161-168. doi:10.1016/j.lindif.2015.01.008

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

Gur, R. C., Alsop, D., Glahn, D., Petty, R., Swanson, C. L., Maldjian, J. A., et al. (2000). An fMRI study of sex differences in regional activation to a verbal and a spatial task. *Brain and Language, 74*, 157–170. doi:10.1006/brln.2000.2325

Gur, R. C., & Gur, R. E. (2007). Neural substrates for sex differences in cognition. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 189–198).Washington, DC: American Psychological Association.

*Guthrie, J. T., Klauda, S. L., & Ho, A. N. (2013). Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. *Reading Research Quarterly, 48*(1), 9-26. doi:10.1002/rrq.035

*Guvercin, O., Tekkaya, C., & Sungur, S. (2010). A cross age study of elementary students' motivation towards science learning. *Hacettepe University Journal of Education, 39*, 233-243.

Haaken, J. (1988). Field dependence research: A historical analysis of a psychological construct. *Signs: Journal of Women in Culture and Society, 13*(2), 311-330. doi:10.1086/494408

*Hackett, G. (1992). Gender, ethnicity, and social cognitive factors predicting the academic achievement of students in engineering. *Journal of Counseling Psychology, 39*(4), 527-538. doi:10.1037//0022-0167.39.4.527

*Hackett, G., Betz, N. E., O'Halloran, M., & Romac, D. S. (1990). Effects of verbal and mathematics task performance on task and career self-efficacy and interest. *Journal of Counseling Psychology, 37*(2), 169-177. doi: 10.1037//0022-0167.37.2.169

Haier, R. (2007). Brains, bias, and biology: Follow the data. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more women in science? Top researchers debate the evidence* (pp. 113–120). Washington, DC: American Psychological Association.

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2004). Structural brain variation and general intelligence. *NeuroImage, 23,* 425–433. doi: 10.1016/j.neuroimage.2004.04.025

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2005). The neuroanatomy of general intelligence: Sex matters. *NeuroImage, 25,* 320–327. doi:10.1016/j.neuroimage.2004.11.019

Haines, E., Deaux, K., & Lofaro, N. (2016). The Times They Are a-Changing … or Are They Not? A Comparison of Gender Stereotypes, 1983–2014. *Psychology of Women Quarterly, 40*(3). pp. 353–363, doi:10.1177/0361684316634081

Hakim, C. (1979). *Occupational Segregation: A Study of the Separation of Men and Women's Work in Britain, the United States and Other Countries,* Research Paper 9, London: Department of Employment.

Halle, T. G., Kurtz-Costes, B., & Mahoney, J. L. (1997). Family influences on school achievement in low-income, African American children. *Journal of Educational Psychology, 89,* 527–537. doi:10.1037//0022-0663.89.3.527

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). New York, NY: Psychology Press.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, *8*(1), 1-51.

Hampson, E., & Moffat, S. D. (2005). The psychobiology of gender: The cognitive effects of reproductive hormones in the adult nervous system. In R. J. Sternberg, A. Eagley, & A. Beal (Eds.), *The Psychology of Gender* (2nd ed., pp. 38–64). New York: Guilford Press.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S.(2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology, 111*(5), 745. doi:10.1037/pspp0000075

Harackiewicz, J. M., Rozek, C. S., Hulleman, C. S., & Hyde, J. S. (2012). Helping parents to motivate adolescents in mathematics and science an experimental test of a utility-value intervention. *Psychological Science, 23*(8), 899-906. doi: 10.1177/0956797611435530

Harding, S. (1998). *Is science multicultural? Postcolonialisms, feminisms, and epistemologies*. Bloomington: Indiana University Press.

Hare-Mustin, R. T., & Marecek, J. (1990). *Making a difference: Psychology and the construction of gender.* New Haven, CT: Yale University Press.

*Harty, H., Samuel, K., & Beall, D. (1986). Exploring relationships among four science teaching-learning affective attributes of sixth grade students. *Journal of Research in Science Teaching, 23*(1), 51-60. doi: 10.1002/tea.3660230106

*Hassan, G. (2008). Attitudes toward science among Australian tertiary and secondary school students. *Research in Science & Technological Education, 26*(2), 129-147. doi:10.1080/02635140802034762

*Henderson, B. B., Marx, M. H., & Kim, Y. C. (1999). Academic interests and perceived competence in American, Japanese and Korean children. *Journal of Cross-Cultural Psychology, 30*(1), 2-50. doi:10.1177/0022022199030001002

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33,* 61–83. doi:10.1017/ S0140525X0999152X

Higgins, J. P. T., & Green, S. (Eds). (2011). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins, J., Thompson, S., Deeks, J., & Altman, D. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557–560.doi:10.1136/bmj.327.7414.557

Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In *Girls at Puberty* (pp. 201-228). Springer US.

Hines, M., Fane, B. A., Pasterski, V. L., Mathews, G. A., Conway, G. S., & Brook, C. (2003). Spatial abilities following prenatal androgen abnormality. *Psychoneuroendocrinology, 28,* 1010–1026. doi: 10.1016/s0306-4530(02)00121-x

Hines, M., & Kaufman, F. R. (1994). Androgen and the development of human sex-typical behavior—Rough-and-tumble play and sex of preferred playmates in children with CAH. *Child Development, 65,* 1042–1053. doi: 10.2307/1131303

Hogervorst, E., Bandelow, S., & Moffat, S. D. (2005). Increasing testosterone levels and effects on cognitive functions in elderly men and women. *Drug Targets—CNS & Neurological Disorders, 4,* 531–540. doi: 10.2174/1568007054322049

Hollingworth, L. S. (1918). Comparison of the sexes in mental traits. *Psychological Bulletin, 15,* 427–432. doi:10.1037/h0072007

*Hong, Z. R., & Lin, H. S. (2011). An investigation of students' personality traits and attitudes toward science. *International Journal of Science Education*, 33(7), 1001-1028. doi:10.1080/09500693.2010.524949

Hugdahl, K., Thomsen, T., & Ersland, L. (2006). Sex differences in visuo-spatial processing: An fMRI study of mental rotation. *Neuropsychologia, 44,* 1575–1583. doi:10.1016/j.neuropsychologia.2006.01.026

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*(4), 880. doi:10.1037/a0019506

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410-1412. doi: 10.1126/science.1177067

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60(*6), 581. doi:10.1037/0003-066X.60.6.581

Hyde, J. S. (2007). New directions in the study of gender similarities and differences. *Current Directions in Psychological Science, 16*(5), 259-263. doi: 10.1111/j.1467-8721.2007.00516.x

Hyde, J. S. (2012). Nation-level indicators of gender equality in psychological research: Theoretical and methodological issues. *Psychology of Women Quarterly, 36*(2), 145-148. doi: 10.1177/0361684312441448

Hyde, J. S. (2013). Gender similarities and differences. *Annual review of psychology, 65*, 373-398. doi:10.1146/annurev-psych-010213-115057

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139-155. doi:10.1037/0033-2909.107.2.139

Hyde, J. S., & Linn, M. C. (2006). Gender similarities in mathematics and science. *Science, 314*(5799), 599-600.

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences*, *106*(22), 8801-8807.

*Inda, M., Rodriguez, C., & Pena, J. V. (2013). Gender differences in applying social cognitive career theory in engineering students. *Journal of Vocational Behavior*, *83*(3): 346-355. doi: 10.1016/j.jvb.2013.06.010

Intemann, K. (2009). Why diversity matters: Understanding and applying the diversity component of the National Science Foundation's broader impacts criterion. *Social Epistemology*, *23*(3-4), 249-266. doi: 10.1080/02691720903364134

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

Jacobs, J. A., & Lim, S. T. (1992). Trends in occupational and industrial sex segregation in 56 countries, 1960-1980. *Work and Occupations*, *19*(4), 450-486. doi:10.1177/0730888492019004006

Jaffee, S., & Hyde, J. S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin, 126,* 703–726. doi:10.1037//0033-2909.126.5.703

Jenkins, S. R. (2000). Introduction to the special issue: Defining gender, relationships, and power. *Sex Roles, 42,* 467–493. doi:10.1023/A:1007010604246

Jerrim, J. & Schoon, I. (2014). Do teenagers want to become scientists? A comparison of gender differences in attitudes towards science, career expectations, and academic skill across 29 countries. In I. Schoon, & J. S. Eccles (Eds.), *Gender differences in aspirations and attainment: A life course perspective* (pp. 247-266). Cambridge, United Kingdom: Cambridge University Press.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-26. doi: 10.3102/0013189x033007014

*Jovanovic, J., & King, S. S. (1998). Boys and girls in the performance-based science classroom: Who's doing the performing? *American Educational Research Journal 35*(3), 477-496. doi:10.3102/00028312035003477

*Kahle, J. B., & Damnjanovic, A (1994). The effect of inquiry activities on elementary students' enjoyment, ease, and confidence in doing science: An analysis by sex and race. *Journal of Women and Minorities in Science and Engineering, 1*(1). 17-28. doi:10.1615/jwomenminorscieneng.v1.i1.20

Kane, J. M., and J. E. Mertz. (2012). Debunking myths about gender and mathematics performance. *Notices of the AMS* 59:10-21. doi:10.1090/noti790

*Keller, C. (2001). Effect of teachers' stereotyping on students' stereotyping of mathematics as a male domain. *The Journal of Social Psychology*, *141*(2), 165-173. doi:10.1080/00224540109600544

*Kelley, M. J., & Decker , E. O. (2009). The current state of motivation to read among middle school students. *Reading Psychology, 30*(5), 466-485. doi: 10.1080/02702710902733535

Kennedy, J. P., Lyons, T., & Quinn, F. (2014). The continuing decline of science and mathematics enrolments in Australian high schools. *Teaching Science, 60*(2), 34-46. Accessed online December 2016http://eprints.qut.edu.au/73153/1/Continuing_decline_of_science_proof.pdf

*Khoury, G. A., & Voss, B. E. (1985). *Factors influencing high school students' science enrollments patterns: Academic abilities, parental influences, and attitudes toward science*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching.

*Kilic-Bebek, E. (2010). *Explaining math achievement: Personality, motivation, and trust* (Unpublished doctoral dissertation). Cleveland State University, Cleveland.

*Kim, M. S., & Seo, Y. S. (2014). Social cognitive predictors of academic interests and goals in South Korean engineering students. *Journal of Career Development, 41*(6): 526-546. doi:10.1177/0894845313519703

Kimura, D. (1999). *Sex and cognition.* Cambridge, MA: MIT Press.

Kindermann, T. A. (1993). Natural peer groups as contexts for individual development: The case of children's motivation in school. *Developmental Psychology*, *29*(6), 970. doi: 10.1037//0012-1649.29.6.970

*Kissau, S. (2006). Gender differences in motivation to learn French. *Canadian Modern Language Review*, *62*(3), 22. doi:10.3138/cmlr.62.3.401

Klasen, S. (2006) UNDP’s gender-related measures: some conceptual problems and possible solutions. *Journal of Human Development, 7*(2), 243–74. doi:10.1080/14649880600768595

KPMG (2016). *She’s Price(d)less: The economics of the gender pay gap.* Report prepared for Diversity Council Australia (DCA) and the Workplace Gender Equality Agency (WGEA). Retrieved from https://home.kpmg.com/content/dam/kpmg/au/pdf/2016/gender-pay-gap-economics-full-report.pdf

Knickmeyer, R. C., & Baron-Cohen, S. (2006). Fetal testosterone and sex differences in typical social development and in autism. *Journal of Child Neurology, 21,* 825–845. doi:10.1177/08830738060210101601

*Koh, C. K. (2012). *Adolescents' motivation to study music as compared to other school subjects: A Singaporean perspective*. University of Illinois, Urbana Champaign.

*Koohang, A. A. (1989). A study of attitudes toward computers: Anxiety,confidence, liking, and perception of usefulness. *Journal of Research on Computing in Education, 22*(2), 137-150. doi: 10.1080/08886504.1989.10781909

*Kyttala, M., & Bjorn, P. M. (2010). Prior mathematics achievement, cognitive appraisals and anxiety as predictors of Finnish students' later mathematics performance and career orientation. *Educational Psychology, 30*(4), 431-448. doi:10.1080/01443411003724491

Lane, K. A., Goh, J. X., & Driver-Linn, E. (2012). Implicit science stereotypes mediate the relationship between gender and academic participation. *Sex Roles, 66*(3-4), 220-234. doi:10.1007/s11199-011-0036-z

Lareau, A. (2003). *Unequal childhoods: Race, class and family life.* Berkeley: University of California Press.

*Lau, S. (2003). *Cognitive abilities and motivational processes in high school students' science achievement and engagement.* (Unpublished doctoral dissertation). Stanford University, Stanford.

Leaper, C., Farkas, T., & Brown, C. S. (2012). Adolescent girls' experiences and gender-related beliefs in relation to their motivation in math/science and English. *Journal of Youth and Adolescence, 41*(3), 268-282. doi: 10.1007/s10964-011-9693-z

Leaper, C., & Van, S. R. (2008). Masculinity ideology, covert sexism, and perceived gender typicality in relation to young men's academic motivation and choices in college. *Psychology of Men & Masculinity, 9*(3), 139. doi:10.1037/1524-9220.9.3.139

*Lee, W., Lee, M. J., & Bong, M. (2014). Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemporary Educational Psychology, 39*(2), 86-99. doi:10.1016/j.cedpsych.2014.02.002

*Lefevre, J. A., Kulak, A. G., & Heymans, S. L. (1992). Factors influencing the selection of university majors varying in mathematical content. *Canadian Journal of Behavioural Science-Revue Canadienne Des Sciences Du Comportement, 24*(3), 276-289. doi:10.1037/h0078742

*Leibham, M. B., Alexander, J. M., & Johnson, K. E. (2013). Science interests in preschool boys and girls: Relations to later self-concept and science achievement. *Science Education, 97*(4), 574-593. doi:10.1002/sce.21066

*Lent, R. W., et al. (2005). Social cognitive predictors of academic interests and goals in engineering: Utility for women and students at historically Black universities. *Journal of Counseling Psychology, 52*(1), 84-92. doi: 10.1037/0022-0167.52.1.84

*Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1993). Predicting mathematics-related choice and success behaviors: Test of an expanded social cognitive model. *Journal of Vocational Behavior, 42*(2), 223-236. doi:10.1006/jvbe.1993.1016

*Levin, J., & Fowler, H. (1984). Sex, grade, and course differences in attitudes that are related to cognitive performance in secondary science. *Journal of Research in Science Teaching, 21*(2), 151-166. doi:10.1002/tea.3660210207

*Levin, J., & Klindienst, D. (1983). *Differences in attitudes between academic continuing and academic terminal secondary science students*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (Dallas, TX, April, 1983).

Lewin, K. (1938). *The conceptual representation and measurement of psychological forces*. Durham, NC: Duke University Press.

Lewin, M. (1984). *In the shadow of the past: Psychology portrays the sexes: A social and intellectual history*. Columbia University Press.

Lewontin, R. (2000). *It ain't necessarily so: The dream of the human genome and other illusions*. New York: New York Review of Books.

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

*Lindsay, H. A. (2002). *Factors related to achievement in sophomore organic chemistry at the University of Arkansas* (Unpublished doctoral dissertation). University of Arkansas, Fayetteville.

*Liou, P. Y. &, Kuo. P. J. (2014). Validation of an instrument to measure students' motivation and self-regulation towards technology learning. *Research in Science & Technological Education, 32*(2): 79-96. doi: 10.1080/02635143.2014.893235

Lips, H. M. (2017). *A New Psychology of Women: Gender, Culture, and Ethnicity. 4th edition.* Waveland Press Inc: Long Grove, Illinois

Lopez-Claros, A., & Zahidi, S. (2005). *Women's empowerment: Measuring the global gender gap*. Geneva: World Economic Forum. Retrieved from http://www.weforum.org/pdf/Global_Competitiveness_Reports/Reports/gender_gap.pdf

Lubienski, S. T., Crane, C. C., & Robinson, J. P. (2011). *A longitudinal study of gender and mathematics using ECLS data.* Final report (grant#

R305A080147) submitted to the National Center for Education Research, Institute of Education Sciences, Washington, DC.

Lubienski, S. T., Robinson, J. P., Crane, C. C., & Ganley, C. M. (2013). Girls' and Boys' Mathematics Achievement, Affect, and Experiences: Findings from ECLS-K. *Journal for Research in Mathematics Education, 44*(4), 634-645. doi:10.5951/jresematheduc.44.4.0634

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

Luster, T., & McAdoo, H. (1996). Family and child influences on educational attainment: A secondary analysis of the high/scope Perry Pre-school data. *Developmental Psychology, 32,* 26–39. doi: 10.1037//0012-1649.32.1.26

Lykes, M. B. (2000). Possible contributions of a psychology of liberation: Whither health and human rights?. *Journal of Health Psychology*, *5*(3), 383-397. doi: 10.1177/135910530000500312

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Mack, J., & Walsh, B. (2013). *Mathematics and science combinations NSW HSC 2001-2011 by gender.* Technical Paper, University of Sydney. Retrieved from http://www.maths.usyd.edu.au/u/SMS/MWW2013.pdf

Mack, J. & Wilson, R. (2015). Trends in mathematics and science subject combinations in the NSW HSC 2001-2014 by Gender. Technical paper, University of Sydney. Retrieved from: http://www.maths.usyd.edu.au/u/SMS/MMW2015.pdf

Machin S, Pekkarinen T (2008). Global sex differences in test score variability. *Science,* 322:1331–1332. doi: 10.1126/science.1162573

Magnusson, E., & Maracek, J (2012). Gender and culture in psychology: Theories and practices. New York: Cambridge University Press

*Makrakis, V. (1992). Cross-cultural comparison of gender differences in attitude towards computers in Japan and Sweden. *Scandinavian Journal of Educational Research, 36*(4), 275-287. doi: 10.1080/0031383920360403

Malouf, M. A., Migeon, C. J., Carson, K. A., Petrucci, L., & Wisniewski, A. B. (2006). Cognitive outcome in adult women affected by CAH due to 21-

hydroxylase deficiency. *Hormone Research, 65,* 142–150. doi: 10.1159/000091793

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

Mann, A., & DiPrete, T. A. (2016). The consequences of the national math and science performance environment for gender differences in STEM aspiration. *Sociological Science*, *3*, 568. doi:10.15195/v3.a25

Mann, A., Legewie, J., & DiPrete, T. A. (2015). The role of school performance in narrowing gender gaps in the formation of STEM aspirations: a cross-national study. *Frontiers in Psychology*, *6*, 171. doi:10.3389/fpsyg.2015.00171

Manning, J. T. (2002). *Digit ratio: A pointer to fertility, behavior, and health*. Rutgers University Press.

*Marinak, B. A., & Gambrell, L. B. (2010). Reading motivation: Exploring the elementary gender gap. *Literacy Research and Instruction*, *49*(2), 129-141. doi:10.1080/19388070902803795

Marks, J. L., Lam, C. B., & McHale, S. M. (2009). Family patterns of gender role attitudes. *Sex Roles, 61*(3-4), 221-234. doi:10.1007/s11199-009-9619-3

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38*(1), 299-337. doi:10.1146/annurev.psych.38.1.299

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology*, *81*(3), 417. doi:10.1037//0022-0663.81.3.417

Marsh, H. W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, *2*(2), 77-172. doi:10.1007/BF01322177

Marsh, H. W. (1991). Failure of high ability schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal*, *28*(2), 445-480. doi:10.3102/00028312028002445

Marsh, H. W. (1993). The multidimensional structure of academic self-concept: Invariance over gender and age. *American Educational Research Journal, 30,* 841-860. doi:10.3102/00028312030004841

*Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J., Abelfattah, F., Leung, K. C., Xu, M. K., Nagengast, B., & Parker, P. D. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology, 105*(1), 108-128. doi:10.1037/a0029907

Marsh, H. W., Bornmann, L., Mutz, R., Hans-Dieter, D., & O'Mara, A. (2009). Gender effects in the peer reviews of grant proposals: A comprehensive meta-analysis comparing traditional and multilevel approaches. *Review of Educational Research, 79*, (3), 1290-1326. doi:10.3102/0034654309334143

*Marsh, H. W., Trautwein, U., Ludtke, O., Koller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2): 397-416. doi:10.1111/j.1467-8624.2005.00853.x

*Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal*, *35*(4), 705-738. doi:10.3102/00028312035004705

*Matthews, G. D. (2009). *Mapping the relationship between a learner's belief system and their ability to learn competencies* (Unpublished doctoral dissertation). Capella University, Minneapolis.

McGraw, R., Lubienski, S. T., & Strutchens, M. E. (2006). A closer look at gender in NAEP mathematics achievement and affect data: Intersections with achievement, race/ethnicity, and socioeconomic status. *Journal for Research in Mathematics Education, 37*(2), 129-150. Retrieved from http://www.jstor.org/stable/30034845

McKeever, W. F., Rich, D., Deyo, R., & Conner, R. (1987). Androgens and spatial ability: Failure to find a relationship between testosterone and ability measures. *Bulletin of the Psychonomic Society, 25*, 438–440. doi:10.3758/bf03334734

*Medeiros, D. J. (2012). *The influence of female social models in corporate STEM initiatives on girls' math and science attitudes* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia.

*Meece, J. L. (1981). *Individual differences in the affective reactions of middle and high school students to mathematics: A social cognitive perspective.* (Unpublished doctoral dissertation). University of Michigan, Ann Arbor.

*Meelissen, M. and H. Luyten (2008). The Dutch gender gap in mathematics: Small for achievement, substantial for beliefs and attitudes. *Studies in Educational Evaluation, 34*(2), 82-93. doi: 10.1016/j.stueduc.2008.04.004

Melkas, H., & Anker, R. (1997). Occupational segregation by sex in Nordic countries: An empirical investigation. *International Labour Review, 136*(3), 341-365. Retrieved from http://heinonline.org.ezproxy1.acu.edu.au/HOL/Page?public=false&handle=hein.journals/intlr136&page=341&collection=journals

*Miller, M. (2010). *An investigation of perceived anxiety toward new software technologies among teachers in a Mississippi rural city school district* (Unpublished doctoral dissertation). Mississippi State University, Starkville.

*Mills, N., Pajares, F., & Herron, C. (2007). Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language Learning, 57*(3), 417-442. doi:10.1111/j.1467-9922.2007.00421.x

*Miura, I. T. (1987). The relationship of computer self-efficacy expectations to computer interest and course enrollment in college. *Sex Roles, 16*(5). doi:10.1007/bf00289956

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educational Research and Evaluation, 21,* 343–370. doi:10.1080/13803 611.2015.1057161.

Moffat, S. D., Zonderman, A., Metter, E., Blackman, M., Harman, S., & Resnick, S. M. (2002). Longitudinal assessment of serum free testosterone concentration predicts memory and cognitive status in men. *Journal of Clinical Endocrinology and Metabolism, 87,* 5001–5007. doi: 10.1210/jc.2002-020419

Moss-Racusin, C. A. , Dovidio, J. F., Brescoll, V. L., M. J. Graham, & Handelsman. J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences, 109*(41), 16474-16479. doi:10.1073/pnas.1211286109

*Nagy, G., Trautwein, U., Baumert, J., Koller, O., & Garrett, J. (2006). Gender and course selection in upper secondary education: Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation, 12*(4), 323-345. doi:10.1080/13803610600765687

Nagy, G., Watt, H. M., Eccles, J. S., Trautwein, U., Lüdtke, O., & Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: Different countries, different trajectories?. *Journal of Research on Adolescence, 20(*2), 482-506. doi:10.1111/j.1532-7795.2010.00644.x

Nash, A., & Grossi, G. (2007). Picking Barbie™'s brain: Inherent sex differences in scientific ability?. *Journal of Interdisciplinary Feminist Thought, 2*(1), 5.

National Center for Education Statistics (2015). "Table 318.30: Bachelor's, Master's and Doctor's Degrees Conferred by Postsecondary Institutions, By Sex of Student and Discipline Division: 2013-14", *Digest of Education Statistics*. Retrieved from https://nces.ed.gov/programs/digest/d15/tables/dt15_318.30.asp

National Innovation and Science Agenda, Australian Government (2017). *National innovation and science agenda*. Retrieved from http://www.innovation.gov.au/page/agenda

National Science Foundation, Division of Science Resources Statistics. (2012). *Women, minorities, and persons with disabilities in science and engineering* (Special Report NSF 11-309). Arlington, VA. Retrieved from http://www.nsf.gov/statistics/wmpd/sex

National Science and Technology Council, Office of Science Technology Policy, US. Government Washington D.C. (2013). https://www.whitehouse.gov/sites/default/files/microsites/ostp/stem_stratplan_2013.pdf

*Negishi, M. (2007). *A cross-cultural, multilevel study of inquiry-based instruction effects on conceptual understanding and motivation in physics* (Unpublished doctoral dissertation). Mississippi State University, Starkville.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics, 122,* 10671101. doi:10.1162/qjec.122.3.1067

Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *The Journal of Economic Perspectives, 24,* 129-144. doi:10.1257/jep.24.2.129

Noble, J. H. (2006). Meta-analysis: methods, strengths, weaknesses, and political uses. *Journal of Laboratory and Clinical Medicine, 147*(1), 7-20. doi:10.1016/j.lab.2005.08.006

Nosek, B., Smyth, F., Sriram, N., Lindner, N., Devos, T., Ayala, A., . . . Greenwald, A. G. (2009). National differences in gender science stereotypes predict national sex differences in science and achievement. *Proceedings of the National Academy of Sciences, 106,* 10593-10597. doi:10.1073/pnas.0809921106

*Oberman, P. S. (2002). *Academic help-seeking in the high school computer science classroom: Relationship to motivation, achievement, gender, and ethnicity* (Unpublished doctoral dissertation). Emory University, Atlanta.

*Obrentz, S. B. (2012). *Predictors of science success: The impact of motivation and learning strategies on college chemistry performance* (Unpublished doctoral dissertation). Georgia State University, Atlanta.

Office of the Chief Scientist (2014). *Science, technology, engineering and mathematics: Australia's future.* Australian Government, Canberra. Retrieved from http://www.chiefscientist.gov.au/wp-content/uploads/STEM_AustraliasFuture_Sept2014_Web.pdf

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist, 41*, 181–206. doi:10.1207/s15326985ep4103_4

Page, S. E. (2007). Making the difference: Applying a logic of diversity. *The Academy of Management Perspectives*, *21*(4), 6-20. doi:10.5465/amp.2007.27895335

*Pajares, F., & Miller, M. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86*(2). 193-203. doi:10.1037//0022-0663.86.2.193

*Pajares, F., & Viliante, G. (1997). Influence of self-efficacy on elementary students' writing. *Journal of Educational Research, 90*(6): 353-360. doi: 10.1080/00220671.1997.10544593

*Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology, 24*(4), 390-405. doi:10.1006/ceps.1998.0995

*Pajares, F., & Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: A function of gender orientation? *Contemporary Educational Psychology, 26*(3), 366-381. doi:10.1006/ceps.2000.1069

Parker, J. (2017). *Australia's jobs market highly gender-segregated, little change over past 20 years.* Australian Broadcasting Commission. Retrieved from http://www.abc.net.au/news/2017-02-10/men-entering-health-pink-collar-jobs/8258424

Parker, P. D., Jerrim, J., Anders, J., & Astell-Burt, T. (2016). Does living closer to a university increase educational attainment? A longitudinal study of aspirations, university entry, and elite university enrolment of Australian youth. *Journal of Youth and Adolescence*, *45*(6), 1156-1175. doi: 10.1007/s10964-015-0386-x

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-termachievement outcomes. *Educational Psychology, 34*(1), 29-48. doi: 10.1080/01443410.2013.797339

Parker, P., Nagy, G., Trautwein, U., & Ludtke, O. (2014). Predicting careeraspirations and university majors from academic ability and self-concept: Alongitudinal applications of the internal-external frame of reference model. InI. Schoon, & J. S. Eccles (Eds.), *Gender differences in aspirations and attainment: A life course perspective* (pp. 247-266). Cambridge, UnitedKingdom: Cambridge University Press.

Parker, P. D., Schoon, I., Tsai, Y., Nagy, G., Trautwein, U., & Eccles, J. (2012). Achievement, agency, gender, and socioeconomic background as predictors of postschool choices: A multi-context study. *Developmental Psychology, 48*(6), 1629–1642. doi:10.1037/a0029167

Parsons, J. E., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development, 53,* 310–321. doi: 10.2307/1128973

*Paslov, L. S. (2007). *The effect of a piloted middle school pre-engineering program on girls' interest and achievement in mathematics* (Unpublished doctoral dissertation) Southern Connecticut State University, New Haven.

*Patrick, H., Mantzicopoulos, P., & Samarapungayan, A. (2009). Motivation for learning science in Kindergarten: Is there a gender gap and does integrated Inquiry and Literacy Instruction Make a Difference. *Journal of Research in Science Teaching, 46*(2), 166-191. doi: 10.1002/tea.20276

*Peklaj, C., Podlesek, A., & Pecjak, S. (2014). Gender, previous knowledge, personality traits and subject-specific motivation as predictors of students' math grade in upper-secondary school. *European Journal of Psychology of Education, 30*(3), 313-330. doi:10.1007/s10212-014-0239-0

*Pell, A. W. (1985). Enjoyment and attainment in secondary school physics. *British Educational Research Journal, 11*(2), 123-132. doi: 10.1080/0141192850110205

Penner, A. M. (2008). Gender differences in extreme mathematical achievement: An international perspective on biological and social factors. *American Journal of Sociology, 114*(S1), S138-S170. doi:10.1086/589252

Perez, T., Cromley, J. G., Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology*, *106*(1), 3015-329. doi:10.1037/a0034027

Pettigrew, T. F. (1991). Toward unity and bold theory: Popperian suggestions for two persistent problems of social psychology. In C. W. Stephan, W. G. Stephan & T. F. Pettigrew (Eds.), The future of social psychology (pp. 13–27). New York, NY: Springer-Verlag.

Pinker, S. (2008). *The sexual paradox: Men, women, and the real gender gap.* Scribner: New York/London/Toronto/Sydney

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

*Pinxten, M., Marsh, H. W., De Fraine, B., Van den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology*, *84*(1), 152-174. doi:10.1111/bjep.12028

*Plante, I., de la Sablonniere, R., Aronson, J. M., & Theoret, M. (2013). Gender stereotype endorsement and achievement-related outcomes: The role of competence beliefs and task values. *Contemporary Educational Psychology, 38*(3), 225-235. doi:10.1016/j.cedpsych.2013.03.004

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

Priess-Groben, H. A., & Hyde, J. S. (2017). Implicit theories, expectancies, and values predict mathematics motivation and behavior across high school and college. *Journal of Youth and Adolescence*, *46*(6), 1318-1332. doi:10.1007/s10964-016-0579-y

Prilleltensky, I. (2008). The role of power in wellness, oppression, and liberation: The promise of psychopolitical validity. *Journal of Community Psychology, 36*(2), 116-136. doi:10.1002/jcop.20225

*Quihuis, G. (2002). *Understanding the factors that influence high science achievers' academic choices and intent to pursue or opt out of the hard sciences* (Unpublished doctoral dissertation). Stanford University, Stanford.

*Quinn, F., & Lyons, T. (2011). High school students' perceptions of school science and science careers: A critical look at a critical issue. *Science Education International*, 225-238.

Raudenbush & Bryk, 1985 Raudenbush, S. W., & Bryk, A. S. (1985). Empirical Bayes meta-analysis. *Journal of Educational and Behavioral Statistics, 10,* 75-98. doi:10.2307/1164836

*Rech, J. F. (1994). A comparison of the mathematics attitudes of Black students according to grade level, gender, and academic achievement. *Journal of Negro Education, 63*(2), 212-220. doi:10.2307/2967384

Reeve, J., & Deci, E. L (1996). Elements within the competitive situation that affect intrinsic motivation. *Personality and Social Psychology Bulletin, 22,* 24-33. doi:10.1177/0146167296221003

Resnick, S. M., Berenbaum, S. A., Gottesman, I. I., & Bouchard, T. J. (1986). Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology, 22,* 191–198. doi:10.1037/0012-1649.22.2.191

Ridgeway, C. L. (2009). Framed before we know it: How gender shapes social relations. *Gender & Society, 23*(2), 145-160. doi:10.1177/0891243208330313

Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system a theoretical perspective on gender beliefs and social relations. *Gender & Society, 18*(4), 510-531. doi:10.1177/0891243204265269

Ripa, C. P. L., Johannsen, T. H., Mortensen, E. L., & Muller, J. (2003). General cognitive functions, mental rotations ability, and handedness in adult females with CAH. *Hormones and Behavior, 44*, 72.

*Riconscente, M. M. (2014). Effects of perceived teacher practices on Latino high school students' interest, self-efficacy, and achievement in mathematics. *Journal of Experimental Education, 82*(1), 51-73. doi: 10.1080/00220973.2013.813358

*Riegle-Crumb, C., Moore, C., & Ramos-Wada, A. (2011). Who wants to have a career in science or math? Exploring adolescents' future aspirations by gender and race/ethnicity. *Science Education, 95*(3), 458-476. doi:10.1002/sce.20431

Reis, H. T., & Carothers, B. J. (2014). Black and white or shades of gray: Are gender differences categorical or dimensional?. *Current Directions in Psychological Science, 23*(1), 19-26. doi:10.1177/0963721413504105

Reis, H. T., & Wright, S. (1982). Knowledge of sex-role stereotypes in children aged 3 to 5. *Sex Roles*, *8*(10), 1049-1056. doi:10.1007/bf00290999

Reuben, E., Sapienza, P., & Zingales. L. How stereotypes impair women's careers in science. *Proceedings of the National Academy of Sciences, 111*(12), 4403-4408. doi:10.1073/pnas.1314788111

Robnett, R. D., & Leaper, C. (2013). Friendship groups, personal motivation, and gender in relation to high school students' STEM career interest. *Journal of Research on Adolescence, 23*(4), 652-664. doi:10.1111/jora.12013

Romanes G J (1887). Mental differences between men and women. *Nineteenth Century* 21:654 – 672.

Roos, P. A. (1985). *Gender and work: A comparative analysis of industrial societies.* Albany, NY: State University of New York Press.

Rosenfeld, R. A., & Kalleberg, A. L. (1991). Gender inequality in the labor market: A cross-national perspective. *Acta Sociologica, 34*(3), 207-225. doi: 10.1177/000169939103400304

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Vol. 6). Newbury Park, CA: Sage.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). New York, NY: Russell Sage Foundation.

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

Rushton, J. P. (1992a). Cranial capacity related to sex, rank, and race in a stratified random sample of 6,325 U.S. military personnel. *Intelligence, 16,* 401–413. doi:10.1016/0160-2896(92)90017-l

Rushton, J. P. (1992b). The brain size IQ debate. *Nature*, *360*(6402), 292. doi: 10.1038/360292a0

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68–78. doi:10.1037/0003-066x.55.1.68

Sáinz, M., & López-Sáez, M. (2010). Gender differences in computer attitudes and the choice of technology-related occupations in a sample of secondary students in Spain. *Computers & Education, 54*(2), 578-587. doi: 10.1016/j.compedu.2009.09.007

Sandberg, S. (2013). *Lean in: Women, work, and the will to lead* (First edition.). New York: Alfred A. Knopf.

Sanders, G., Bereczkei, T., Csatho, A., & Manning, J. (2005). The ratio of the 2nd to 4th finger length predicts spatial ability in men but not women. *Cortex, 41,* 789–795. doi:10.1016/s0010-9452(08)70297-1

*Sasson, I., & Cohen, D. (2013). Assessment for effective intervention: Enrichment science academic program. *Journal of Science Education and Technology, 22*(5), 718-728. doi:10.1007/s10956-012-9425-5

Savin-Williams, R. C. (2014). An exploratory study of the categorical versus spectrum nature of sexual orientation. *The Journal of Sex Research, 51*(4), 446-453. doi:10.1080/00224499.2013.871691

Schafer, R. (1974). Problems in Freud's psychology of women. *Journal of the American Psychoanalytic Association, 22*(3), 459-485. doi: 10.1177/000306517402200301

Schattman, L., & Sherwin, B. B. (2007). Effects of the pharmacologic manipulation of testosterone on cognitive functioning in women with polycystic ovary syndrome: A randomized, placebo-controlled treatment study. *Hormones and Behavior, 51,* 579–686. doi:10.1016/j.yhbeh.2007.02.002

Scheibinger, L. (1987). The history and philosophy of women in science: A review essay. *Signs: Journal of Women in Culture and Society*, *12,* 305–332.

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

Science in Australia Gender Equality (2017). *Gender Equality in STEMM*. Retrieved from https://www.sciencegenderequity.org.au/gender-equity-in-stem/

*Selkirk, L. C., Bouchey, H. A., & Eccles, J. S. (2011). Interactions among domain-specific expectancies, values, and gender: Predictors of test anxiety during early adolescence. *The Journal of Early Adolescence*, *31*(3), 361-389.

Sells, L. W. (1980). Mathematics: The invisible filter. *Engineering Education, 70*(4), 340-341.

*Senler, B., & Sungur, S. (2009). Parental influences on students' self-concept, task value beliefs, and achievement in science. *The Spanish Journal of Psychology, 12*(1), 106-117. doi:10.1177/0272431610363156

*Sha, S. L. (2012). *Sociocultural and motivational factors affecting Asian American females studying physics and engineering in high school.* (Unpublished doctoral dissertation). University of Southern California, Los Angeles.

Sharps, M. J., Price, J. L., & Williams, J. K. (1994). Spatial cognition and gender: Instructional and stimulus influences on mental image rotation performance. *Psychology of Women Quarterly, 18*(3), 413-425. doi:10.1111/j.1471-6402.1994.tb00464.x

*Shashaani, L., & Khalili, A. (2001). Gender and computers: similarities and differences in Iranian college students' attitudes toward computers. *Computers & Education*, *37*(3-4), 363-375. doi:10.1016/s0360-1315(01)00059-8

*Sheldrake, R., Mutjaba, T., Reiss, M. J. (2014). Calibration of self-evaluations of mathematical ability for students in England aged 13 and 15, and their intentions to study non-compulsory mathematics after age 16. *International Journal of Educational Research, 64*, 49-61. doi:10.1016/j.ijer.2013.10.008

*Sherman, J. (1981). Girls' and boys' enrollments in theoretical math courses: A longitudinal study. *Psychology of Women Quarterly, 5*(5), 681-689. doi: 10.1111/j.1471-6402.1981.tb01092.x

Shields, S. (1975). Functionalism, Darwinism, and the psychology of women. *American Psychologist*, *30*(7), 739. doi:10.1037/h0076948

*Siegle, D., & Reis, S. M. (1998). Gender differences in teacher and student perceptions of gifted students' ability and effort. *Gifted Child Quarterly*, *42*(1), 39-47. doi:10.1177/001698629804200105

Sigerist H. E. (1951). *A history of medicine. Primitive and archaic medicine*. New York : Oxford University Press.

Sikora, J., & Pokropek, A. (2012). Gender segregation of adolescent science career plans in 50 countries. *Science Education, 96*(2), 234-264. doi: 10.1002/sce.20479

Sinclair, S., Hardin, C. D., & Lowery, B. S. (2006). Self-stereotyping in the context of multiple social identities. *Journal of Personality and Social Psychology, 90*(4), 529. doi:10.1037/0022-3514.90.4.529

*Skaalvik, S., & Skaalvik, E. M. (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. *Sex Roles, 50*(3-4), 241-252. doi:10.1023/B:SERS.0000015555.40976.e6

Skinner, B. F. (1953). *Science and human behavior.* Simon and Schuster.

Slabbekoorn, D., van Goozen, S. H., Megens, J., Gooren, L. J. G., & Cohen-Kettenis, P. T. (1999). Activating effects of cross-sex hormones on cognitive functioning. *Psychoneuroendocrinology, 24,* 432–447. doi:10.1016/s0306-4530(98)00091-2

Smeding, A. (2012). Women in science, technology, engineering, and mathematics (STEM): An investigation of their implicit gender stereotypes and stereotypes' connectedness to math performance. *Sex Roles, 67*(11-12), 617-629. doi: 10.1007/s11199-012-0209-4

*Smith, J. K., Smith, L. F., Gilmore, A., & Jameson, M. (2012). Students' self-perception of reading ability, enjoyment of reading and reading achievement. *Learning and Individual Differences, 22*(2), 202-206. doi: 10.1016/j.lindif.2011.04.010

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*(9), 752-760. doi:10.1037//0003-066x.32.9.752

Spelke, E. (2005). Sex differences in intrinsic aptitude for mathematics and science? A critical review. *American Psychologist, 60,* 950–958. doi:10.1037/0003-066x.60.9.950

Spierings, N. (2012). The inclusion of quantitative techniques and diversity in the mainstream of feminist research. *European Journal of Women's Studies, 19,* 331–347. doi:10.1177/1350506812443621

Squire, C. (1989). *Significant differences: Feminism in psychology.* Routledge: London and New York.

Stake, J. E. (2006). The critical mediating role of social encouragement for science motivation and confidence among high school girls and boys. *Journal of Applied Social Psychology, 36(*4), 1017-1045. doi:10.1111/j.0021-9029.2006.00053.x

Stake, J. E., & Nickens, S. D. (2005). Adolescent girls' and boys' science peer relationships and perceptions of the possible self as scientist. *Sex Roles, 52*(1-2), 1-11. doi:10.1007/s11199-005-1189-4

*Stanisavljevic, D., Trajkovic, G., Marinkovic, J., Bukumiric, Z., Cirkovic, A., & Milic, N. (2014). Assessing attitudes towards statistics among medical students: Psychometric properties of the Serbian Version of the Survey of Attitudes Towards Statistics (SATS). Plos One, *9(*11), 7. doi:10.1371/journal.pone.0112567

*Steiner, L. A. (2007). *The effect of personal and epistemological beliefs on performance in a college developmental mathematics class.* Kansas State University, Manhattan.

*Steinmayr, R., & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality, 22*(3), 185-209.

*Steinmayr, R., Wirthwein, L, Schone, C. (2014). Gender and numerical intelligence: Does motivation matter? *Learning and Individual Differences, 32,* 140-147.

*Stephanou, G. (2008). Students' value beliefs, performance expectations, and school performance: The effect of school subject and gender. *Hellenic Journal of Psychology 5*(3), 231-257. doi:10.1002/per.676

Sterpellone L. (2002). *Greek medicine.* 2nd ed. Noceto: Essebiemme

Stevens, J. S., & Hamann, S. (2012). Sex differences in brain activation to emotional stimuli: a meta-analysis of neuroimaging studies. *Neuropsychologia, 50*(7), 1578-1593. doi:10.1016/j.neuropsychologia.2012.03.011

*Stevens, T., Wang, K., Olivarez, A., & Hamman, D. (2007). Use of self-perspectives and their sources to predict the mathematics enrollment intentions of girls and boys. *Sex Roles,* 56(5-6), 351-363. doi:10.1007/s11199-006-9180-2

Stevenson, H. W., Chen, C., & Booth, J. (1990). Influences of schooling and urban-rural residence on gender differences in cognitive abilities and academic achievement. *Sex Roles, 23*(9-10), 535-551. doi: 10.1007/BF00289767

*Stevenson, H. W., & Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading. *Child Development*, *57*(3), 646-659. doi:10.2307/1130343

Stinson, S. (1985). Sex differences in environmental sensitivity during growth and development. *American Journal of Physical Anthropology*, *28*(S6), 123-147. doi:10.1002/ajpa.1330280507

Stoet, G., Bailey, D. H., Moore, A. M., & Geary, D. C. (2016). Countries with higher levels of gender equality show larger national sex differences in mathematics anxiety and relatively lower parental mathematics valuation for girls. *PloS one*, *11*(4), e0153857. doi:10.1371/journal.pone.0153857

Stotsky, J., Shibuya, S., Kolovich, L., & Kebhaj, S. (2016). *Trends in women's advancement and gender equality.* International Monentary Fund Working Paper, Washington DC, IMF).

Syed, M. (2010). Disciplinarity and methodology in intersectionality theory and research. *American Psychologist, 65,* 61–62. doi:10.1037/a0017495

*Tapia, M., & Marsh, G. E. (2000). *Effect of gender, achievement in mathematics, and ethnicity on attitudes toward mathematics.* Paper presented at the Annual Meeting of the Mid-South Educational Research Association.

Teddlie, C., & Tashakkori, A. (Eds.). (2010). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

*Tenenbaum, H. R., & Leaper, C. (2003). Parent-child conversations about science: The socialization of gender inequities?. *Developmental Psychology, 39*(1), 34. doi:10.1037//0012-1649.39.1.34

*Terwilliger, J. S., & Titus, J. C. (1995). Gender differences in attitudes and attitude changes among mathematically talented youth. *Gifted Child Quarterly, 39*(1), 29-35. doi:10.1177/001698629503900105

Thilers, P., MacDonald, S. W. S., & Herlitz, A. (2006). The association between endogenous free testosterone and cognitive performance: A population-based study. *Psychoneuroendocrinology, 31*, 565–576. doi: 10.1016/j.psyneuen.2005.12.005

Thorndike, E. L. (1914). *Educational psychology* (Vol. 3). New York: Teachers College, Columbia University.

*Thorndike-Christ, T. (1991). *Attitudes toward mathematics: Relationships to mathematics achievement, gender, mathematics course-taking plans, and career interests* (Unpublished doctoral dissertation). Western Washington University, Bellingham.

*Tissot, S. L. (1997). *An examination of ethnic variation in gender differences in mathematics attitudes and performance.* University of Maryland.

Tolman, E. C. (1932). *Purposive behavior in animals and men*. New York: Appleton-Century-Crofts

Trent, L. M. Y., Cooney, G., Russell, G., & Warton, P. M. (1996). Significant others' contribution to early adolescents' perceptions of their competence. *British Journal of Educational Psychology, 66*, 95–107. doi:10.1111/j.2044-8279.1996.tb01179.x

*Tsai, C. C., & Lin, C. C. (2004). Taiwanese adolescents' perceptions and attitudes regarding the internet: Exploring gender differences. *Adolescence San Diego, 39*(156), 725.

U.S. Bureau of Labor Statistics (2014). *STEM 101: Intro into tomorrow's jobs. Occupational Outlook Quarterly.* Retrieved from https://www.bls.gov/careeroutlook/2014/spring/art01.pdf

U.S. Department of Commerce (2012). *The competitiveness and innovative capacity of the United States*. U.S. Department of Commerce; Washington, D.C.

U.S. Department of Education. (2010). *Science, technology, engineering, and math: Education for global leadership.* Washington, DC: U.S. Department of Education.

*Uitto, A. (2014). Interest, attitudes and self-efficacy beliefs explaining upper-secondary school students' orientation towards biology-related careers.

International Journal of Science and Mathematics Education, 12(6), 1425-1444. doi:10.1007/s10763-014-9516-2

*Urhahne, D., Ho, L. H., Parchmann, I., & Nick, S. (2012). Attempting to predict success in the qualifying round of the International Chemistry Olympiad. *High Ability Studies, 23*(2), 167-182. doi:10.1080/13598139.2012.738324

Van Den Noortgate, W., & Onghena, P. (2003). Multilevel meta-analysis: A comparison with traditional meta-analytical procedures. *Educational and Psychological Measurement*, 63(5), 765-790. doi: 10.1177/0013164402251027

Van Goozen, S. H. M., Cohen-Kettenis, P. T., Gooren, L. J. G., Frijda, N. H., & Van de Poll, N. E. (1994). Activating effects of androgens on cognitive performance. *Neuropsychologia, 32*, 1153–1157. doi:10.1016/0028-3932(94)90099-x

Van Goozen, S. H. M., Cohen-Kettenis, P. T., Gooren, L. J. G., Frijda, N. H., & Van de Poll, N. E. (1995). Gender differences in behaviour: Activating effects of cross-sex hormones. *Psychoneuroendocrinology*, *20,* 343–363. doi: 10.1016/0306-4530(94)00076-x

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

*Vekiri, I. (2013). Information science instruction and changes in girls' and boy's expectancy and value beliefs: In search of gender-equitable pedagogical practices. *Computers & Education*, *64*, 104-115. doi: 10.1016/j.compedu.2013.01.011

*Vekiri, I., & Chronaki, A. (2008). Gender issues in technology use: Perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. *Computers & Education, 51*(3), 1392-1404. doi: 10.1016/j.compedu.2008.01.003

*Visser, D. (1986). *Sex differences in mathematics achievement and participation.* Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).

von Hippel, P. (2015). The heterogeneity statistic I^2 can be biased in small meta-analyses. *BMC Medical Research Methodology, 15*(1), 35. doi: 10.1186/s12874-015-0024-z

Walby, S., Armstrong, J., & Strid, S. (2012). Intersectionality: Multiple inequalities in social theory. *Sociology, 46,* 224–240. doi:10.1177/0038038511416164

Walker, E., Hernandez, A. V., & Kattan, M. W. (2008). Meta-analysis: Its strengths and limitations. *Cleveland Clinic Journal of Medicine*, 75(6), 431. doi: 10.3949/ccjm.75.6.431

*Walles, R. L. (2009). *The road to mathematics in elementary school: Social and cognitive influences on performance and response to intervention.* University of Massachusetts Amherst.

*Wang, J. (2012). *Untangling the relations among high school students' motivation, achievement and advanced course-taking in mathematics: Using structural equation modeling with complex samples.* University of Houston.

Wang, X. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal, 50(*5), 1081-1121. doi:10.3102/0002831213488622

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

*Wang, M. T., Degol, J., & Ye, F. F. (2015). Math achievement is important, but task values are critical, too: examining the intellectual and motivational factors leading to gender disparities in STEM careers. *Frontiers in Psychology*, *6*, 9. doi: 10.3389/fpsyg.2015.00036

Wang, M., & Kelly, S. (2014).Gender differences in personal aptitudes and motivational beliefs for achievement in and commitment to math and science fields. In I. Schoon, & J. S. Eccles (Eds.), *Gender differences in aspirations and attainment: A life course perspective* (pp. 247-266). Cambridge, United Kingdom: Cambridge University Press.

Wang, J., & Staver, J. R. (2001). Examining relationships between factors of science education and student career aspiration. *The Journal of Educational Research, 94*(5), 312-319. doi:10.1080/00220670109598767

Watt. H. M. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in the 7^{th}-through 11^{th} grade Australian students. *Child Development, 75*(5), 1556-1574. doi: 10.1111/j.1467-8624.2004.00757.x

Watt, H. M., & Eccles, J. S. (2008). *Gender and occupational outcomes: Longitudinal assessments of individual, social, and cultural influences.* American Psychological Association.

Watt, H. M., Eccles, J. S., & Durik, A. M. (2006). The leaky mathematics pipeline for girls: A motivational analysis of high school enrolments in Australia and the USA. *Equal Opportunities International, 25*(8), 642-659. doi: 10.1108/02610150610719119

*Weger-Guntharp, H. D. (2009). *Learner motivations and preferences: Realities in the language classroom* (Unpublished doctoral dissertation). Georgetown University, Georgetown.

*Weinberg, A. E., Basile, C. G., & Albright, L. (2011). The effect of an experiential learning program on middle school students' motivation toward mathematics and science. *RMLE Online: Research in Middle Level Education, 35*(3), 1-12.

*Weinburgh, M. H. (2000). *Gender, ethnicity, and grade level as predictors of middle school students' attitudes toward science* (Unpublished doctoral dissertation). Georgia State University, Atlanta.

Weiner, B. (1989/1980). *Human motivation.* Lawrence Erlbaum Associates, Hillsdale: NJ.

*Weisgram, E. S., & Bigler, R. S. (2006). Girls and science careers: The role of altruistic values and attitudes about scientific tasks. *Journal of Applied Developmental Psychology, 27*(4), 326-348. doi: 10.1016/j.appdev.2006.04.004

Weisgram, E. S., & Bigler, R. S. (2007). Effects of learning about gender discrimination on adolescent girls’ attitudes toward and interest in science. *Psychology of Women Quarterly, 31*(3), 262-269. doi:10.1111/j.1471-6402.2007.00369.x

Weisstein, N. (1968). *Kinder, Kuche, Kirche as scientific law: Psychology constructs the female* (pp. 228-245). Boston, MA: New England Free Press.

Weisstein, N. (1968/1973). Psychology constructs the female; or, the fantasy life of the male psychologist (with some attention to his friends, the male biologist and the male anthropologist). *Feminism & Psychology, 3*(2), 194-210. doi:10.1177/0959353593032005

Wenner, G. (2003). Comparing poor, minority elementary students' interest and background in science with that of their white, affluent peers. *Urban Education, 38*(2), 153-172. doi:10.1177/0042085902250483

Wentzel, K. R. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90,* 202–209. doi:10.1037//0022-0663.90.2.202

Westen, D. (1998). The scientific legacy of Sigmund Freud: toward a psychodynamically informed psychological science. *Psychological Bulletin, 124*(3), 333. doi:10.1037//0033-2909.124.3.333

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*(3), 265-310.

*Wigfield, A., & J. S. Eccles (1994). Children's competence beliefs, achievement values, and general self-esteem: Change across elementary and middle school. *Journal of Early Adolescence*, *14*(2), 107-138. doi:10.1177/027243169401400203

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

Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review, 30*(1), 1-35. doi:10.1016/j.dr.2009.12.001

*Wilhelm, S., & Brooks, D. M. (1980). The relationship between pupil attitudes toward mathematics and parental attitudes toward mathematics. *Educational Research Quarterly, 5*(2), 8-16.

Wilkins, J. L., & Ma, X. (2003). Modeling change in student attitude toward and beliefs about mathematics. *The Journal of Educational Research, 97*(1), 52-63. doi:10.1080/00220670309596628

Wittig, M. A. (1985). Metatheoretical dilemmas in the psychology of gender. *American Psychologist, 40,* 800–811. doi:10.1037//0003-066x.40.7.800

*Wolters, C. A., Denton, C. A., York, M. J., & Francis, D. J. (2014). Adolescents' motivation for reading: Group differences and relation to standardized achievement. *Reading and Writing: An Interdisciplinary Journal*, *27*(3), 503-533. doi:10.1007/s11145-013-9454-3

*Wolters, C. A., & Pintrich, P. R. (1998). Contextual differences in student motivation and self-regulated learning in mathematics, English, and social studies classrooms. *Instructional Science, 26*(1-2), 27-47. doi:10.1007/978-94-017-2243-8_6

Wood, W., & Eagly, A. H. (2002). Biological construction of sex differences and similarities in behavior. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 46, pp. 55-124). San Diego, CA: Academic Press.

Woolley, H. B. T. (1903). *The mental traits of sex: An experimental investigation of the normal mind in men and women.* Chicago: University of Chicago Press.

Woolley, H. T. (1914). The psychology of sex. *Psychological Bulletin, 11*, 353–379.

*Wong, K. Y. (2013). *Gender differences in scientific literacy of HKPISA 2006: A multidimensional differential item functioning and multilevel mediation study.* (Unpublished doctoral dissertation). The Chinese University of Hong Kong, Shatin, Hong Kong.

*Wong, S. L., & Hanafi, A. (2007). Gender differences in attitudes towards Information Technology among Malaysian student teachers: A case study at Universiti Putra Malaysia. *Educational Technology & Society, 10*(2), 158-169.

*Woods-McConney, A., Oliver, M. C., McConney, A., Schibeci, R., & Maor, D. (2014). Science engagement and literacy: A retrospective analysis for students in Canada and Australia. *International Journal of Science Education, 36*(10), 1588-1608.

Workplace Gender Equality Agency (2016). *Australia's gender equality scorecard: Key findings from the workplace gender equality agency's 2015-16 reporting data.* Australian Government. https://www.wgea.gov.au/sites/default/files/80653_2015-16-gender-equality-scorecard.pdf

Xie Y, Killewald A.A. (2012). *Is American Science in Decline?* Cambridge, MA.

*Yancy, Y. G. (2013). *The effects of project-based learning activities on intrinsic motivation and skill acquisition of rural middle school math students* (Unpublished doctoral dissertation). Union University, Jackson.

*Yang, F. Y., Tseng, J. S., & Lin, M. H. (2012). The interaction between junior-high students' academic and social motivations and the influences of the motivational factors on science performance. *Asia-Pacific Education Researcher, 21*(1), 92-106.

*Yeung, A. S., Lau, S., & Nie, Y. (2011). Primary and secondary students' motivation in learning English: Grade and gender differences. *Contemporary Educational Psychology, 36*(3), 246-256. doi:10.1016/j.cedpsych.2011.03.001

Yoder, J. D., & Kahn, A. S. (2003). Making gender comparisons more meaningful: A call for more attention to social context. *Psychology of Women Quarterly, 27*(4), 281-290. doi:10.1111/1471-6402.00108

*Zachai, J. (1999). *Adult learners' math self-concept as a barrier to passing California State University's Entry Level Mathematics (ELM) Test* (Unpublished doctoral dissertation). University of San Francisco, San Francisco.

Zarrett, N., Malanchuk, O., Davis-Kean, P. E., & Eccles, J. (2006). Examining the gender gap in IT by race: Young adults' decisions to pursue an IT career. In *Women and information technology: Research on underrepresentation*. Eds. J. McGrath and W. Asprey. MIT Press: Cambridge, Massachusetts, London

Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist*, *70*(1), 10-20. doi: 10.1037/a0038208.

Zurbriggen, E. L., & Capdevila, R. (2010). The personal and the political are feminist: Exploring the relationships among feminism, psychology, and political life. *Psychology of Women Quarterly, 34*(4), 458-459. doi: 10.1111/j.1471-6402.2010.01595.x

* References included in meta-analysis of Study 1

APPENDICES

Appendix A

Search Terms and Search Strategy for Meta-analysis

ERIC Search Strategy

(math* OR "verbal ability" OR English OR Science OR STEM) AND (gender OR sex) AND ("self-concept" OR expectancy OR "expectancy for success" OR "math self-concept" OR "verbal self-concept" OR "English self-concept" OR "science self-concept" OR "self-efficacy" OR competenc* OR attitude OR "expectancy value theory" OR "EVT") AND (“task value” OR value OR interest OR “intrinsic motivation” OR “intrinsic value” OR enjoyment OR importance OR “attainment value” OR usefulness OR "incentive value" OR "utility value" OR "extrinsic motivation")

Web of Science Search Strategy

(math* OR "verbal ability" OR English OR Science OR STEM) AND (gender OR sex) AND ("self-concept" OR expectancy OR "expectancy for success" OR "math self-concept" OR "verbal self-concept" OR "English self-concept" OR "science self-concept" OR "self-efficacy" OR competenc* OR attitude OR "expectancy value theory" OR "EVT") AND (“task value” OR value OR interest OR “intrinsic motivation” OR “intrinsic value” OR enjoyment OR importance OR “attainment value” OR usefulness OR "incentive value" OR "utility value" OR "extrinsic motivation

Psychinfo Search Strategy

1. math*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
2. "verbal ability".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
3. English.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
4. science.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
5. STEM.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
6. 1 or 2 or 3 or 4 or 5
7. gender.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
8. sex.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
9. 7 or 8
10. "self-concept".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
11. "expectancy".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
12. "expectancy for success".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
13. "math self-concept".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
14. "verbal self-concept".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
15. "English self-concept".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
16. "self-efficacy".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
17. competenc*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
18. attitude*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
19. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18
20. "task value".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
21. value.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
22. interest.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

23. "intrinsic motivation".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
24. "intrinsic value".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
25. "enjoyment".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
26. "importance".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
27. "attainment value".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
28. usefulness.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
29. "incentive value".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
30. "utility value".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
31. "extrinsic motivation".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
32. cost.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
33. 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32
34. "expectancy value theory".mp.
35. "EVT".mp.
36. 34 or 35
37. 6 and 9 and 19 and 33
38. 36 or 3

Appendix B

Eligibility Criteria for Meta-Analysis Studies (Study 1)

1) Includes both expectancy and value variables

Must include <u>ONE Expectancy variable</u> (examples include, self-concept or self-efficacy, perceived competence or ability, self-confidence, or an expectancy for success measure) AND at least <u>ONE value variable</u> (task value, interest, intrinsic value, attainment value, incentive value, utility value, extrinsic value, instrumental value, intrinsic motivation, importance, relevance)

2) The expectancy and value variables both need to be domain specific to the areas of math, science/technology and literacy

In other words, expectancy and value measures need to be specific to maths, literacy, or science/technology. General measures of expectancy or value (including general academic) are to be excluded.

Examples of domain specific measures include: maths self-concept, perceived competence in physics, reading self-efficacy, interest in computers, value of engineering etc.

Examples of measures that are not domain specific: STEM self-efficacy, a generalised measure including more than one domain (e.g., both science and maths together), general academic self-concept

3) Quantitative measures and statistical information required

Quantitative measures of gender/EVT variables (qualitative to be excluded). In other words, there needs to be some kind of quantitative result reported for the relationship between gender and the expectancy value variables (e.g., Means for males and females, correlation between gender and EVT variables, effect sizes of gender, regression path coefficients between gender and variables)

4) Full-text English language

Full-text English language (too many practical issues with non-English to be included unfortunately – this will be highlighted as a limitation of the current study)

5) New and original data

We will look at this once full-text is over in relation to duplicate data sets, but for now exclude systematic reviews, meta-analyses or literature reviews that summarise existing data as opposed to presenting new findings.

6) Participant Groups

The only restriction on participants is that the sample must include both male and female participants (single sex samples are to be excluded). There are no other restrictions on age or any other characteristic of participants. However, specify if sample is from a particular group (e.g., chemistry students, disadvantaged students, students from a particular ethnic group within country, students from rural/urban location)

7) Date of Publication and Publication Type

No restrictions on date of publication and publication type (theses, dissertations, and conference papers are acceptable in addition to peer reviewed articles and books)

Appendix C

Coding Information for Moderators Included in Meta-Analysis (Study 1)

Moderator	**Categories within Moderator**
Social Class	1 = Majority of sample identified as working class
	2 = Majority of sample identified as middle class
	3 = Majority of sample identified as upper class
	**Note: ‘Majority of sample’ is defined as >50% of sample from which effect sizes were extracted as being described by authors as either working class/low SES, middle class, or upper class/high SES.*
Ethnic minority %	1 = Lowest (0-25% ethnic minority sample)
	2 = Low (26-50% ethnic minority sample)
	3 = High (51-75% ethnic minority sample)
	4 = Highest (76-100% ethnic minority sample)
	**Note: ‘Ethnic minority percentage’ is defined as the percentage of sample from which effect sizes were extracted that were described as belonging to an ethnic minority of the country in which the study took place.*
Age group	1 = Elementary (mean age $\geq$ 5 and < 11 years)
	2 = Middle School (mean age $\geq$ 11 and < 14 years)
	3 = High School (mean age $\geq$ 14 and < 18 years)
	4 = Young Adult (mean age $\geq$ 18 and < 26 years)
	5 = Adult (26 years and above)
	**Note: Mean age of samples from which effect sizes were extracted was used. Grades were used to categories where age was not available.*
GII 2014	0 = GII value under .10
	1 = GII value under .20

	2 = GII value under .30
	3 = GII value above .30
	**Note: Higher values reflect higher national levels of gender inequality and greater disparities between women and men.*
National Level of Gender Segregation Between Arts and Science Graduates	1 = % difference of female representation is above -10. 2 = % difference of female representation is below -20. 3 = % difference of female representation is below -30. 4 = % difference of female representation is below -40.
	**Note: Gender segregation value was calculated by subtracting the % of females enrolled in humanities from the % of females enrolled in science. Data was taken from the World Bank data. (See World Bank Data http://data.worldbank.org/indicator to download data on gender statistics).*
Era	1 = Publication before 1980s 2 = Publication date during 1980s 3 = Publication date during 1990s 4 = Publication date during 2000s 5 = Publication date during or after 2010
Population Type	0 = Special population (e.g., students enrolled in elective or voluntary courses) 1 = Standard population (e.g., students in compulsory classes)
Publication Type	1 = Peer-reviewed journal article or book chapter 2 = Dissertation, thesis or unpublished conference paper
Reliability	0 = Cronbach's alpha was below .70 or not reported 1 = Cronbach's alpha was above .70

Appendix D

PISA Items Used

PISA self-concept items

The index of mathematics self-concept asks students to rate the extent to which they strongly agreed, agreed, disagreed, or strongly disagreed with the following statements:

I am just not good at mathematics. (R)

I get good grades in mathematics.

I learn mathematics quickly.

I have always believed that mathematics is one of my best subjects.

In my mathematics class, I understand even the most difficult work.

R = Reverse scored item

PISA self-efficacy items

The index of mathematics self-efficacy asks students to rate the extent to which they feel very confident, confident, not very confident, and not confident about doing the following tasks:

Using a train timetable to work out how long it would take to get from one place to another.

Calculating how much cheaper a TV would be after a 30% discount.

Calculating how many square meters of tiles you need to cover a floor.

Understanding graphs presented in newspapers.

Solving an equation like $3x + 5 = 17$.

Finding the actual distance between two places on a map with a 1:10,000 scale.

Solving an equation like $2(x + 3) = (x + 3)(x - 3)$.

Calculating the petrol consumption rate of a car.

PISA intrinsic motivation items

The index of intrinsic motivation to learn mathematics asks students to rate the extent to which they strongly agreed, agreed, disagreed, or strongly disagreed with the following statements:

I enjoy reading about mathematics

I look forward to my mathematics

I do mathematics because I enjoy it

I am interested in the things I learn in mathematics

PISA instrumental motivation items

The index of instrumental motivation to learn mathematics asks students to rate the extent to which they strongly agreed, agreed, disagreed, or strongly disagreed with the following statements:

Making an effort in mathematics is worth it because it will help me in the work that I want to do later on

Learning mathematics is worthwhile for me because it will improve my career prospects

Mathematics is an important subject for me because I need it for what I want to study later on

I will learn many things in mathematics that will help me get a job

PISA anxiety items

The index of mathematics anxiety asks students to rate the extent to which they strongly agreed, agreed, disagreed, or strongly disagreed with the following statements:

I often worry that it will be difficult for me in mathematics classes

I get very tense when I have to do mathematics homework

I get very nervous doing mathematics problems

I feel helpless when doing a mathematics problem

I worry that I will get poor grades in mathematics

All items are provided by the PISA 2003 Data Analysis Manual (OECD)

OECD (2005). *PISA 2003 Data Analysis Manual: SPSS users.*

SUPPLEMENTARY MATERIAL

Table S1

Meta-Analyses and Moderation Analyses for Math

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Math expectancy for success		150	-0.27	-0.31	-0.23	0.04		0.92	2046.31
Social class	**p = 0.00**						0.32		
Low SES		25	-0.19	-0.28	-0.10	0.02		0.65	56.87
Middle SES		19	-0.25	-0.33	-0.17	0.00		0.15	23.97
High SES		6	-0.77	-1.24	-0.31	0.31		0.94	85.09
Age	p = 0.45						0.05		
Elementary age		20	-0.31	-0.39	-0.24	0.01		0.53	45.80
Middle school age		48	-0.25	-0.34	-0.17	0.07		0.95	545.24
High school age		54	-0.32	-0.38	-0.26	0.03		0.93	1184.59
Young adult		19	-0.24	-0.36	-0.12	0.04		0.82	56.45
Era	p = 0.38						0.04		
< 1980s		14	-0.33	-0.44	-0.21	0.03		0.69	42.55
1980s		13	-0.19	-0.27	-0.11	0.01		0.56	31.25
1990s		45	-0.27	-0.33	-0.22	0.01		0.64	170.52
2000s		47	-0.31	-0.40	-0.22	0.08		0.97	1605.20
Current		30	-0.23	-0.30	-0.16	0.03		0.85	174.00

% Ethnic minority	p = 0.07						0.14		
Lowest		38	-0.33	-0.44	-0.22	0.09		0.94	228.72
Low		8	-0.24	-0.46	-0.02	0.09		0.98	924.04
High		4	0.05	-0.17	0.28	0.04		0.77	15.85
Highest		15	-0.17	-0.23	-0.12	0.00		0.00	19.68
National Gender Equality	p = 0.11						0.06		
Highest level of gender equality		19	-0.38	-0.45	-0.31	0.02		0.85	115.89
High level of gender equality		27	-0.28	-0.34	-0.23	0.01		0.73	89.25
Medium level of gender equality		91	-0.27	-0.32	-0.19	0.07		0.94	1477.19
Low level of gender equality		8	-0.16	-0.22	-0.10	0.00		0.58	25.26
Gender Segregation in Graduates	p = 0.41						0.05		
Medium levels of gender segregation		41	-0.22	-0.34	-0.11	0.12		0.97	1291.72
High levels of gender segregation		11	-0.32	-0.40	-0.23	0.01		0.83	59.42
Highest levels of gender segregation		12	-0.37	-0.47	-0.27	0.01		0.65	22.23
Sample Type	p = 0.50						0.00		
Special sample (e.g., elective STEM)		25	-0.24	-0.36	-0.13	0.07		0.87	146.23
Standard sample		125	-0.28	-0.32	-0.24	0.04		0.92	1891.85
Reliability	**p = 0.01**						0.08		
<.70 or not reported		55	-0.19	-0.24	-0.14	0.02		0.71	215.39
≥.70		93	-0.32	-0.37	-0.26	0.05		0.94	1598.16

Math task value		44	-0.14	-0.21	-0.06	0.05		0.93	567.35
Social class	**p = 0.00**						0.65		
Low SES		8	-0.03	-0.12	0.06	0.00		0.00	7.29
High SES		5	-0.67	-1.00	-0.34	0.11		0.85	36.13
Age	p = 0.61						0.02		
Middle school age		23	-0.17	-0.29	-0.04	0.08		0.96	508.48
High school age		11	-0.17	-0.27	-0.07	0.01		0.83	21.34
Young adult		5	-0.04	-0.20	0.11	0.01		0.59	9.69
Era	p = 0.30						0.10		
1980		4	-0.12	-0.16	-0.09	0.00		0.00	3.69
1990s		5	0.08	-0.04	0.19	0.00		0.00	6.62
2000s		21	-0.20	-0.35	-0.06	0.10		0.98	523.92
Current		14	-0.06	-0.12	0.00	0.00		0.02	12.74
% Ethnic minority	p = 0.26						0.08		
Lowest		12	-0.26	-0.48	-0.03	0.14		0.96	86.06
Highest		5	0.00	-0.19	0.19	0.01		0.46	7.32
National Gender Equality	p = 0.81						0.04		
Highest level of gender equality		5	-0.18	-0.36	0.00	0.03		0.91	26.37
High level of gender equality		9	-0.23	-0.24	-0.21	0.00		0.00	5.40
Medium level of gender equality		25	-0.16	-0.29	-0.03	0.09		0.95	299.92

Gender Segregation in Graduates	p = 0.79						0.00		
Medium levels of gender segregation		18	-0.22	-0.39	-0.05	0.11		0.96	173.90
High levels of gender segregation		5	-0.23	-0.26	-0.20	0.00		0.00	4.44
Sample Type	p = 0.92						0.00		
Special sample (e.g., elective STEM)		8	-0.15	-0.30	0.00	0.03		0.76	22.87
Standard sample		36	-0.14	-0.23	-0.05	0.05		0.95	537.11
Reliability	p = 0.11						0.09		
<.70 or not reported		16	-0.05	-0.15	0.05	0.03		0.81	195.27
≥.70		28	-0.19	-0.29	-0.08	0.06		0.95	259.68
Math intrinsic value		79	-0.17	-0.22	-0.12	0.04		0.91	1020.50
Social class	**p = 0.02**						0.38		
Low SES		6	-0.05	-0.11	0.01	0.00		0.00	13.17
Middle SES		17	-0.08	-0.19	0.03	0.02		0.54	26.75
High SES		5	-0.47	-0.75	-0.19	0.08		0.80	23.22
Age	p = 0.67						0.06		
Elementary age		18	-0.16	-0.24	-0.09	0.01		0.50	45.79
Middle school age		28	-0.15	-0.23	-0.07	0.04		0.92	279.34
High school age		20	-0.22	-0.32	-0.12	0.05		0.95	528.14
Young adult		10	-0.16	-0.32	0.01	0.05		0.82	42.45
Era	p = 0.58						0.03		
1990s		26	-0.15	-0.24	-0.06	0.03		0.79	180.48
2000s		36	-016	-0.22	-0.10	0.03		0.92	597.02

Current		13	-0.23	-0.38	-0.07	0.07		0.93	140.44
% Ethnic minority	p = 0.42						0.12		
Lowest		13	-0.23	-0.41	-0.06	0.08		0.92	71.79
Low		4	-0.15	-0.31	0.01	0.03		0.93	253.96
High		4	0.00	-0.12	0.11	0.00		0.00	4.22
Highest		5	-0.15	-0.36	0.06	0.04		0.72	16.87
National Gender Equality	**p = 0.03**						0.20		
Highest level of gender equality		13	-0.32	-0.45	-0.18	0.05		0.94	181.78
High level of gender equality		19	-0.17	-0.23	-0.11	0.01		0.64	51.93
Medium level of gender equality		39	-0.13	-0.21	-0.05	0.05		0.91	522.59
Low level of gender equality		5	-0.04	-0.11	0.02	0.00		0.46	14.23
Gender Segregation in Graduates	p = 0.21						0.22		
Medium levels of gender segregation		26	-0.12	-0.22	-0.02	0.05		0.92	375.41
High levels of gender segregation		9	-0.13	-0.20	-0.05	0.01		0.71	34.16
Highest levels of gender segregation		10	-0.29	-0.41	-0.17	0.01		0.73	43.00
Sample Type	p = 0.62						0.00		
Special sample (e.g., elective STEM)		11	-0.14	-0.28	0.00	0.04		0.80	48.64
Standard sample		68	-0.18	-0.23	-0.12	0.04		0.92	968.82
Reliability	p = 0.11						0.06		
<.70 or not reported		31	-0.11	-0.18	-0.04	0.02		0.74	167.98
≥.70		46	-0.21	-0.28	-0.14	0.05		0.93	763.28

Math utility value		60	-0.08	-0.13	-0.02	0.03		0.89	563.94
Social class	**p = 0.03**						1.00		
Low SES		14	0.05	-0.04	0.14	0.00		0.00	13.77
Middle SES		4	-0.10	-0.20	-0.01	0.00		0.00	2.28
Age	p = 0.43						0.04		
Middle school age		17	-0.10	-0.20	0.00	0.01		0.79	33.82
High school age		34	-0.10	-0.18	-0.03	0.04		0.93	497.10
Young adult		6	0.03	-0.14	0.20	0.02		0.70	13.60
Era	**p = 0.02**						0.29		
< 1980s		14	-0.18	-0.30	-0.06	0.04		0.72	49.09
1980s		8	-0.20	-0.37	-0.03	0.04		0.87	47.84
1990s		17	-0.09	-0.19	0.02	0.02		0.72	95.84
2000s		14	0.03	-0.05	0.10	0.01		0.86	174.28
Current		7	-0.02	-0.15	0.10	0.02		0.77	23.43
% Ethnic minority	**p = 0.03**						0.33		
Lowest		21	-0.13	-0.23	-0.03	0.04		0.85	67.24
Low		4	0.09	-0.07	0.25	0.02		0.91	128.71
Highest		9	0.16	0.07	0.24	0.00		0.00	12.74
National Gender Equality	p = 0.08						0.15		
High level of gender equality		7	-0.03	-0.12	0.05	0.01		0.70	28.03
Medium level of gender equality		44	-0.05	-0.12	0.02	0.03		0.87	299.51
Low level of gender equality		5	-0.25	-0.41	-0.08	0.03		0.90	32.08

Sample Type	p = 0.09						0.08		
Special sample (e.g., elective STEM)		10	0.04	-0.08	0.17	0.01		0.59	16.75
Standard sample		50	-0.10	-0.16	-0.04	0.03		0.91	540.85
Reliability	p = 0.45						0.01		
<.70 or not reported		17	-0.06	-0.12	0.00	0.00		0.00	21.91
≥.70		43	-0.09	-0.15	-0.02	0.04		0.90	542.00

Table S2

Moderation Analyses for Science

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Science expectancy for success		58	-0.18	-0.26	-0.10	0.04		0.60	794.65
Social Class	**p = 0.01**						0.64		
Low SES		7	-0.14	-0.27	-0.01	0.00		0.13	6.41
Middle SES		4	-0.56	-0.86	-0.27	0.07		0.83	26.44
Age	**p = 0.03**						0.17		
Elementary age		4	0.05	-0.03	0.14	0.00		0.00	2.51
Middle school age		22	-0.09	-0.18	0.00	0.04		0.92	455.49
High school age		25	-0.24	-0.33	-0.14	0.05		0.95	184.42
Young adult		5	-0.31	-0.51	-0.11	0.04		0.81	47.01
Era	p = 0.32						0.06		
1980s		4	-0.24	-0.34	-0.14	0.00		0.00	0.56
2000s		26	-0.19	-0.26	-0.12	0.03		0.92	461.49
Current		25	-0.10	-0.23	0.02	0.08		0.94	276.23
% Ethnic minority	p = 0.28						0.15		
Lowest		12	-0.26	-0.38	-0.14	0.03		0.80	53.38
Low		5	-0.18	-0.38	0.02	0.04		0.96	98.60
Highest		8	-0.09	-0.23	-0.06	0.03		0.62	23.60
National Gender Equality	**p = 0.00**						0.27		
High level of gender equality		17	-0.24	-0.36	-0.12	0.06		0.94	190.19

Medium level of gender equality		32	-0.15	-0.23	-0.08	0.03		0.86	190.22
Low level of gender equality		6	0.11	0.00	0.22	0.01		0.82	44.33
Gender Segregation in Graduates	**p = 0.01**						0.33		
Low levels of gender segregation		4	0.06	-0.15	0.27	0.04		0.91	73.45
Medium levels of gender segregation		28	-0.11	-0.20	-0.02	0.04		0.90	186.55
High levels of gender segregation		6	-0.30	-0.37	-0.23	0.01		0.72	66.48
Sample Type	p = 0.35						0.02		
Special sample (e.g., elective STEM)		9	-0.24	-0.39	-0.08	0.04		0.81	39.96
Standard sample		49	-0.15	-0.22	-0.08	0.05		0.94	739.07
Reliability	p = 0.61						0.01		
<.70 or not reported		16	-0.14	-0.27	0.00	0.07		0.91	288.32
≥.70		38	-0.18	-0.26	-0.10	0.05		0.93	336.38
Science task value		28	-0.01	-0.08	0.06	0.03		0.88	424.73
Age	p = 0.26						0.10		
Middle school age		16	0.03	-0.05	0.12	0.02		0.87	286.34
High school age		10	-0.05	-0.20	0.10	0.03		0.93	27.83
Era	p = 0.24						0.04		
2000s		18	-0.03	-0.10	0.04	0.02		0.89	341.30
Current		9	0.08	-0.12	0.27	0.07		0.93	79.41
% Ethnic minority	p = 0.09						0.60		
Lowest		6	-0.09	-0.27	0.08	0.02		0.77	14.57
Highest		4	0.13	-0.03	0.30	0.00		0.00	3.13

National Gender Equality	**p = 0.00**						0.72		
High level of gender equality		7	-0.14	-0.16	-0.12	0.00		0.00	13.51
Medium level of gender equality		14	0.02	-0.07	0.12	0.02		0.78	104.72
Low level of gender equality		6	0.16	0.13	0.19	0.00		0.00	9.84
Gender Segregation in Graduates	**p = 0.01**						0.52		
Medium levels of gender segregation		15	0.05	-0.04	0.14	0.02		0.81	100.20
High levels of gender segregation		4	-0.15	-0.18	-0.12	0.00		0.10	4.96
Reliability	p = 0.56						0.06		
<.70 or not reported		9	0.03	-0.10	0.16	0.04		0.84	150.44
≥.70		17	-0.02	-0.11	0.08	0.02		0.86	83.48
Science intrinsic value		39	-0.21	-0.32	-0.11	0.04		0.62	575.22
Age	p = 0.97						0.00		
Elementary age		4	-0.22	-0.57	0.13	0.11		0.93	31.35
Middle school age		14	-0.16	-0.25	-0.07	0.03		0.88	297.93
High school age		15	-0.14	-0.29	0.00	0.08		0.97	216.47
Young adult		5	-0.13	-0.27	0.01	0.02		0.62	16.63
Era	p = 0.15						0.07		
2000s		20	-0.18	-0.26	-0.10	0.03		0.93	418.43
Current		16	-0.08	-0.22	0.07	0.08		0.94	142.11
% Ethnic minority	p = 0.85						0.00		
Lowest		7	-0.25	-0.37	-0.13	0.01		0.46	11.62
Low		5	-0.33	-0.58	-0.08	0.07		0.98	52.47

National Gender Equality	p = 0.10						0.08		
High level of gender equality		16	-0.10	-0.23	0.03	0.07		0.95	230.86
Medium level of gender equality		20	-0.22	-0.31	0.13	0.02		0.84	113.49
Gender Segregation in Graduates	p = 0.56						0.00		
Medium levels of gender segregation		17	-0.22	-0.32	-0.12	0.03		0.89	158.53
High levels of gender segregation		6	-0.16	-0.23	-0.09	0.01		0.71	65.88
Sample Type	p = 0.98						0.00		
Special sample (e.g., elective STEM)		4	-0.15	-0.34	0.04	0.03		0.72	12.55
Standard sample		35	-0.15	-0.23	-0.07	0.05		0.94	561.26
Reliability	p = 0.51						0.01		
<.70 or not reported		12	-0.19	-0.34	-0.05	0.05		0.89	215.47
≥.70		23	-0.14	-0.24	-0.03	0.06		0.94	232.47
Science utility value		16	-0.05	-0.12	0.02	0.01		0.80	141.87
Era	p = 0.64						0.02		
2000s		6	-0.05	-0.15	0.04	0.01		0.85	109.17
Current		7	-0.02	-0.15	0.11	0.02		0.80	19.62
National Gender Equality	p = 0.71						0.00		
High level of gender equality		6	-0.04	-0.12	0.04	0.01		0.63	48.48
Medium level of gender equality		8	-0.02	-0.14	0.11	0.02		0.79	19.87
Sample Type	p = 0.29						0.11		
Special sample (e.g., elective STEM)		5	0.01	-0.15	0.18	0.02		0.68	12.38
Standard sample		11	-0.07	-0.15	0.00	0.01		0.79	128.90

Table S3

Moderation Analyses for Computing and Physical Sciences

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	τ^2	R^2	I^2	*Q Statistic*
Computing expectancy for success		22	-0.44	-0.60	-0.28	0.13		0.97	198.76
Age	p = 0.19						0.16		
High school age		7	-0.33	-0.60	-0.05	0.11		0.98	76.02
Young adult		9	-0.62	-0.90	-0.33	0.15		0.94	76.36
Era	p = 0.92						0.03		
1990s		4	-0.36	-0.76	0.04	0.10		0.93	15.82
2000s		8	-0.48	-0.80	-0.16	0.19		0.99	135.78
Current		8	-0.40	-0.59	-0.20	0.05		0.91	32.72
National Gender Equality	p = 0.38						0.10		
High level of gender equality		10	-0.33	-0.55	-0.11	0.10		0.96	87.90
Medium level of gender equality		9	-0.43	-0.50	-0.35	0.00		0.00	16.54
Sample Type	**p = 0.02**						0.27		
Special sample (e.g., elective STEM)		13	-0.61	-0.83	-0.39	0.12		0.92	82.58
Standard sample		9	-0.23	-0.41	-0.05	0.06		0.95	84.50

Reliability	p = 0.47						0.05		
<.70 or not reported		9	-0.51	-0.75	-0.26	0.13		0.95	100.12
≥.70		13	-0.38	-0.59	-0.17	0.11		0.97	89.96
Physical sciences expectancy for success		16	-0.43	-0.56	-0.29	0.05		0.93	60.00
Age	p = 0.84						0.37		
High school age		5	-0.48	-0.59	-0.37	0.00		0.18	4.21
Young adult		7	-0.43	-0.67	-0.19	0.07		0.87	20.49
Era	p = 0.09						0.25		
2000s		5	-0.23	-0.58	0.12	0.13		0.98	30.57
Current		7	-0.51	-0.62	-0.41	0.00		0.39	6.85
National Gender Equality	p = 0.10						0.29		
Highest level of gender equality		5	-0.56	-0.64	-0.48	0.00		0.00	2.29
Medium level of gender equality		8	-0.35	-0.58	-0.12	0.08		0.94	34.33
Sample Type							0.00		
Special sample (e.g., elective STEM)	p = 0.93	9	-0.42	-0.62	-0.22	0.05		0.83	23.13
Standard		7	-0.43	-0.62	-0.24	0.05		0.94	36.73
Reliability	p = 0.17						0.11		
< 0.70 or not reported		8	-0.51	-0.67	-0.35	0.03		0.84	25.51
≥.70		8	-0.32	-0.55	-0.09	0.07		0.95	31.44

Table S4

Moderation Analyses for Verbal

Variable	*ANOVA p-value*	*ESs*	*d*	*Lower 95% CI*	*Upper 95% CI*	t^2	R^2	I^2	*Q Statistic*
Verbal expectancy for success		65	0.17	0.11	0.23	0.04		0.92	353.17
Social class	p = 0.23						0.17		
Low SES		10	0.23	0.12	0.34	0.02		0.64	35.34
Middle SES		18	0.13	0.03	0.24	0.02		0.59	38.04
Age	p = 0.69						0.03		
Elementary age		22	0.22	0.12	0.32	0.02		0.77	43.77
Middle school age		16	0.14	0.02	0.26	0.06		0.94	150.07
High school age		10	0.23	0.07	0.38	0.06		0.96	98.16
Young adult		13	0.21	0.14	0.28	0.00		0.10	17.76
Era	p = 0.27						0.08		
1990s		25	0.20	0.13	0.28	0.02		0.66	52.43
2000s		21	0.22	0.12	0.33	0.04		0.95	77.06
Current		15	0.12	0.02	0.21	0.03		0.85	109.87

Ethnic Minority % of Sample	p = 0.85						0.01		
Lowest		15	0.14	0.04	0.24	0.03		0.84	129.52
Highest		7	0.13	-0.10	0.37	0.06		0.80	19.89
National Gender Equality Level	p = 0.98						0.00		
Highest gender equality		9	0.17	0.03	0.31	0.04		0.92	60.56
High gender equality		19	0.16	0.04	0.28	0.05		0.93	86.15
Medium gender equality		37	0.18	0.10	0.26	0.05		0.91	192.45
Gender Segregation in Graduates	p = 0.60`						0.00		
Medium levels of gender segregation		20	0.20	0.10	0.30	0.03		0.87	64.74
High levels of gender segregation		6	0.20	0.12	0.27	0.00		0.00	6.75
Highest levels of gender segregation		11	0.14	0.05	0.23	0.00		0.04	11.80
Sample Type	p = 0.36						0.02		
Special sample (e.g., elective STEM)		11	0.23	0.12	0.35	0.02		0.70	30.83
Standard sample		54	0.16	0.09	0.23	0.05		0.94	316.17
Reliability	p = 0.74						0.00		
<.70 or not reported		34	0.16	0.06	0.26	0.06		0.90	184.88
≥.70		31	0.18	0.10	0.25	0.03		0.90	167.08

Verbal task value		22	0.48	0.34	0.62	0.10		0.96	315.92
Social class	p = 0.67						0.03		
Low SES		4	0.36	0.07	0.65	0.08		0.87	57.91
Middle SES		4	0.50	-0.08	1.07	0.33		0.96	164.71
Age	p = 0.68						0.00		
Middle school age		9	0.47	0.21	0.73	0.15		0.98	225.91
Young adult		7	0.48	0.40	0.56	0.00		0.00	7.29
Era (not any data for prior 1980s)	p = 0.59						0.03		
2000s		12	0.46	0.40	0.52	0.00		0.00	24.50
Current		5	0.41	0.16	0.67	0.08		0.94	90.20
% Ethnic minority	p = 0.96						0.00		
Lowest		8	0.48	0.18	0.78	0.18		0.97	277.32
Highest		4	0.48	0.08	0.89	0.12		0.89	16.26
Sample Type	p = 1.00						0.00		
Special sample (e.g., elective STEM)		9	0.47	0.34	0.60	0.02		0.66	20.60
Standard sample		13	0.48	0.28	0.69	0.13		0.98	290.93
Reliability	p = 0.18						0.08		
<.70 or not reported		4	0.27	0.06	0.47	0.03		0.80	10.13

≥.70		18	0.53	0.37	0.68	0.10		0.96	298.37
Verbal intrinsic value		35	0.32	0.24	0.40	0.04		0.90	196.98
Social Class	p = 0.49						0.11		
Low SES		4	0.41	0.24	0.58	0.02		0.64	15.26
Middle SES		15	0.33	0.19	0.47	0.03		0.69	38.69
Age	p = 0.14						0.19		
Elementary age		17	0.42	0.31	0.54	0.02		0.75	33.51
Middle school age		6	0.30	0.17	0.43	0.02		0.86	42.72
High school age		7	0.24	0.06	0.41	0.05		0.95	80.01
Young adult		4	0.23	0.13	0.32	0.00		0.00	5.59
Era (not any data for prior 1980s)	**p = 0.02**						0.33		
1990s		18	0.40	0.30	0.51	0.01		0.65	34.99
2000s		7	0.43	0.36	0.51	0.00		0.00	6.66
Current		10	0.21	0.08	0.34	0.04		0.88	76.50
National Gender Equality Level	p = 0.47						0.08		
Highest gender equality		6	0.25	0.09	0.41	0.03		0.89	35.26
High gender equality		15	0.38	0.24	0.52	0.04		0.92	56.76
Medium gender equality		14	0.30	0.19	0.41	0.03		0.84	66.20
Gender Segregation in Graduates	**p = 0.04**						0.65		

Medium levels of gender segregation		7	0.25	0.10	0.40	0.01		0.78	14.68
High levels of gender segregation		6	0.52	0.37	0.67	0.01		0.80	9.42
Highest levels of gender segregation		9	0.36	0.25	0.46	0.00		0.00	8.80
Reliability	**p = 0.01**						0.37		
<.70 or not reported		25	0.39	0.31	0.47	0.02		0.74	65.84
≥.70		10	0.19	0.07	0.31	0.03		0.88	47.67

Table S5

Summary of Studies Included in the Meta-Analysis

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Abarbanel, T. (2008)	Canada	601	311/290	13.5	Middle class	--	Math
Adeyemi, A. A., & Adeniyi, O. M. (2010)	Nigeria	150	69/81	17	--	--	Math
Adkinson, J. E. (2008)	United States	105	53/5	11	--	53% White; 21% African American; 18% Hispanic; 8% other	Math
Alliman-Brissett, A. E. (2007)	United States	108	55/53	13	Lower-middle income	90% African American	Math
Anderman, E. (1992)	United States	678	--	12	Blue collar district	--	Verbal (literacy)
Asonye, E. I. B. (2003)	Nigeria	102	46/56	--	--	--	Math
Baker, L., & Wigfield, A. (1999)	United States	371	178/192	11	54% of sample receiving free/reduced price lunch	52% White; 46% African American; 2% other	Verbal (reading)
Battle, E. (1966)	United States	500	--	13.5	--	--	Math
Beyer, S. (2014)	United States	1319	447/872	20.5	--	83% White; 7% Hispanic; 5% African American; 4% Asian American; >1% Native American	Computing

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Bhanot, R. T. & Jovanovic, J. (2009)	United States	114	54/60	12.21	--	--	Science
Black, J. (2008) – 1	United States	114	53/61	8.5	Substantial % of sample receiving free lunch	Approx. 82% of total sample from ethnic minority	Verbal
Black, J. (2008) – 2	United States	134	72/62	10.5	Substantial % of sample receiving free lunch	Approx. 82% of total sample from ethnic minority	Verbal
Boe, M. V., & Henriksen, E. K. (2013) - 1	Norway	558	312/246	18.5	--	--	Science (Physics)
Boe, M. V., & Henriksen, E. K. (2013) - 2	Norway	278	89/189	19.5	--	--	Science (Physics)
Bonitz, V. S., et al. (2010)	United States	45	15/30	23.4	--	--	Computing
Bonnot, V. and J. C. Croizet (2007)	France	174	100/74	18.86	--	--	Math and Verbal
Bouffard, T., et al. (2003)	Canada	115	63/52	7.04	Sample from middle class area	PPAGS-C	Math and Verbal (Reading)
Branom, C. M. (2014)	United States	23,035	11,718/11,317	14.5	--	53.2% White	Math

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Britner, S. L. (2002)	United States	268	107/161	12.5	--	100% African American	Science
Brockman, G. (2007)	United States	114	173/79	22.75	--	Approx. 98% of total sample from ethnic minority	Math
Broome, P. (2001)	Germany	595	327/268	14.8	--	--	Science (Physics)
Campbell, N. K. and G. Hackett (1986)	United States	120	60/60	22.75	--	--	Math
Campos-Sanchez, A., et al. (2014) - 1	Spain	132	43/89	18.0	--	--	Science (Histology)
Campos-Sanchez, A., et al. (2014) – 2	Spain	125	44/81	18.0	--	--	Science (Histology)
Campos-Sanchez, A., et al. (2014) – 3	Spain	110	31/79	18.0	--	--	Science (Histology)
Cavas, P. (2011)	Turkey	376	188/188	12.5	--	--	Science
Chan, J. Y. (2015)	United States	158	47/111	22.75	--	53.2% White	Math and Science (Chemistry)

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Cheong, Y. F., et al. (2004)	United States	314	250/64	--	--	Approx. 54.78% of total sample from ethnic minority	Computing
Chiu, M. S. (2011)	Taiwan	247	124/123	14	--		Science
Chouinard, R., & Roy, N. (2008)	Canada	1,130	452/678	14.85	--	--	Math
Chouinard, R., et al. (1999) -1	Canada	800	388/412	13.71	--	--	Math
Chouinard, R., et al. (1999) -2	Canada	1,085	515/570	16.21	--	--	Math
Chow, S. J. et al. (2013)	Brunei	324	141/183	16.44	--	--	Science
Coffin, R. J., & P. D. MacIntyre (1999)	Canada	111	31/79	21.3	--	88% White; 4% Asian; 2% Aboriginal Canadian	Computing
Copping, K. E. (2012)	United States	318	136/182	15.8	--	Approx. 61% of sample from ethnic minority	Math, Science and Verbal
Crane, L. R., et al. (2000) - 1	United States	170	84/86	--	--	100% Asian American	Reading

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Crane, L. R., et al. (2000) - 2	United States	55	34/21	--	--	100% African American	Reading
Crane, L. R., et al. (2000) - 3	United States	118	45/73	--	--	100% Hispanic	Reading
Crane, L. R., et al. (2000) - 4	United States	1,281	647/635	--	--	100% White	Reading
Cribbs, J. D. (2013)	United States	10,492	6,925/4,197	22.75	--	66.7% White	Math
Crombie, G., et al. (2002)	Canada	187	155/32	16.3	--	--	Computing
Crombie, G. & Armstrong (1999)	Canada	67	38/8	15.5	--	--	Computing
Cupani & Pautassi (2013)	Argentina	543	304/239	13.23	Middle class	--	Math
Deacon, M. M. (2012)	United States	231	95/136	13.16	> 50% of sample eligible for free or reduced lunch	42.9% White	Math
DeBacker, T. K. and R. Nelson (1999)	United States	157	69/80	15	--	--	Science

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Deemer, E. D., et al. (2014)	United States	55	34/21	--	--	100% African American	Reading
Crane, L. R., et al. (2000) - 3	United States	118	45/73	--	--	100% Hispanic	Reading
Crane, L. R., et al. (2000) - 4	United States	1,281	647/635	--	--	100% White	Reading
Cribbs, J. D. (2013)	United States	10,492	6,925/4,197	22.75	--	66.7% White	Math
Crombie, G., et al. (2002)	Canada	187	155/32	16.3	--	--	Computing
Crombie, G. & Armstrong (1999)	Canada	67	38/8	15.5	--	--	Computing
Cupani & Pautassi (2013)	Argentina	543	304/239	13.23	Middle class	--	Math
Deacon, M. M. (2012)	United States	231	95/136	13.16	> 50% of sample eligible for free or reduced lunch	42.9% White	Math
DeBacker, T. K. and R. Nelson (1999)	United States	157	69/80	15	--	--	Science

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Deemer, E. D., et al. (2014)	United States	1,368	635/733	20.7	--	61.6% White	Science
Denner, J., et al. (2014)	United States	741	550/191	24.85	--	Black 5%; Asian/Pacific Islander 32%; Hispanic/Latino 20%; White 48%	Computing
Desy et al. (2009)	United States	376	199/169	--	--	75.5% White; Asian 12.2%; Black 7.4%	Science
DeWitt, J., et al. (2011)	England	298	136/162	22.75	2 out of 4 schools deprived; 1 from high socio-economic area	37 Black students; 83 White students; 151 Asian (south asian) students; 26 mixed race and Chinese students	Science
Dickhauser, O., & J. Stiensmeier-Pelster (2002)	Germany	200	100/100	16.3	--	--	Science
Dlamini, M. S. (1998)	Swaziland	941	438/478	15	--	--	Math
Doube, W., & Lang, C. (2012)	Australia	85	47/38	22.75	--	42% Australian domestic students and 58% international students	Computing
Dogbey, G. (2011)	United States	288	78/210	27.06	> 50% of sample eligible for free or reduced lunch	59.8% of students had annual household incomes less than $40,000;	Math

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Eccles, J., et al. (1993)	United States	865	420/444	7.83	Middle class	> 95% White	Math and Reading
Elisha-Primo, I., et al (2010)	Israel	469	130/339	--	--	83.2% Hebrew speakers; 28 Russian speakers; 21 Arabic speakers; 20 Amharic, Spanish and French speakers	Verbal (English)
Elmore, P. B., & Vasu, E. S (1980)	United States	188	83/98	--	--	--	Math
Else-Quest, N. M., et al. (2013) - 1	United States	102	54/48	16.19	Economically disadvantaged	100% White	Math and Science
Else-Quest, N. M., et al. (2013) - 2	United States	96	55/41	16.19	Economically disadvantaged	100% African American	Math and Science
Else-Quest, N. M., et al. (2013) - 3	United States	84	39/45	16.19	Economically disadvantaged	100% Latino/Latina	Math and Science
Else-Quest, N. M., et al. (2013) - 4	United States	85	35/50	16.19	Economically disadvantaged	100% Asian American	Math and Science
Ethington, C. A. (1991)	United States	869	401/460	13.5	--	--	Math
Falco, L. D. (2008)	United States	153	78/75	11.5	High SES	87% White	Math

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Falco, L. D., et al (2008) - 1	United States	112	57/55	11.5	High SES	93% White; 4% Hispanic; <1% Asian American; <1% Eastern Indian	Math
Falco, L. D., et al (2008) - 2	United States	116	57/59	11.5	High SES	93% White; 4% Hispanic; <1% Asian American; <1% Eastern Indian	Math
Feather, N. (1988)	Australia	444	183/260	22.79	--	--	Math and Verbal (English)
Fennema, E., & Sherman, J (1977) - 1	United States	528	271/257	14.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 2	United States	472	224/248	14.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 3	United States	350	169/181	14.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 4	United States	520	243/277	14.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 5	United States	396	217/179	15.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 6	United States	330	158/172	15.5	Working class	Nearly all students described as white	Math

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Fennema, E., & Sherman, J (1977) - 7	United States	347	184163	15.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 8	United States	229	109/120	15.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 9	United States	264	149/115	16.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 10	United States	185	112/73	16.5	Working class	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 11	United States	160	87/73	16.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 12	United States	130	68/62	16.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 13	United States	97	64/33	17.5	--	Nearly all students described as white	Math
Fennema, E., & Sherman, J (1977) - 14	United States	73	64/33	17.5	--	Nearly all students described as white	Math
Ferla, J., et al. (2009)	Belgium	8,796	--	15.5	--	--	Math

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Frenzel, A. C., et al. (2007)	Germany	2,053	1036/1017	11.7	--	--	Math
Fryer, L. K. (2015)	Japan	381	246/135	--	--	--	Verbal (English)
Fullerton, J. A., & Umphrey, D. (2001)	United States	275	99/176	20.5	--	85% White	Math (Statistics)
Ganley C. M. & Lubienski, S. T. (2016)	United States	7,040	3,460/3,580	8.5	--	--	Math
Gasco, J., et al. (2014) - 1	Spain	192	88/104	--	--	--	Math
Gasco, J., et al. (2014) - 2	Spain	211	95/116	--	--	--	Math
Githua, B. N., & Mwangi, J. G. (2003)	Kenya	649	320/329	--	--	--	Math
Glynn, S. M., et al. (2011) - 1	United States	367	127/240	22.75	--	7% African American; 3.1% Hispanic; 0.6% Multiracial; 0.2% Native American	Science
Glynn, S. M., et al. (2011) - 2	United States	313	98/215	22.75	--	Same as above (taken from total sample)	Science

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Greene, B. A., et al. (1999) - 1	United States	191	83/108	16.5	Sample from middle class area	81% Caucasian; 8% Native American; 5% Hispanic; 4% African American; 2% Asian	Math
Greene, B. A., et al. (1999) - 2	United States	167	63/104	16.5	Sample from middle class area	81% Caucasian; 8% Native American; 5% Hispanic; 4% African American; 2% Asian	Math
Guo, J., et al. (2015) -1	Hong Kong (‘99)	5,179	2,626/2,553	14.4	--	85% White	Math
Guo, J., et al. (2015) -2	Hong Kong (‘03)	4,972	2,466/2,506	14.4	--	--	Math
Guo, J., et al. (2015) - 3	Hong Kong (‘07)	3,470	1,721/1,749	14.4	--	--	Math
Guthrie, J. T., et al. (2013)	United States	1,159	615/544	12.5	24% qualified for free or reduced price school lunches	78% European American; 19% African American; 3% Asian	Verbal (Reading)
Guvercin, O., et al. (2010) -1	Turkey	1,114	593/551	11.5	--	--	Science
Guvercin, O., et al. (2010) -2	Turkey	1,035	510/525	13.5	--	--	Science

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Hackett, G. (1992) - 1	United States	125	99/26	19.7		100% European American	Engineering
Hackett, G. (1992) - 2	United States	42	32/10	19.7		100% Mexican American	Engineering
Hackett, G. (1992) - 3	United States	9	5/4	19.7	--	100% African American	Engineering
Hackett, G. (1992) - 4	United States	21	13/8	19.7	--	100% Asian American	Engineering
Hackett, G., et al. (1990) - 1	United States	38	18/20	20.55	--	--	Math
Hackett, G., et al. (1990) - 2	United States	37	18/19	20.55	--	--	Math
Hackett, G., et al. (1990) - 3	United States	36	17/19	20.55	--	--	Math
Hackett, G., et al. (1990) - 4	United States	38	18/20	20.55	--	--	Math

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Harty, H., et al. (1986)	United States	228	110/118	11.5	--	100% White	Science
Hassan, G. (2008)	Australia	1,568	818/750	18.35	Middle class	--	Science
Henderson, B. B., et al. (1999) - 1	United States	104	51/53	7.5	Middle class		Math and Verbal
Henderson, B. B., et al. (1999) - 2	United States	75	37/38	8.5	Middle class		Math and Verbal
Henderson, B. B., et al. (1999) - 3	United States	129	56/73	9.5	Middle class	--	Math and Verbal
Henderson, B. B., et al. (1999) - 4	United States	100	54/46	10.5	Middle class	--	Math and Verbal
Henderson, B. B., et al. (1999) - 5	South Korea	80	42/38	7.5	Middle class	--	Math and Verbal
Henderson, B. B., et al. (1999) - 6	South Korea	84	46/38	8.5	Middle class	--	Math and Verbal
Henderson, B. B., et al. (1999) - 7	South Korea	89	50/39	9.5	Middle class	--	Math and Verbal

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Henderson, B. B., et al. (1999) - 8	South Korea	83	48/35	10.5	Middle class	--	Math and Verbal
Henderson, B. B., et al. (1999) - 9	Japan	52	30/22	7.5	Middle class	--	Math and Verbal
Henderson, B. B., et al. (1999) - 10	Japan	65	31/34	8.5	Middle class	--	Math and Verbal
Henderson, B. B., et al. (1999) - 11	Japan	56	29/27	9.5	Middle class	--	Math and Verbal
Henderson, B. B., et al. (1999) - 12	Japan	51	27/24	10.5	Middle class	--	Math and Verbal
Hong, Z.-R., & Lin, H. S. (2011) - 1	Taiwan	922	444/432	11	--	--	Science
Hong, Z.-R., & Lin, H. S. (2011) - 2	Taiwan	499	253/232	14	--	--	Science
Hong, Z.-R., & Lin, H. S. (2011) - 3	Taiwan	577	144/425	17	--	--	Science

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Hong, Z.-R., & Lin, H. S. (2011) - 4	Taiwan	878	350/516	17	--	--	Science
Inda, M., et al. (2013)	Spain	579	416/163	20	--	100% Spanish	Engineering
Jovanovic, J., & King, S. S. (1998)	United States	165	78/87	12.21	--	76% Euro-American; 5% African American; 4% Latino; 3% Asian American; 12% other	Science
Kahle, J. B., & Damnjanovic, A. (1994) - 1	United States	348	202/146	10	Lower middle class	100% White	Physical and Biological Science
Kahle, J. B., & Damnjanovic, A. (1994) - 2	United States	321	152/169	10	Lower middle class	100% African American	Physical and Biological Science
Keller, C. (2001)	Switzerland	6,602	3,006/3,021	14.5	--	--	Math
Kelley, M. J., & Decker, E. O.	United States	1,080	555/525	12.18	--	74% White; 10% Hispanic; 7% Black; 5% Multiracial; 4% Asian	Verbal (Reading)
Khoury, G. A., & Voss, B. E. (1985)	United States	237	110/127	15.5	--	16% African American, Middle Eastern, Asian American and Native American	Science

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Kilic-Bebek, E. (2010)	United States	162	68/94	22.75	--	69% White	Math
Kim, M. S., & Seo, Y. S. (2014)	South Korea	660	519/141	21.35	--		Engineering
Kissau, S. (2006)	Canada	490	236/254	14	--	--	Verbal (French)
Koh, C. K. (2012)	Singapore	1,733	388/412	--	--	92% Chinese	Math, Verbal (English), Science
Koohang, A. A. (1989)	United States	81	34/47	22.75	--	--	Computing
Kyttala, M., & Bjorn, P. M. (2010)	Finland	116	64/52	13.5	--	--	Math
Lau, S. (2003)	United States	318	158/160	16	--	49% European-American, remainder of sample from ethnic minority	Science
Lee, W., et al. (2014)	South Korea	500	246/253	14.5	--	Approx. 61% of sample from ethnic minority	Math, Science and Verbal (Korean and English)
Lefevre, J. A., et al.	Canada	126	59/67	21.7	--	--	Math

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Leibham, M. B., et al. (2013)	United States	116	66/50	8.4	Middle/middle-upper class	4.31% ethnic minority	Science
Lent, R. W., et al. (2005)	United States	487	352/120	18.82	--	Approx. 87% identified as Black	Engineering
Lent, R. W., et al. (1993)	United States	166	59/107	19.58	--	85% White	Math
Levin, J., & Fowler, H. (1984)	United States	988	489/499	17.05	Lower-middle class community	Mainly white community	Science
Levin, J., & Klindienst, D. (1983)	United States	163	89/74	16.96	--	--	Science
Lindsay, H. A. (2002) - 1	United States	93	45/48	22.75	--	--	Science (Chemistry)
Lindsay, H. A. (2002) - 2	United States	100	41/59	22.75	--	--	Science (Chemistry
Liou, P. Y., & Kuo, P. J.	Taiwan	1,822	1,020/802	16	--	--	Computing (Technology)
Lupart, J. L., et al. (2004)	Canada	1,419	657/762	14	--	--	Math, Science, and Verbal (English)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Makrakis, V. (1992) - 1	Sweden	303	159/144	14.5	--	--	Computing
Makrakis, V. (1992) - 2	Japan	470	204/266	14.5	--	--	Computing
Marinak, B. A., & Gambrell, L. B. (2010)	United States	288	143/145	8	--	50% White; 30% African American; 20% Asian	Verbal (reading)
Marsh, H. W., et al. (2013) - 1	Saudi Arabia	4,269	2,006/2,263	14	--	Mainly white community	Math and Science
Marsh, H. W., et al. (2013) - 2	Jordan	5,251	2,468/2,783	14	--	--	Math and Science
Marsh, H. W., et al. (2013) - 3	Oman	4,752	2,519/2,233	14	--	--	Math and Science
Marsh, H. W., et al. (2013) - 4	Egypt	6,582	3,357/3,225	14	--	--	Math and Science
Marsh, H. W., et al. (2013) - 5	United States	7,593	3,797/3,796	14	--	--	Math and Science
Marsh, H. W., et al. (2013) - 6	England	4,048	1,943/2,105	14	--	--	Math and Science

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Marsh, H. W., et al. (2013) - 7	Australia	4,103	2,257/1,846	14	--	--	Math and Science
Marsh, H. W., et al. (2013) - 8	Scotland	4,205	2,060/2,145	14	--	--	Math and Science
Marsh, H. W., et al. (2005)	Germany	2,264	1,132/1,132	13.7	--	95% White	Math
Marsh, H. W., et al. (1998)	United States	5,847	2,885/2,962	14.5	--	--	Math and Verbal (English)
Matthews, G. (2003)	United States	57	31/26	20.55	--	--	--
Medeiros, D. J. (2012)	United States	336	171/165	16	--	83% Hispanic, Asian, Black, American Indian, Native Hawaiian	Math and Science
Meece, J. L. (1981)	United States	254	122/132	13.5	Middle to upper middle class	--	Math
Meelissen, M., & Luyten, H. (2008)	The Netherlands	2,908	1,476/1,432	9.5	--	--	Math
Miller, M. (2010)	United States	84	20/64	>25	--	--	Computing

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Mills, N., et al. (2007)	United States	303	89/214	22.75	--	--	Verbal (French)
Miura, I. T. (1987)	United States	368	104/264	22.75	--	--	Computing
Nagy, G., et al. (2006)	Germany	1,148	459/689	15.5	--	--	Math and Biology
Negishi, M. (2007)	Japan	616	459/689	16	--	--	Science
Negishi, M. (2007)	United States	108	413/203	17	--	--	Science
Oberman, P. S. (2002)	United States	314	250/64	16.5	--	50.34% African American	Computing
Meece, J. L. (1981)	United States	254	122/132	13.5	Middle to upper middle class	--	Math
Meelissen, M., & Luyten, H. (2008)	The Netherlands	2,908	1,476/1,432	9.5	--	--	Math
Miller, M. (2010)	United States	84	20/64	>25	--	--	Computing

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Mills, N., et al. (2007)	United States	303	89/214	22.75	--	--	Verbal (French)
Miura, I. T. (1987)	United States	368	104/264	22.75	--	--	Computing
Nagy, G., et al. (2006)	Germany	1,148	459/689	15.5	--	--	Math and Biology
Negishi, M. (2007)	Japan	616	459/689	16	--	--	Science
Negishi, M. (2007)	United States	108	413/203	17	--	--	Science
Oberman, P. S. (2002)	United States	314	250/64	16.5	--	50.34% African American	Computing
Obrentz, S. B. (2012)	United States	413	145/248	22.75	--	44.8% White; 38.5% Asian/Pacific Islander; 10.2% Black; 4.1% Hispanic; 3.1% Other	Science (Chemistry)
Pajares, F., & Miller, M. (1994)	United States	350	121/229	22.75	--	--	Math
Pajares, F., & Valiante (1999)	United States	742	366/376	12.5	Middle class	Primarily white sample	Verbal (Writing)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Pajares, F., Valiante, G. (2001)	United States	497	247/250	12.50	Middle class	Primarily white	Verbal (Writing)
Pajares, F., & Viliante, G. (1997)	United States	218	103/115	10.50	--	--	Verbal (Writing)
Paslov, L. S. (2007)	United States	156	79/77	13	--	43% of sample from ethnic minority	Math
Patrick, H., et al. (2009)	United States	49	--	5.5	Most students received reduced/free lunch	12.5% Hispanic; 16.7% African American and 6.3% Other	Science
Peklaj, C., et al. (2014)	Slovenia	386	149/244	15.67	--	--	Math
Pell, A. (1985)	United Kingdom	274	108/135	10.5	--	--	Science (Physics)
Pinxten, M. et al. (2014)	Belgium	4,724	2,339/2,385	9.5	--	--	Math
Plante, I., et al. (2013)	Canada	770	385/385	12	Sample from low socioeconomic area	Canadians of French Caucasian ancestry.	Math and Verbal (Language Arts)
Preckel, F., et al. (2008)	Germany	362	181/181	12.77	--	--	Math

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Quihuis, G. (2002)	United States	90	45/45	16	90% of parents college graduates. >50% parents completed masters, PhD, legal or medical degree.	77% White, 23% non-white students	Science
Quinn, F., & Lyons, T. (2011)	Australia	3,800	1,900/1,900	15	--	--	Science
Rech, J. F. (1994)	United States	251	124/126	11.5	--	100% African American	Math
Risconscente, M. M. (2014)	United States	326	140/186	15	Most students received reduced/free lunch	100% Latino/a	Math
Riegle-Crumb, C., et al. (2011) - 1	United States	5,698	2,472/2,956	13.5	--	100% White	Math
Riegle-Crumb, C., et al. (2011) - 2	United States	2,598	1,232/1,366	13.5	--	100% Black and Latino/a	Math
Sasson, I., & Cohen, D. (2013)	Israel	66	20/30	14.5	--	--	Science (Physics)
Selkirk, L. C., et al. (2011)	United States	1,953	937/1,016	11.5	Lower-middle class	95% White	Math and Verbal (English)

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Senler, B., & Sungur, S. (2009)	Turkey	502	249/253	11.5	--	--	Science
Sha, S. L. (2012)	United States	440	268/172	15	--	1.6% White and 5% White in each school	Science (Physics)
Shashaani, L., Khalili, A. (2001)	Iran	375	155/220	20	--	--	Computing
Sheldrake, R., et al. (2014)	England	2,490	140/186	15	--	--	Math
Sherman (1981) – 1	United States	116	36/80	17.5	--	--	Math
Sherman (1981) – 2	United States	120	52/68	17.5	--	--	Math
Sherman (1981) – 3	United States	95	64/31	17.5	--	--	Math
Siegle, D., & Reis, S. M. (1998)	United States	5,385	2,709/2,676	11.5	--	--	Math, Science and Verbal (Language Arts)
Skaalvik, S., & Skaalvik, E. M. (2004) - 1	Norway	277	129/148	10.9	--	--	Math and Verbal

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Skaalvik, S., & Skaalvik, E. M. (2004) - 2	Norway	239	115/124	13.95	--	--	Math and Verbal
Skaalvik, S., & Skaalvik, E. M. (2004) - 3	Norway	264	128/136	15.89	--	--	Math and Verbal
Skaalvik, S., & Skaalvik, E. M. (2004) - 4	Norway	127	48/79	27.20	--	--	Computing
Smith, J. K., et al. (2012) - 1	New Zealand	480	240/240	8.5	--	--	Verbal (Reading)
Smith, J. K., et al. (2012) - 2	New Zealand	480	240/240	12.5	--	--	Verbal (Reading)
Stanisavljevic, D., et al. (2014)	Serbia	417	146/271	22.75	--	--	Math (Statistics)
Steiner, L. A. (2007)	United States	159	64/95	24	--	37% from ethnic minority	Math
Steinmayr, R., & Spinath, B. (2008)	Germany	342	138/204	16.94	Medium-high socioeconomic status	Majority Caucasian	Math and Verbal (German)

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Steinmayr, R., et al. (2014)	Germany	305	154/150	17.54	--	--	Math
Stephanou, G. (2008)	Greece	350	160/190	13.3	--	--	Math, Language and Physics
Stevens, T., et al. (2007)	United States	438	201/237	13.5	--	45.3% Mexican American; 45.6% European American; 7.6% other	Math
Stevenson, H. W., & Newman, R. S. (1986)	United States	255	57/48	8.5	--	--	Math and Verbal (Reading)
Tapia, M., & Marsh, G. E.	Mexico	545	302/245	15.5	--	--	Math
Tenenbaum, H. R., & Leaper, C. (2003)	United States	52	26/26	12.45	Middle income	Predominately European American	Science
Terwilliger, J. S., & Titus, J. C. (1995) - 1	United States	1,054	568/486	13.2	--	--	Math
Terwilliger, J. S., & Titus, J. C. (1995) - 1	United States	113	77/36	13.2	--	--	Math

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Thorndike-Christ, T. (1991)	United States	1,516	722/794	16.94	--	--	Math
Tissot, S. L. (1997) - 1	United States	73	41/32	11.5	Lower middle class	100% White	Math
Tissot, S. L. (1997) - 2	United States	45	25/21	11.5	Lower middle class	100% African American	Math
Tissot, S. L. (1997) - 3	United States	33	17/16	11.5	Lower middle class	100% Asian American	Math
Tissot, S. L. (1997) - 4	United States	58	21/37	11.5	Lower middle class	100% Hispanic	Math
Tissot, S. L. (1997) - 5	United States	89	42/47	13.5	Lower middle class	100% White	Math
Tissot, S. L. (1997) - 6	United States	71	28/43	13.5	Lower middle class	100% African American	Math
Tissot, S. L. (1997) - 7	United States	37	17/20	13.5	Lower middle class	100% Asian American	Math
Tissot, S. L. (1997) - 8	United States	60	28/32	13.5	Lower middle class	100% Hispanic	Math

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Tsai, C.-C., & Lin, C-C. (2004)	Taiwan	636	327/309	17	--	--	Math
Uitto, A (2014)	Finland	321	164/157	17	--	--	Math and Science (Biological/ Physical Sciences)
Urhahne, D., et al. (2012)	Germany	52	38/14	17.27	--	--	Science (Chemistry)
Vekiri, I. (2010)	Greece	301	135/166	13.5	37.4% upper-middle class; 41.2% middle class; 21.4% low SES (based on father's occupation)	--	Computing
Vekiri, I. (2013)	Greece	261	117/144	12	55% of students were from upper middle class; 28% in a school serving upper middle class and middle SES; 17% in a school serving low SES	--	Computing
Vekiri, I., & Chronaki, A. (2008)	Greece	340	174/166	10.5	23.5% from upper middle class background; 29.1% middle class; 47.4% low SES (based on father occupation and education	--	Computing

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Visser, D. (1986) – 1	South Africa	824	389/435	12.4	--	--	Math
Visser, D. (1986) – 2	South Africa	781	376/405	14.4	--	--	Math
Walles, R. L. (2009)	United States	88	46/42	8	--	--	Verbal (Reading)
Wang, J. (2012)	United States	8,976	4,756/4,220	15.5	--	37% ethnic minority	Math
Watt, H. M. G. (Steps Database)	Australia	1,136	639/497	12	--	--	Math
Weger-Guntharp, H. D. (2009)	United States	131	55/76	10.5	Middle class	100% international students in USA	Verbal (English)
Weinberg, A. E., et al (2011)	United States	336	158/158	--	--	54% ethnic minority; 46% non-minority	Science and Math
Weinburgh, M. H. (2000)	United States	1,034	517/517	12.5	--	45.26% from ethnic minority	Science
Weisgram, E. S., & Bigler, R. S.	United States	158	64/94	13.7	--	42% from ethnic minority	Science

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Wigfield, A., & Eccles, J. S. (1994) - 1	United States		276/310	--	Lower middle class	> 95% White	Math and Verbal (Reading)
Wigfield, A., & Eccles, J. S. (1994) - 2	United States	781	818/957	--	Lower middle class	> 95% White	Math and Verbal (Reading)
Wilhelm, S., & Brooks, D. M. (1980)	United States	241	112/129	13.5	Middle class	--	Math
Wolters, C. A., et al. (2014)	United States	406	220/186	15.56	69% economically disadvantaged	61% Hispanic, White 21%, 17% African American, 1% Asian, 1% other ethnicities	Verbal (Reading)
Wolters, C. A., Pintrich, P. R. (1998)	Australia	545	265/280	12.6	From working class suburb	95% Caucasian	Math and Verbal (English)
Wong, K. Y. (2006)	Hong Kong	4,626	2,280/2,346	15	--	55.6% Hong Kong; 18.7% first generation immigrants; 24.4% second generation immigrants	Science
Wong, S. L., & Hanafi, A. (2007)	Malaysia	102	29/73	21.06	Predominately middle class	--	Computing
Woods-McConney, A., et al. (2006) - 1	Australia	14,170	7,227/6,943	15	--	--	Science

(continued)

Author and Data Point	Country	Sample	Gender Ratio (M/F)	Age	Social class	Ethnicity %	Domains measured
Woods-McConney, A., et al. (2006) - 2	Canada	22,646	11,097/11,549	15	--	--	Science
Yancy, Y. G. (2013)	United States	56	27/29	11.5	51% of sample were economically disadvantaged	> 95% White	Math
Yang, F. Y., Tseng, J. S., & Lin, M. H. (2012)	Taiwan	277	122/140	13.5	Middle class	--	Science
Yeung, A. S., et al. (2011) - 1	Singapore	2,288	1,193/1,095	11	--	Chinese 75%, Malay 18%, Indian 5%, other 2%	Verbal (English)
Yeung, A. S., et al. (2011) -2	Singapore	1,926	919/1,007	16	--	Chinese 75%, Malay 18%, Indian 5%, other 2%	Verbal (English)
Zachai, J.	United States	111	50/61	33.3	--	10% from ethnic minority	Math

Supplementary Materials: Data Analysis with LSAY

The first time wave of Longitudinal Survey of Australian Youth (LSAY) utilises the responses of PISA participants on math, reading, and science achievement tests, as well as attitudes towards math. To measure achievement, the PISA survey provides five plausible values for each domain of achievement (e.g., math, science, and reading). Thus, this study contained five separate data-sets in this research, each containing one of five possible plausible values for math, reading, and science achievement. I used the formulas provided by Rubin (1987), whereby analyses were ran on each dataset separately, and parameter estimates were generated from the average estimates across the five datasets with standard errors corrected for the between plausible value variance.

LSAY is a large, complex database, and thus there are a number of recommendations provided to researchers who plan on analysing the data (see the website provided by the National Centre for Vocational Education Research, 2017). A critical issue in a large longitudinal database is attrition of participants. LSAY recommends that to counteract this issue researchers use attrition weights provided in the database to account for the effect of participant drop out. LSAY also recommends the use of sample weights that are included in the data provided to researchers. LSAY's sample weights ensure that the sample matches the population distribution. Using sample weights helps to ensure that any conclusions drawn from analyses are not altered by oversampling or undersampling of parts of the population. Finally, in order to be able to compare the estimates in the regression across different variables, I standardised the scales included in analyses (M = 0; SD = 1).

Supplementary Materials: Interview Schedule for Study 4

Categorical outcome variables for longitudinal analysis:

C104. In previous interviews we recorded that you did science, or maths subjects in Yr 12. Are you currently studying a science, engineering, mathematics or IT-related course?

1. Yes (Go to C105)
2. No (Go to C107)

C105. How important were each the following in your decision to study science, engineering, maths or IT:

a You were good at science, engineering, maths or IT.

b You wanted to pursue a career in science, engineering, maths or IT.

c You were influenced by career advice provided by your teachers or career advisers

d You were influenced by having good science or maths teachers in high school.

e You were influenced by your parents.

f You were influenced by one or more science related ‘experiences’ at high school.

g You have an employer who supports your study in this field.

h The course you are doing provides a good basis for employment in areas you like.

Was it….(READ OUT)

1 Very Important

2 Important

3 Neither Important or unimportant

4 Not Important

5 Not at all important

6 DON’T KNOW/CAN’T SAY

PRE 105a IF C105e = 1,2, CONTINUE

ELSE GO TO C106

C105a. Do you have a parent or close relative with a career in this field?

1. Yes

2. No

C106. What other factors (if any) influenced your decision to study science, engineering, maths or IT? (PROBE FULLY – RECORD VERBATIM)

C107. How important were each of the following in your decision NOT to study science, engineering, mathematics or IT after leaving school:

a You had no desire to work in science, engineering, maths or IT

b You thought those fields would not get you a well paid job

c You were influenced by advice from teachers and career advisers to consider other careers.

d Your parents influenced you to do other things.

e Science and maths teachers didn't inspire you enough to consider careers in science, engineering, maths or IT.

f Your friends did not study science, engineering, maths or IT.

g You were influenced against science, engineering, maths or IT by the negative image of them in the community.

Was it….(READ OUT)

1 Very Important

2 Important

3 Neither Important or unimportant

4 Not Important

5 Not at all important

6 DON'T KNOW/CAN'T SAY

PRE 107a IF C107g =1, 2 ASK C107a

ELSE GO TO C108

C107a Can you describe the image that you found off-putting?

(PROBE FULLY – RECORD VERBATIM)

C108. What other factors (if any) influenced your decision not to study science,

engineering, maths or IT?

(PROBE FULLY – RECORD VERBATIM)

C109. What would need to change for you to consider choosing to study science,engineering, maths or IT?

(PROBE FULLY – RECORD VERBATIM)

Study 4 Supplementary Tables

Table S6

Mean Rating of Importance for Choosing STEM

Ratings of importance for choosing STEM course	Female Mean	Male Mean	Total Mean
You were good at science, engineering, maths or IT	1.95 (0.99)	1.84 (0.91)	1.90 (0.95)
You wanted to pursue a career in science, engineering, maths or IT	2.07 (1.21)	2.02 (1.01)	2.04 (1.11)
You were influenced by career advice provided by your teachers or career advisers	3.24 (1.26)	3.21 (1.21)	3.23 (1.24)
You were influenced by having good science or maths teachers in high school	2.22 (1.09)	2.24 (1.09)	2.23 (1.09)
You were influenced by your parents	2.68 (1.26)	2.53 (1.18)	2.60 (1.22)
You were influenced by one or more science related 'experiences' at high school	2.74 (1.19)	2.67 (1.18)	2.70 (1.19)
You have an employer who supports your study in this field	3.26 (1.38)	3.47 (1.31)	3.37 (1.34)
The course you are doing provides a good basis for employment in areas you like	1.61 (0.87)	1.52 (0.68)	1.56 (0.77)

Ratings on 1-6 scale, 1= most important, 5 = not at all important 6 = don't know

Table S7

Mean Rating of Importance in not Choosing STEM

Ratings of importance in not choosing STEM	Female Mean	Male Mean	Total Mean
You had no desire to work in science, engineering, maths or IT	2.74 (1.38)	2.78 (1.34)	2.76 (1.36)
You thought those fields would not get you a well paid job	3.80 (0.93)	3.54 (1.07)	3.70 (1.00)
You were influenced by advice from teachers and career advisers to consider orher careers	3.24 (1.21)	3.25 (1.21)	3.24 (1.21)
Your parents influenced you to do other things.	3.44 (1.18)	3.26 (1.18)	3.37 (1.18)
Science and maths teachers didn't inspire you enough to consider careers in science, engineering, maths or IT	3.53 (1.10)	3.52 (1.10)	3.53 (1.10)
Your friends did not study science, engineering, maths or IT	3.99 (0.75)	3.91 (0.80)	3.96 (0.77)
You were influenced against science, engineering, maths or IT by the negative image of them in the community	4.10 (.67)	3.99 (0.75)	4.06 (0.70)

Ratings on 1-6 scale, lower scores are indicate higher importance, measures with statistically significant mean differences (p = <.05) highlighted

Table S8

Coding Framework, Frequency Counts and Significance for Motivation to Study STEM

	Female n	Expected n	Male n	Expected n	*p*
Previous exposure to STEM	19	18.8	22	22.2	0.95
Family, friends, mentors	25	20.6	20	24.4	0.17
Concern for society and environment	7	3.2	0	3.8	0.00
Attainment value (long-term goal)	19	19.7	24	23.3	0.17
Utility value (lifestyle)	10	7.8	7	9.2	0.27
Utility value (career opportunities)	40	40.8	49	48.2	0.85
Utility value (financial gain)	8	18.3	32	21.7	0.00
Intrinsic value	103	105.5	127	124.5	0.64
Expectancy for success	7	9.2	13	10.8	0.32
Lack of other options	7	3.7	1	4.3	0.02

Note. Statistically significant differences ($p < .05$) in counts bolded.

Table S9

Coding Framework, Frequency Counts and Significance for Barriers to Study STEM

	Female n	Expected n	Male n	Expected n	*p*
Lack of intrinsic value	115	110.1	60	64.9	0.31
Attracted to other non-STEM areas	63	71.1	50	41.9	0.67
Lack of perceived competence	37	38.4	24	22.6	0.70
Perception of being better in non-STEM areas	15	11.9	4	7.1	0.14
Low grades	6	10.1	10	5.9	0.03
Teacher influence	5	3.8	1	2.2	0.30
Lack of career opportunities or pathways	8	8.8	6	5.2	0.65
Family and friends	6	5.7	3	3.3	0.81
STEM too hard	15	15.7	10	9.3	0.76

Note. Statistically significant differences (p <.05) in counts bolded.

Table S10

Coding Framework, Frequency Counts and Significance for What Would Need to Change to Study STEM

	Female n	Expected n	Male n	Expected n	*p*
No suggestion	172	175.4	118	114.6	0.62
Change in affect or interest towards STEM	175	159.7	89	104.3	0.02
Change in interests/goals in non-STEM area	55	59.3	43	38.7	0.35
Change in perceived competence	53	44.8	21	29.2	0.04
Increased exposure	17	19.4	15	12.6	0.39
Better career opportunities	17	24.8	24	16.2	0.01
Better teaching	21	16.3	6	10.7	0.06
Change to content or curriculum	13	15.7	13	10.3	0.27
Financial incentive	12	21.8	24	14.2	0.00
Easier entry	5	6.7	6	4.3	0.31
Supportive and inclusive environment	12	9.1	3	5.9	0.12
Flexibility in study and work	16	13.9	7	9.1	0.37
Change in self or personality	14	13.9	9	9.1	0.97
Relevancy to life	10	9.7	6	6.3	0.87

Note. Statistically significant differences ($p < .05$) in counts bolded.

References for Supplementary Material

National Centre for Vocational Education Research (2017). *Longitudinal Surveys of Australian Youth.* Retrieved from https://www.lsay.edu.au/

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.