



Gender Gap in STEM Pathways: The Role of Gender-Segregated Schooling in Mathematics and Science Performance

Alice Smith¹ · Tanya Evans¹

Received: 3 October 2023 / Accepted: 26 March 2024 / Published online: 18 April 2024
© The Author(s) 2024

Abstract

According to the World Economic Forum, the occupational gender gap with men dominating well-remunerated STEM careers is a key factor in the gender pay gap. However, the problem of underrepresentation of women in STEM fields is mitigated or reversed in some Eastern European and Islamic countries. This disparity suggests that cultural factors are at play. Gender-segregated schooling, which is commonly practiced in Islamic countries, is often seen as a contributing factor. However, the issue of gender-segregated education sparks much controversy. Critics argue evidence is lacking, and segregation may foster sexism and hinder social skills.

In New Zealand, there is a significant opportunity to conduct research on gender segregation, as approximately 14.43% of girls and 14.35% of boys attend single-gender schools for their secondary (high school) education. In contrast to the USA and Australia, where single-gender schooling is predominantly offered by private and/or Catholic schools, the majority of New Zealand gender-segregated schools are state schools.

Utilising the TIMSS 2019 dataset, our analysis revealed that girls from low socio-economic backgrounds and boys from higher socio-economic backgrounds who attend single-gender schools exhibit significantly higher mathematics and science achievement scores compared to their counterparts in co-educational settings. This success is particularly striking for girls in low socio-economic settings. Given mixed international evidence, this suggests that there must be protective factors at single-gender New Zealand schools, positively influencing these students. We examine potential factors setting apart girls-only schools to provide an explanation of our results: (1) a mitigated influence of gender socialisation practices and stereotype threat and (2) explicit messaging aimed at challenging stereotypes and empowering girls to actively engage in STEM fields.

Keywords STEM education · Gender differences · Single-sex schools · Single-gender schools · TIMSS

Introduction

In nearly all countries in the OECD, there is a greater proportion of tertiary-educated women than men (Encinas-Martin, 2020). Specifically for New Zealand, the overall number of young women enrolled in Bachelor of Science courses is greater than the number of young men (Bray & Timewell, 2011), and women are well represented in biological and health science professions (Bray & Timewell, 2011; Huyer, 2015; Ministry of Business & Employment, 2014). However, women are still less likely to study physical science, technology, engineering, and mathematics than men and, consequently, remain a minority in these fields (Encinas-Martin, 2020; Ministry of Business & Employment, 2014; “Women at work: 1991–2013,” 2015).

A recent New Zealand government report estimated that in the technology industry only 27% of those working in digital roles are women (Hindle & Muller, 2021). This percentage has not increased since 2017, and in 2019, only 25% of those who graduated with IT degrees in New Zealand were women (Hindle & Muller, 2021). Recruiting women into the field of engineering also continues to be a struggle, with men still making up the majority of graduates (Docherty et al., 2020).

According to the World Economic Forum, the occupational gender gap with men dominating well-remunerated STEM careers is a key factor in the gender pay gap (Schwab et al., 2017). However, UNESCO reports that the problem of underrepresentation of women in STEM fields is mitigated or reversed in some Eastern or Islamic countries (Huyer, 2015). This disparity suggests that cultural factors are at play. It is posited that, at some stage before tertiary education, girls encounter discouragement from pursuing STEM pathways in many Western nations. This hypothesis posits that peers, parents, and educators play a key role in influencing young people’s career choices with mounting evidence supporting this assertion (Dasgupta & Stout, 2014; Ing, 2014; Lazarides & Watt, 2015; Lazarides et al., 2016; Muenks et al., 2020; Watson et al., 2015; Watt et al., 2019). Next, we summarise the existing literature exploring the various school factors that contribute to gender differences in occupational outcomes.

Societal Stratification: Gender Socialisation Practices and Stereotype Threat

The stereotype of male superiority in mathematics-related disciplines has persisted for generations, perpetuating the notion that men inherently possess a greater aptitude for the subjects despite mounting evidence against this misconception (Else-Quest et al., 2010; Fennema & Sherman, 1977). In fact, the recent data from a large-scale international assessment of 15-year-olds showed that boys significantly outperformed girls in mathematics in less than half of the 79 countries and economies that participated in PISA 2018. In science, boys’ performance was significantly better than girls’ performance in only six countries/economies, whereas the opposite, girls doing better, was observed in 35 countries/economies

(Encinas-Martin, 2020). Despite the evidence of closing the achievement gap, the hard-to-change age-old stereotypes profoundly impact societal stratification, instilling gendered socialisation practices.

It is well documented that young men are more likely to aspire to mathematics-related careers than young women (Watt, 2010, 2016; Watt et al., 2012, 2019). In contrast, young women express preferences for careers that bring to the fore their social needs and involve interacting with people (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000); that appear to be socially meaningful and relate to helping others, such as nursing (Eccles, 2007); or that would be compatible with child-rearing responsibilities (Jozefowicz et al., 1993). However, the degree of such societal stratification is varied among different countries and cultures, and is shaped by numerous religious, political and other historical factors (Huyer, 2015).

In New Zealand, there has been a longstanding presence of foundational beliefs about mathematics that favour boys (Watson et al., 2015). Gender stereotypes in subject choice persisted in the New Zealand curriculum for decades despite the Feminist movement, with girls typically choosing arts and boys dominating in sciences. This pattern can be traced back to an early twentieth century government report, which revealed that many parents held the belief that girls required only half as much education as boys, reflecting prevailing societal attitudes about women's inferiority (Fry, 1985). A measure of academic success for girls was their knowledge of "home science" and morality. As such, girls were given little instruction in mathematics and science, and few women were available to teach the subjects and act as role models. In this strongly gendered curriculum, generations of New Zealand girls have been subjected to messages of diminished expectation for their achievements in mathematics and science (Bradstreet, 2000). Despite major efforts in the 1990s to reform the New Zealand curriculum to achieve gender equality in education (Bradstreet, 2000), recent research indicates that the outcomes have not been uniformly successful (Watson et al., 2015).

Such cultural discrepancy is especially concerning because these negative stereotypes have the potential to hinder performance on mathematics tests and induce anxiety through the social psychology phenomenon called *stereotype threat* (Spencer et al., 1999, 2016; Steele, 1997). Stereotype threat refers to a situational predicament in which individuals are at risk of conforming to a negative stereotype about their social group, which in turn may lead to diminished performance in a particular domain. This phenomenon arises when individuals are aware of a negative stereotype associated with a group they belong to (such as gender, race, or age) and fear that their behaviour or performance will confirm that stereotype. This heightened anxiety and concern can adversely affect their performance, creating a self-fulfilling prophecy. The implication is that stereotype threat could be a contributing factor to gender disparities in advanced mathematics performance.

The Direct Impact of Teacher Gender

Traditionally, it was common for girls-only schools in New Zealand, as well as in many other countries, to have a preference for hiring women as teachers. This

preference was often based on the belief that women-teachers could serve as positive role models for girls and create a supportive and nurturing learning environment. In contrast, co-educational schools are generally less inclined to exhibit such gender-based biases in their hiring practices.

Notably, recent quantitative research examined the impact of teacher gender in a sample of New Zealand co-educational primary schools (Watson et al., 2015). Using hierarchical linear modelling, they found that in the classes taught by men, girls' mathematics scores were significantly lower than those of boys, which was not the case in classes taught by women. Given that the researchers did not identify variations in teachers' expectations for mathematics performance based on gender, they proposed a plausible hypothesis that other factors may be influencing this outcome, such as the presence of a man-teacher might inadvertently reinforce gender stereotypes implying inferior mathematical ability, consequently triggering stereotype threat among students (Steele, 1997; Watson et al., 2015). As described above, stereotype threat undermines performance and may lead to disengagement from the subject as individuals struggle with the fear of confirming a negative stereotype associated with their social group. With caution, the researchers proposed potential negative consequences for some girls' engagement with mathematics as they progress through their education if they are taught by a man. Consequently, they concluded that having a teacher of the same gender may benefit girls' mathematics achievement.

Theoretical Foundation

On the global stage, for four decades now, Eccles and her associates have formulated and scrutinised a comprehensive model, known as the Expectancy-Value model, that elucidates the social-cognitive mechanisms underpinning disparities in both individual and gender-related involvement in STEM fields (Eccles, 1983, 2007; Wigfield & Eccles, 2000). The core premise of the model is that engagement in an activity can be predicted by the expectancy a person has for succeeding at it, as well as the value they ascribe to the activity (Eccles & Wigfield, 2002). Numerous studies have collectively confirmed that beliefs linked to expectations (such as perceived competence, perceived talent, and self-concept), along with intrinsic/utility/importance values, play a predictive role in engagement, choices related to achievement, and career aspirations (for review, see Watt, 2010, 2016).

Confidence, self-concept and self-efficacy in mathematics are strong predictors of achievement (Michaelides et al., 2019), with boys tending to report higher levels than girls in many Western countries (Else-Quest et al., 2010). In Arab countries, however, gender differences in achievement and other education-related factors tend to favour girls (Marsh et al., 2013). Girls in these countries also reported having higher educational and career aspirations than boys (Marsh et al., 2013). Using the 2015 Trends in International Mathematics and

Science Study (TIMSS), Michaelides et al. (2019) discovered that in Iran, girls were overrepresented in both the high achievement and high-motivation groups, which was in sharp contrast to the results from other countries. One notable characteristic of the Iranian education system is that all girls attend girls-only schools where they are taught by women, and all boys attend boys-only schools where they are taught by men (Marsh et al., 2013; Michaelides et al., 2019). The absence of co-educational learning environments could potentially explain why the proportion of girls studying mathematics and science in Arab countries such as Iran is either on par with or surpasses that of boys (Huyer, 2015).

Research Questions

In New Zealand, approximately 14.43% of girls and 14.35% of boys attend single-gender schools for their secondary education (MoE, 2022). Out of 2544 schools in New Zealand, 64 are girls-only schools and 52 are boys-only schools spread across the country's main urban areas. The majority of gender-segregated schools are state (public) schools, which is different from the practice in the USA and Australia, where single-gender schooling is mainly found in private and/or Catholic schools.

Recent New Zealand findings from Docherty et al. (2020) have revealed that of all girls enrolled in engineering degrees at the University of Canterbury between 2005 and 2017, 56% had attended girls-only schools. This is almost five times higher than expected (Docherty et al., 2020). Pringle et al. (2010) also found that girls from girls-only schools in Auckland ranked more men-dominated professions in their top ten occupational choices than girls from co-educational schools in the same city. This contradicts the findings of a comprehensive study by Park et al. (2018) of Seoul schools in South Korea, where assignment to single-gender or co-educational high schools is random. They found significantly positive effects of all-boys schools consistently across different STEM outcomes but not for girls, which underscores the impact of cultural variability.

The aforementioned findings from Canterbury (Docherty et al., 2020) suggest that gender-segregated schools in New Zealand may somehow allow girls to engage more positively with STEM subjects traditionally dominated by men, and whilst there is extensive research in the literature on gender differences in achievement, there is very little on the effect that school type may have on these differences. The present study has been designed to investigate the role of school type (while controlling for socio-economic status) on the mathematics and science achievement of Year 9 girls and boys in New Zealand and will address the following research questions:

RQ1: Does the TIMSS 2019 mathematics performance of Year 9 girls and boys at single-gender and co-educational schools depend on the school type, controlling for socio-economic status?

RQ2: Does the TIMSS 2019 science performance of Year 9 girls and boys at single-gender and co-educational schools depend on the school type, controlling for socio-economic status?

RQ3: Are there any interaction effects between school type and socio-economic status on mathematics and science performance of Year 9 girls and boys?

Method

The study used data from the Trends in International Mathematics and Science Study (TIMSS) 2019 (<https://timssandpirls.bc.edu/timss2019/index.html>), which is a large-scale assessment of student achievement in mathematics and science. In line with the objectives of our study, we limited our focus to the New Zealand Year 9 cohort (TIMSS Year 8) since the data was collected at the end of the first year in New Zealand secondary schools. The sample consisted of 5,918 secondary schools' students, selected through a two-staged random sampling process. There were 906 girls attending girls-only schools (56.51% State [Public], 33.33% State-Integrated, 10.15% Private) and 970 boys attending boys-only schools (76.39% State [Public], 15.05% State-Integrated, 8.56% Private).

In New Zealand, the decile rating system was used to classify the socio-economic status of a school's roll (ranging 1–10), with decile 1 indicating that the school has one of the highest proportions of students from low socio-economic backgrounds (MoE, 2022). In the 2019 TIMSS data, the decile of each participating student's school was coded into the variable "decile". However, due to sampling, the data set did not include single-gender and co-educational schools from each decile. Because of this, the "decile" variable was recoded into three categories: "low decile" for decile 1–3 schools, "medium decile" for decile 4–7 schools, and "high decile" for decile 8–10 schools. As such, the sample used for the study is presented in Table 1.

TIMSS utilises plausible values as a statistical technique to estimate students' proficiency levels in mathematics and science. Plausible values are multiple imputed scores assigned to each student based on their responses to some assessment items and the overall distribution of scores. These values allow researchers to account for the uncertainty inherent in estimating student proficiency and provide more accurate representations of achievement at both individual and group levels. This is done to ensure robust and reliable comparisons of educational outcomes across different jurisdictions. In 2019 TIMSS cycle, five plausible values were estimated. In our analysis the "average mathematics achievement score" was calculated by averaging all five plausible values.

The study employed a two-way ANOVA analytical model to investigate the influence of school type and decile on the average mathematics and science achievement scores across all students collectively, followed by separate analyses for each gender. The variable "gender school type", comprising four

Table 1 TIMSS 2019 New Zealand Year 9 sample by gender school type and decile category

School and Student Gender	Decile category	Number of schools	Number of students
Girls – Single-Gender School	Low Decile (1–3)	2	83
	Mid Decile (4–7)	7	324
	High Decile (8–10)	12	499
	Total	21	906
Girls – Co-Educational School	Low Decile (1–3)	26	516
	Mid Decile (4–7)	38	828
	High Decile (8–10)	26	625
	Total	90	1969
Boys – Single-Gender School	Low Decile (1–3)	3	128
	Mid Decile (4–7)	7	356
	High Decile (8–10)	10	486
	Total	20	970
Boys – Co-Educational School	Low Decile (1–3)	26	484
	Mid Decile (4–7)	39	919
	High Decile (8–10)	26	670
	Total	91	2073
Total	Low Decile (1–3)	57	1211
	Mid Decile (4–7)	91	2427
	High Decile (8–10)	74	2280
	Total	222	5918

categories, (1) girls at single-gender girls schools, (2) girls at co-educational schools, (3) boys at single-gender boys schools, and (4) boys at co-educational schools, was constructed by segregating co-educational schools into two distinct categories for each gender.

Results

Mathematics Year 9

First, a two-way ANOVA was conducted to examine the effects of “gender school type” and “decile” on “average mathematics achievement score”. There was a statistically significant interaction between gender school type and decile on “average mathematics achievement score”, $F(6, 5906) = 6.44$, $p < 0.001$, partial $\eta^2 = 0.01$, indicating a small effect size (Richardson, 2011). There was a statistically significant difference in mean “average mathematics achievement score”

for students at low decile schools that are either girls at single-gender girls schools, girls at co-educational schools, boys at single-gender boys schools, or boys at co-educational schools, $F(3, 5906) = 16.26$, $p < 0.001$, partial $\eta^2 = 0.01$, as for students at medium decile schools, $F(3, 5906) = 10.44$, $p < 0.001$, partial $\eta^2 = 0.01$, and for students at high decile schools, $F(3, 5906) = 18.117$, $p < 0.001$, partial $\eta^2 = 0.01$ (Table 2 and Fig. 1). According to the guidelines set by the What Works Clearinghouse, designed specifically for educational settings and differing from those prevalent in the medical sciences, even a small effect size can carry substantial significance and practical relevance for educators (McMillan & Foley, 2019).

Expanding our understanding of group disparities, the pairwise comparisons were run for each simple main effect with reported 95% confidence intervals.

Low Decile

The “average mathematics achievement score” for girls at low decile single-gender schools ($M = 483.05$, $SD = 73.80$), girls at low decile co-educational schools ($M = 418.56$, $SD = 70.96$), boys at low decile single-gender schools ($M = 443.21$, $SD = 73.05$), and boys at low decile co-educational schools ($M = 424.02$, $SD = 79.56$) were compared. Girls at low decile single-gender girls schools had a statistically significantly higher mean “average mathematics achievement score” than any other group: higher than girls at low decile co-educational schools, 64.49 (95% CI, 38.63 to 90.34), $p < 0.001$, Hedges’ $d = 0.90$, 95% CI [0.67, 1.14] with a large effect size (Lakens, 2013); higher than boys at low decile single-gender schools, 39.83 (95% CI, 9.02 to 70.65), $p = 0.004$, Hedges’ $d = 0.54$, 95% CI [0.26, 0.82] with a medium effect size, and higher than boys at low decile co-educational schools, 59.02 (95% CI, 33.05 to 85.00), $p < 0.001$, Hedges’ $d = 0.75$, 95% CI [0.51, 0.99] with a medium effect size. The difference between boys in the two different school types was not significant, with a trivial effect size. Note that Hedges’ d effect size estimate was used instead of Cohen’s d to account for uneven group samples (Lakens, 2013).

Medium Decile

The mean “average mathematics achievement score” for girls at medium decile single-gender girls schools ($M = 509.44$, $SD = 77.44$), girls at medium decile co-educational schools ($M = 494.96$, $SD = 82.67$), boys at medium decile single-gender schools ($M = 523.69$, $SD = 85.66$), and boys at medium decile co-educational schools ($M = 505.89$, $SD = 90.00$) were compared. Girls at medium decile single-gender girls schools had a statistically significantly higher mean “average mathematics achievement score” than girls at medium decile co-educational schools, 14.48 (95% CI, 0.155 to 28.81), $p = 0.046$, with an effect size almost reaching the non-trivial threshold (Hedges’ $d = 0.18$, 95% CI [0.05, 0.32]). Boys at medium decile single-gender boys schools also had a statistically significantly higher mean “average mathematics achievement score” than boys at medium decile co-educational schools, 17.80 (95% CI, 4.15 to 31.45), $p = 0.003$, with small effect size (Hedges’ $d = 0.20$, 95% CI [0.08, 0.32]).

Table 2 Achievement scores by gender school type and decile category

Subject	Decile	Gender	Single-gender		Co-educational		F	Hedges' <i>d</i>	95% <i>CI</i>
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Maths	Low (1–3)	Girls	483.05	73.80	418.56	70.96	$F(1,527) = 58.39^{***}$	0.90	[0.67, 1.14]
		Boys	443.21	73.05	424.02	79.56	$F(1,610) = 6.09^*$	0.25	[0.05, 0.44]
	Mid (4–7)	Girls	509.44	77.44	494.96	82.67	$F(1,1150) = 7.40^{**}$	0.18	[0.05, 0.32]
		Boys	523.69	85.66	505.89	90.00	$F(1,1273) = 10.31^{***}$	0.20	[0.08, 0.32]
	High (8–10)	Girls	528.38	74.61	523.02	81.34	$F(1,1122) = 1.29$	0.07	[−0.05, 0.19]
		Boys	557.76	88.80	531.14	89.90	$F(1,1153) = 24.92^{***}$	0.30	[0.18, 0.42]
Science	Low (1–3)	Girls	501.94	77.68	433.20	79.15	$F(1,597) = 54.20^{***}$	0.87	[0.63, 1.11]
		Boys	453.86	87.31	435.73	86.63	$F(1,610) = 4.42^*$	0.21	[0.01, 0.40]
	Mid (4–7)	Girls	524.49	75.51	511.63	79.68	$F(1,1150) = 6.24^*$	0.16	[0.04, 0.29]
		Boys	539.94	85.74	522.52	91.69	$F(1,1273) = 9.59^{**}$	0.19	[0.07, 0.32]
	High (8–10)	Girls	547.80	72.69	544.47	73.29	$F(1,1122) = 0.57$	0.05	[−0.07, 0.16]
		Boys	569.62	83.32	547.92	85.96	$F(1,1153) = 18.40^{***}$	0.26	[0.14, 0.37]

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

High Decile

In the high decile band, the mean “average mathematics achievement score” for girls at high decile single-gender girls schools ($M=528.38$, $SD=74.61$), girls at high decile co-educational schools ($M=523.02$, $SD=81.34$), boys at high decile single-gender boys schools ($M=557.85$, $SD=88.72$), and boys at high decile co-educational schools ($M=531.14$, $SD=89.90$) were compared. Boys at high decile single-gender boys schools had a statistically significantly higher mean “average mathematics achievement score” than all other groups. In particular, boys at high decile single-gender schools scored higher than boys at co-educational schools, 26.71 (95% *CI*, 13.68 to 39.74), $p < 0.001$ with small effect size (Hedges’ $d=0.30$, 95% *CI* [0.18,0.42]). However, while the girls at high decile single-gender schools scored higher than girls at co-educational schools, the difference was insignificant with a trivial effect size (Hedges’ $d=0.07$, 95% *CI* [-0.05,0.19]).

Overall Gender-Specific Effects for Mathematics

After splitting the data set with respect to gender, two-way ANOVAs were conducted to examine the effects of school type and decile on average mathematics achievement scores separately for girls and boys.

For girls, the results indicated a significant main effect of school type, with girls at single-gender schools performing significantly better than girls at co-educational schools, $F(1, 2869) = 53.18$, $p < 0.001$, partial $\eta^2 = 0.02$. As expected, there was also a significant main effect of the decile category, $F(2, 2869) = 107.42$, $p < 0.001$, partial $\eta^2 = 0.07$. However, a significant interaction $\eta^2 = 0.01$ between decile category and school type, $F(2, 2869) = 16.37$, $p < 0.001$, partial $\eta^2 = 0.01$ was an unexpected finding.

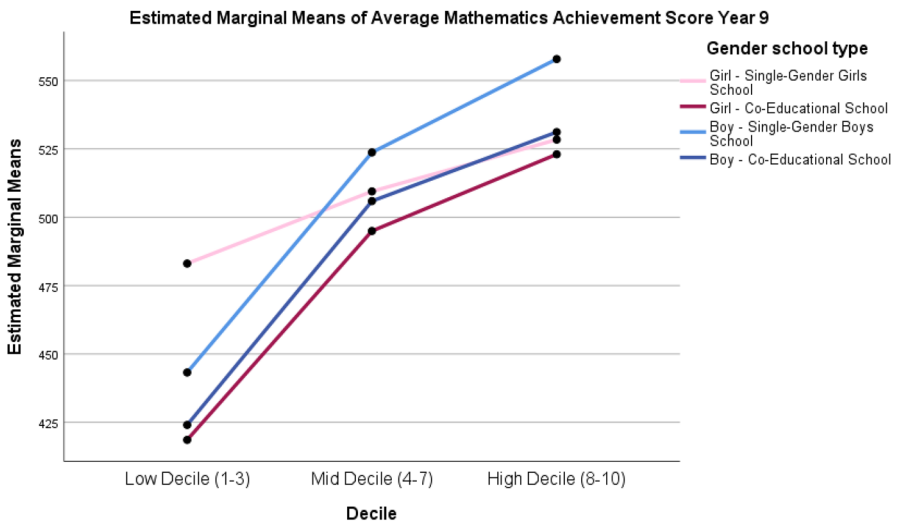


Fig. 1 Interaction between school type and school decile on mathematics achievement

This reflects the variations of the effect sizes for pairwise comparison of single-gender vs co-educational schooling (Table 2): girls at low-decile single-gender schools are performing significantly better than their counterparts in co-ed schools (Hedges' $d=0.90$, 95% $CI [0.67,1.14]$) with a large effect size (Lakens, 2013). Whereas the effect size at mid-decile was close to small (Hedges' $d=0.18$, 95% $CI [0.05,0.32]$), and the high-decile effect (Hedges' $d=0.07$, 95% $CI [-0.05,0.19]$) was trivial.

For boys, the results indicated significant main effect of school type, with boys at single-gender schools performing significantly better than boys at co-ed schools, $F(1, 3036)=30.82$, $p<0.001$, partial $\eta^2=0.01$. There was also a significant large effect for the decile category, $F(2, 3036)=241.41$, $p<0.001$, partial $\eta^2=0.14$, albeit not unexpected. However, in contrast with the girls, no significant interaction between decile category and school type was identified, $F(2, 3036)=0.75$, $p=0.474$. This is reflected in the uniformity of the effect sizes across the decile categories, with boys at single-gender schools performing significantly better than in co-ed schools with similar small effect sizes across all categories (Low-decile: Hedges' $d=0.25$, 95% $CI [0.05,0.44]$; Mid-decile: Hedges' $d=0.20$, 95% $CI [0.05,0.31]$; High-decile: Hedges' $d=0.30$, 95% $CI [0.18,0.42]$).

Science Year 9

A two-way ANOVA was conducted to examine the effects of gender school type and decile on average science achievement scores. There was a statistically significant interaction between gender school type and decile on average science achievement score, $F(6, 5906)=7.499$, $p<0.001$, partial $\eta^2=0.01$. There was a statistically significant difference in mean “average science achievement score” for students at low decile schools that are either girls at single-gender girls schools, girls at co-educational schools, boys at single-gender boys schools, or boys at co-educational schools, $F(3, 5906)=18.389$, $p<0.001$, partial $\eta^2=0.01$, as for students at medium decile schools, $F(3, 5906)=10.096$, $p<0.001$, partial $\eta^2=0.01$, and for students at high decile schools, $F(3, 5906)=10.132$, $p<0.001$, partial $\eta^2=0.01$ (Fig. 2 and Table 2).

The pairwise comparisons were run for each simple main effect with reported 95% confidence intervals.

Low Decile

The mean “average science achievement score” for girls at low decile single-gender girls schools ($M=501.94$, $SD=77.68$), girls at low decile co-educational schools ($M=433.20$, $SD=79.15$), boys at low decile single-gender boys schools ($M=453.86$, $SD=87.31$), and boys at low decile co-educational schools ($M=435.73$, $SD=86.63$) were compared. Girls at low decile single-gender girls schools had a statistically significantly higher mean “average science achievement score” than any other group: higher than girls at low decile co-educational schools, 68.74 (95% CI , 43.04 to 94.45), $p<0.001$, Hedges' $d=0.87$, 95% $CI [0.63,1.11]$) with a large effect size (Lakens, 2013), higher than boys at low decile boys schools,

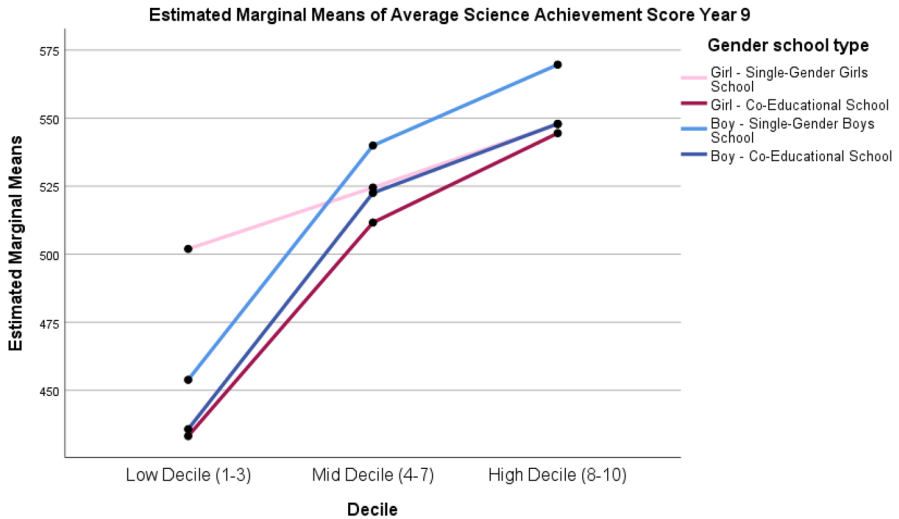


Fig.2 Interaction between school type and school decile on science achievement

48.09 (95% CI, 17.45 to 78.72), $p < 0.001$, Hedges' $d = 0.57$, 95% CI [0.29, 0.85]) with a medium effect size, and higher than boys at low decile co-educational schools, 66.22 (95% CI, 40.39 to 92.04), $p < 0.001$, Cohen's $d = 0.77$, 95% CI [0.54, 1.01]) with a medium effect size. The difference between boys in the two school types was also significant, favouring single-gender schooling with a small effect size (Hedge' $d = 0.21$, 95% CI [0.01, 0.40]).

Medium Decile

The mean “average science achievement score” for girls at medium decile girls schools ($M = 524.49$, $SD = 75.51$), girls at medium decile co-educational schools ($M = 511.63$, $SD = 79.68$, boys at medium decile single-gender boys schools ($M = 539.94$, $SD = 85.74$), and boys at medium decile co-educational schools ($M = 522.53$, $SD = 91.69$) were compared. Boys at medium decile single-gender boys schools had a statistically significantly higher mean “average science achievement score” than boys at medium decile co-educational schools, 17.41 (95% CI, 3.84 to 30.98), $p = 0.004$, Hedges' $d = 0.19$, 95% CI [0.07, 0.32]. Girls at single-gender schools scored higher than girls at co-educational schools, but the difference was not significant, with Hedges' $d = 0.16$, 95% CI [0.04, 0.30].

High Decile

In the high decile category, the means of “average science achievement score” for girls at single-gender girls schools ($M = 547.80$, $SD = 72.69$), girls at co-educational schools ($M = 544.47$, $SD = 73.79$), boys at single-gender boys schools ($M = 569.64$, $SD = 83.23$), and boys at high decile co-educational schools ($M = 547.92$,

$SD=85.96$) were compared. Boys at high decile single-gender boys schools had a statistically significantly higher mean “average science achievement score” than all other groups. In particular, their scores were higher than the scores of boys at high decile co-educational schools 21.72 (95% CI , 8.77 to 34.67), $p < 0.001$, Hedges’ $d=0.26$, 95% CI [$0.14, 0.37$], with small effect size. Girls at high decile single-gender schools scored higher than girls at co-educational schools, but the difference was insignificant, with a trivial effect size (Hedges’ $d=0.05$, 95% CI [$-0.07, 0.16$]).

Overall Gender-Specific Effects for Science

After splitting the data set with respect to gender, two-way ANOVAs were conducted to examine the effects of school type and decile on average science achievement score separately for girls and boys.

For girls, the results indicated a significant main effect of school type, with girls at single-gender schools performing significantly better than girls at co-educational schools, $F(1, 2869)=56.14$, $p < 0.001$, partial $\eta^2=0.02$. As expected, there was significant main effect of decile category, $F(2, 2869)=126.36$, $p < 0.001$, partial $\eta^2=0.08$. However, a significant interaction between decile category and school type, $F(2, 2869)=20.898$, $p < 0.001$, partial $\eta^2=0.01$ was an unexpected finding, which was similar to the intriguing discrepancy identified in mathematics. This is reflected in the variations of the effect sizes for pairwise comparison of single-gender vs co-educational schooling (Table 2): girls at low decile single-gender schools are performing significantly better than in co-ed schools (Hedges’ $d=0.87$, 95% CI [$0.63, 1.11$]) with a large effect size (Lakens, 2013). In contrast, the effect size at mid decile is not so substantial (Hedges’ $d=0.16$, 95% CI [$0.04, 0.30$]) and trivial at high decile (Cohen’s $d=0.05$, 95% CI [$-0.07, 0.16$]).

For boys, the results indicated a significant main effect of school type, with boys at single-gender schools performing significantly better than boys at co-ed schools, $F(1, 3036)=24.72$, $p < 0.001$, partial $\eta^2=0.01$. There was a significant large effect of decile category, $F(2, 3036)=253.11$, $p < 0.001$, partial $\eta^2=0.14$, albeit not unexpected. However, in contrast with the girls, no significant interaction between decile category and school type was identified, $F(2, 3036)=0.17$, $p=0.84$. This is reflected in the uniformity of the effect sizes across the decile categories with boys at single-gender schools performing significantly better than in co-ed schools with similar small (or close to non-trivial) effect sizes (Low-decile: Hedges’ $d=0.21$, 95% CI [$0.01, 0.40$]; Mid-decile: Hedges’ $d=0.19$, 95% CI [$0.07, 0.32$]; High-decile: Hedges’ $d=0.26$, 95% CI [$0.14, 0.37$]).

Discussion

Using the TIMSS 2019 data, the present study investigated the main effect of schooling type (single-gender versus co-ed) and interaction effect of decile (low, mid, and high) on the mathematics and science achievement of Year 9 New Zealand students. The findings revealed that students at single-gender schools perform better than their counterparts in co-educational settings for both mathematics and science,

albeit with varied effect sizes depending on gender and decile. There is a significant interaction between gender school type and decile for girls (but not for boys), reflecting a more pronounced positive effect within low-decile girls-only schools. Mathematics and science achievement scores of girls at low decile single-gender girls' schools are significantly higher (with large effect sizes) than the scores of all other low decile groups (including boys at both co-ed and single-gender settings), which is an intriguing discrepancy. Conversely, among boys, the performance disparities were most prominent in high decile settings, where scores at boys schools significantly outstripped those of their counterparts in other schools.

The consistent observation that boys perform better academically in single-gender schools compared to co-educational settings is well-documented (Park et al., 2018; Rowe, 1990). This effect is further amplified when young boys are taught by men, as evidenced by a large-scale study conducted in South Korea (Park et al., 2018). However, it is important to note that the New Zealand context may present unique factors influencing these outcomes. For example, considering whether single-gender boys schools excel in providing quality education that is tailored to support boys' needs would be worthwhile.

Central to our study's objectives lies a pivotal finding: girls enrolled in single-gender schools demonstrate superior academic performance in both mathematics and science compared to their counterparts in co-educational settings, albeit with varied effect sizes across the socio-economic categories. Particularly striking is the large effect size observed within the low decile settings. Given the contradicting international evidence (Park et al., 2018; Watson et al., 2015), this suggests that there must be protective factors at single-gender New Zealand schools, positively influencing these students. However, the potential skew of a small sample size in the low decile category is worth pointing out – only two low decile schools were included in the TIMSS 2019 data, with 83 girls participating in the study. It could well be that a major factor influencing our results is not due to some generalisable protective factors but rather is caused by the practices of a small number of schools. This warrants further investigation.

Based on the available body of literature, we next discuss the possibilities that could account for our findings.

Manifestation of Stereotype Threat

The results of our study can be interpreted through the lens of the Expectancy Value Theory, which premises that engagement in an activity can be predicted by the expectancy a person has for succeeding at it, as well as the value they ascribe to the activity (Eccles & Wigfield, 2002). Given the observed positive effect of girls-only schooling on girls' mathematics and science performance in a representative sample of Year 9 New Zealand schools, it is plausible to suggest that the negative influence of gender stereotypes ingrained in New Zealand's culture is somewhat mitigated within the environment afforded in girls-only schools. Perhaps the absence of boys in maths/science classes diminish the manifestation of societal stratification and removes the immediacy of stereotype threat.

Research in New Zealand supports this. Case studies in co-educational primary schools highlight boys domination in classroom interactions, with boys receiving more attention from the teacher (Bradstreet, 2000). Moreover, boys are reported to receive positive reinforcements and rewards for behaviours that are considered normal for girls (Bradstreet, 2000). Other studies show boys tend to talk more and receive more praise (Coxon & Jenkins, 1994), prompting arguments from some researchers that girls may experience marginalisation within co-educational settings (Bradstreet, 2000; Coxon & Jenkins, 1994).

Furthermore, a comprehensive global review has revealed that teachers' gender beliefs significantly influence classroom dynamics, often resulting in a bias towards favouring boys over girls in various aspects of academic interaction and engagement (Li, 1999). This underscores the pervasive impact of educators' perceptions on shaping the educational experiences of students along gender lines. This dynamic might be at play within co-educational high school environments in New Zealand, potentially providing an explanation for the superior performance of girls in single-gender schools.

This is combined with a large body of literature that has consistently identified gender differences favouring young boys in their self-perceived mathematical ability, talent, and self-efficacy (e.g., Watt et al., 2012), noting that these gender disparities persist even in the absence of any significant difference in actual performance (Evans et al., 2021; Watt, 2016). However, these self-assured beliefs can potentially manifest in behavioral distinctions, consequently exacerbating the daily ramifications of stereotype threat experienced by girls in co-educational environments.

Explicit Messages to Break Stereotypes and Empower Girls to Participate in STEM

Contemporary girls-only schools purposefully position themselves as a counterbalance to environments marked by gender stereotypes. For example, girls-only St Cuthbert's College outlines all-girl advantage as having "Can-do attitude: where aspirations are not limited by gender" and thriving for "Academic performance: girls' schools are leading the way in STEM subjects (Science, Technology, Engineering, Maths)" (<https://stcuthberts.school.nz/about-st-cuthberts/st-cuthberts-advantage/why-an-all-girls-school/>, retrieved 28/09/2023). Regular activities include speeches and presentations by successful women who are STEM leaders and messages of empowerment by the principal and teachers that refer to breaking gender stereotypes and shattering glass ceilings.

It is known from research in Australian and Canadian secondary school samples that an attainment/utility, or "importance" value, predicts mathematics-related career plans for only young women (Watt, 2016; Watt et al., 2019). Eccles and her colleagues have previously shown that tasks considered socially meaningful and significant are particularly engaging for young women (Eccles, 2007; Eccles et al., 2004). However, mathematics is frequently presented in an abstract and decontextualised manner, which may not resonate strongly with adolescent girls' interests and thereby diminish the perceived value they attach to the subject (Watt, 2016). Adolescents

often hold misconceptions about careers requiring advanced mathematical skills, so their perceptions may be quite inaccurate. As such, an explicit strategic focus characterising girls-only schools to provide information about the mathematics required for different kinds of rewarding careers may enhance young women's interest and valuation of the subject, which, according to the Expectancy-Value Theory, can lead to improved learning engagement.

Study Limitations, Future Research and Recommendations for Practice

TIMSS data is cross-sectional, offering a snapshot of performance at a specific time, limiting the ability to establish causal relationships or capture developmental trajectories. Having access to only Year 9 data limited our ability to control for primary school experiences in terms of school gender type. However, it may not have had an extra explanatory power since only a very small proportion (1.01% of girls and 1.02% of boys) attend New Zealand single-gender primary schools.

In conclusion, the study's compelling finding that girls at low decile and boys at higher deciles attending single-gender schools demonstrate higher mathematics and science achievement scores than their co-educational counterparts holds valuable implications for educational practice and signals the need for further research. The consideration arises regarding whether high decile boys' schools excel in delivering quality education and tailored support for boys' needs, and whether girls in low decile schools merely benefit from an environment with fewer stereotypes. Therefore, it is important to note that the recommendation for educators and policymakers to consider the benefits of creating more gender-segregated educational environments is premature and unjustified. This is because we were unable to control for the prior achievement of students entering at the start of Year 9, and thus, we cannot account for the selection bias – it could be that academically strong learners prefer to enrol into single-gender schools. However, it is worth noting that students in many New Zealand regions have no choice due to zoning since public single-gender schools are the only option. Nevertheless, even if the selection bias is the case, it is plausible to suggest, given the evidence, that single-gender girls schools successfully provide supportive environments that enable girls' achievement in mathematics and science. This success is particularly striking in low-decile settings, which constitutes an unexpected finding deserving of deeper investigation into the unique practices and environments of these schools.

Another limitation is that the data originate from the end of the first year of secondary education. Consequently, we lack information regarding how schooling type and socio-economic status influence outcomes for girls and boys over the subsequent four years of schooling. Therefore, the main implication of this study is the need for further longitudinal research in identifying specific protective factors that constitute beneficial gender-segregated learning environments. Further investigations need to be carried out to examine (1) the role of teacher gender in single-gender schooling environments, (2) specialised pedagogical approaches tailored to the learning preferences and strengths of girls and boys in these schools, (3) the existing mentorship

and role modelling initiatives featuring successful women in STEM fields that inspire and motivate girls to pursue careers in mathematics and science, (4) extra-curricular activities, workshops, and programs focused on STEM subjects that foster a supportive and stimulating learning environment, and (5) strategies in professional development to enable gender-equitable instruction that challenges stereotypes.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data availability The data that support the findings of this study are available from the New Zealand Ministry of Education, but restrictions apply to the availability of these data, which were used under the National TIMSS Data Release and Confidentiality Agreement for the current study and so are not publicly available. The data are, however, available from the corresponding author upon reasonable request and with the permission of the Ministry of Education.

Declarations

Conflict of interest The authors have no conflict of interest to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bradstreet, A. (2000). Gender Equity in the New Zealand Curriculum: A change in focus. *ACE Papers*, 7, 89–105.
- Bray, B., & Timewell, E. (2011). *Women in Science: A 2011 Snapshot*. Hamilton, Association for Women in the Sciences.
- Coxon, E., & Jenkins, K. (1994). *The politics of learning and teaching in Aotearoa-New Zealand*. Dunmore Press.
- Dasgupta, N., & Stout, J. G. (2014). Girls and women in science, technology, engineering, and mathematics: STEMing the tide and broadening participation in STEM careers. *Policy Insights from the Behavioral and Brain Sciences*, 1(1), 21–29. <https://doi.org/10.1177/2372732214549471>
- Docherty, P. D., Fox-Turnbull, W. H., Naswall, K., Homewood, A., Bradley, H., Zaka, P., & Chase, J. G. (2020). The contribution of gender segregated secondary education on the progression to engineering. *Australasian Journal of Engineering Education*, 25(1), 31–38. <https://doi.org/10.1080/22054952.2020.1788254>
- Eccles, J. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and Achievement Motives: Psychological and Sociological Approaches*. (pp. 75–146). W.H. Freeman. <https://books.google.co.nz/books?id=wMtnQgAACAAJ>
- Eccles, J. S. (2007). *Where are all the women? Gender differences in participation in physical science and engineering*. American Psychological Association.
- Eccles, J. S., Vida, M. N., & Barber, B. (2004). The relation of early adolescents' college plans and both academic ability and task-value beliefs to subsequent college enrollment. *The Journal of Early Adolescence*, 24(1), 63–77. <https://doi.org/10.1177/0272431603260919>
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132.

- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, *136*(1), 103–127.
- Encinas-Martin, M. (2020). Why do gender gaps in education and work persist. In: *OECD education and skills today*.
- Evans, T., Thomas, M. O. J., & Klymchuk, S. (2021). Non-routine problem solving through the lens of self-efficacy. *Higher Education Research & Development*, *40*(7), 1403–1420. <https://doi.org/10.1080/07294360.2020.1818061>
- Fennema, E., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. *American Educational Research Journal*, *14*(1), 51–71.
- Fry, R. (1985). *It's different for daughters: A history of the curriculum for girls in New Zealand Schools, 1900–1975*. Educational Research Series No. 65. ERIC.
- Hindle, S., & Muller, G. (2021). *Digital skills Aotearoa: Digital skills for our digital future*. MBIE.
- Huyer, S. (2015). Is the gender gap narrowing in science and engineering? In *UNESCO science report: towards 2030* (pp. 85–103). UNESCO.
- Ing, M. (2014). Can parents influence children's mathematics achievement and persistence in STEM careers? *Journal of Career Development*, *41*(2), 87–103. <https://doi.org/10.1177/0894845313481672>
- Jozefowicz, D. M., Barber, B. L., & Eccles, J. S. (1993). *Adolescent work-related values and beliefs: Gender differences and relation to occupational aspirations*. Biennial meeting of the Society for Research on Child Development.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. <https://doi.org/10.3389/fpsyg.2013.00863>
- Lazarides, R., Viljaranta, J., Aunola, K., Pesu, L., & Nurmi, J.-E. (2016). The role of parental expectations and students' motivational profiles for educational aspirations. *Learning and Individual Differences*, *51*, 29–36. <https://doi.org/10.1016/j.lindif.2016.08.024>
- Lazarides, R., & Watt, H. M. (2015). Girls' and boys' perceived mathematics teacher beliefs, classroom learning environments and mathematical career intentions. *Contemporary Educational Psychology*, *41*, 51–61.
- Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research*, *41*(1), 63–76.
- Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J., Abdelfattah, F., Leung, K. C., Xu, M. K., Nagengast, B., & Parker, P. (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.
- McMillan, J. H., & Foley, J. (2019). Reporting and discussing effect size: Still the road less traveled? *Practical Assessment, Research, and Evaluation*, *16*(1), 14.
- Michaelides, M. P., Brown, G. T., Eklöf, H., & Papanastasiou, E. C. (2019). *Motivational profiles in TIMSS mathematics: Exploring student clusters across countries and time*. Springer Nature.
- Ministry of Business, I., & Employment. (2014). *A nation of curious minds—a national strategic plan for science in society*. New Zealand Government.
- MoE. (2022). *Education Counts*. Retrieved from <https://www.educationcounts.govt.nz/statistics/school-rolls>
- Muenks, K., Peterson, E. G., Green, A. E., Kolvoord, R. A., & Uttal, D. H. (2020). Parents' beliefs about high school students' spatial abilities: Gender differences and associations with parent encouragement to pursue a STEM career and students' STEM career intentions. *Sex Roles*, *82*(9–10), 570–583. <https://doi.org/10.1007/s11199-019-01072-6>
- Park, H., Behrman, J. R., & Choi, J. (2018). Do single-sex schools enhance students' STEM (science, technology, engineering, and mathematics) outcomes? *Economics of Education Review*, *62*, 35–47. <https://doi.org/10.1016/j.econedurev.2017.10.007>
- Pringle, J. K., Rothman, M., Ryan, I., Lin, J., Warner, A., & Sisman, R. (2010). *Students' occupational choice study: Dunedin, Auckland*. Ministry of Women's Affairs.
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147. <https://doi.org/10.1016/j.edurev.2010.12.001>
- Rowe, K. (1990). What are the benefits of single sex maths classes? *Set*(1).
- Schwab, K., Samans, R., Zahidi, S., Leopold, T., Ratcheva, V., Hausmann, R., & Tyson, L. D. (2017). *The global gender gap report 2017* [Report]. World Economic Forum. <https://apo.org.au/node/208501>

- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67(1), 415–437.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35(1), 4–28.
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613.
- Watson, P. W. S., Rubie-Davies, C. M., Meissel, K., Flint, A., Peterson, E. R., Garrett, L., & McDonald, L. (2015). Gendered teacher expectations of mathematics achievement in New Zealand: Contributing to a Kink at the base of the STEM pipeline? *International Journal of Gender, Science and Technology*, 8(1), 82–102. <https://genderandset.open.ac.uk/index.php/genderandset/article/view/394>
- Watt, H. M. (2010). Gender and occupational choice. *Handbook of Gender Research in Psychology: Volume 2. Gender Research in Social and Applied Psychology*, 379–400.
- Watt, H. M. (2016). Gender and motivation. *Handbook of motivation at school*, 320–339.
- Watt, H. M., Bucich, M., & Dacosta, L. (2019). Adolescents' motivational profiles in mathematics and science: Associations with achievement striving, career aspirations and psychological wellbeing. *Frontiers in Psychology*, 10, 990.
- Watt, H. M., Shapka, J. D., Morris, Z. A., Durik, A. M., Keating, D. P., & Eccles, J. S. (2012). Gendered motivational processes affecting high school mathematics participation, educational aspirations, and career plans: A comparison of samples from Australia, Canada, and the United States. *Developmental Psychology*, 48(6), 1594.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>
- Women at work: 1991–2013. (2015). Ministry for Women; Statistics New Zealand.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Alice Smith¹  · Tanya Evans¹ 

✉ Tanya Evans
t.evans@auckland.ac.nz

Alice Smith
asmi887@aucklanduni.ac.nz

¹ Department of Mathematics, University of Auckland, Auckland, New Zealand