

Controlling for Prior Attainment Reduces the Positive Influence that Single-Gender Classroom Initiatives Exert on High School Students' Scholastic Achievements

Charlotte R. Pennington¹ · Linda K. Kaye² · Adam W. Qureshi² · Derek Heim²

© The Author(s) 2017. This article is an open access publication

Abstract Research points to the positive impact that gender-segregated schooling and classroom initiatives exert on academic attainment. An evaluation of these studies which reveal positive effects highlights, however, that students are typically selectively assigned to single- or mixed-gender instructional settings, presenting a methodological confound. The current study controls for students' prior attainment to appraise the efficacy of a single-gender classroom initiative implemented in a co-educational high school in the United Kingdom. Secondary data analysis (using archived data) was performed on 266 middle-ability, 11–12 year-old students' standardized test scores in Languages (English, foreign language), STEM-related (Mathematics, Science, Information and Communication Technology), and Non-STEM subjects (art, music, drama). Ninety-eight students (54, 55% female) were taught in single-gender and 168 (69, 41% female) in mixed-gender classrooms. Students undertook identical tests irrespective of classroom type, which were graded in accordance with U.K national curriculum guidelines. Controlling for students' prior attainment, findings indicate that students do not appear to benefit from being taught in single-gender relative to mixed-gender classrooms in Language and STEM-related subjects. Young women benefitted from being taught in mixed-gender relative to single-gender classes for Non-

STEM subjects. However, when prior ability is not controlled for, the intervention appears to be effective for all school subjects, highlighting the confounding influence of selective admissions. These findings suggest that gender-segregated classroom initiatives may not bolster students' grades. It is argued that studies that do not control for selection effects may tell us little about the effectiveness of such interventions on scholastic achievement.

Keywords Single-gender classrooms · Co-education · Achievement · Gender-achievement gap · Education policy

The gender-achievement gap is well documented in Western cultures across a number of different subject domains (Else-Quest et al. 2010; Stoet and Geary 2013). Current research indicates that females outperform males typically across the majority of school subjects (Mullholland et al. 2004; Voyer and Voyer 2014), particularly in English literacy (Young-Suk et al. 2015). However, there is considerable variation when exploring gender differences in mathematics, with females underperforming in comparison to males at the high end of the distribution (Ceci and Williams 2010; Reilly et al. 2015; Stoet and Geary 2013; Wai et al. 2010).

Many factors have been proposed to account for differences in females' and males' academic performance. For example, boys tend to report higher academic self-efficacy in mathematics (Dai 2001), whereas girls report higher self-efficacy in English literacy (Niederle and Vesterlund 2010; Pajares and Valiante 2001). Furthermore, children's academic self-efficacy has been found to be correlated with parents' and teachers' beliefs of gender-subject competence (Bleeker and Jacobs 2004; Miller et al. 2015; Tiedemann 2002; Wood et al. 2010). Gender differences in academic attainment may arise due to the format of achievement tests, with research

Electronic supplementary material The online version of this article (doi:10.1007/s11199-017-0799-y) contains supplementary material, which is available to authorized users.

✉ Charlotte R. Pennington
c.r.pennington@lancaster.ac.uk; charlotte.pennington@outlook.com

¹ Department of Psychology, Lancaster University, Bailrigg, Lancashire LA4 1YW, UK

² Department of Psychology, Edge Hill University, Ormskirk, UK

suggesting that boys excel on standardized tests relative to girls who do better in coursework-based examinations (Ceci et al. 2009; Kimball 1989). In attempt to reduce performance clefts, other research has moved beyond these factors to examine the direct role of the learning context (Park et al. 2013; Sullivan et al. 2010). Despite being met with considerable controversy (Bigler and Signorella 2011; Pahlke et al. 2014; Pahlke et al. 2013; Signorella and Bigler 2013), one solution that has been proposed is single-gender schooling.

Proponents of single-gender schooling suggest that the segregation of females and males has a positive impact on their academic self-concept (Sullivan 2009), educational transition (Lee and Marks 1990; Park et al. 2013), and attainment and interest (Else-Quest and Peterca 2015). Other research indicates that female students benefit more from single-gender schooling compared to males (Alon and Gelbgiser 2011; Else-Quest and Peterca 2015; Lee and Bryk 1986; Mullholland et al. 2004), with such environments suggested to lessen the impact of gender stereotypes on females' interest and performance in STEM-related subjects (Inzlicht and Ben-Zeev 2003; Shapka and Keating 2003). For example, females report higher competence beliefs and tend to achieve higher grades in mathematics and science when they are taught in single-gender relative to co-educational schools (Eisenkopf et al. 2015; Hoffman 2002).

Those taught in single-gender schools also report fewer experiences of gender stereotyping compared to their mixed-gender counterparts (Pahlke et al. 2014). Gender-segregated learning environments have therefore been suggested to alleviate experiences of *stereotype threat*, a situational phenomenon whereby young women and men apprehend that their performance will be evaluated in line with gender-related expectations (Elizaga and Markman 2008; Huguet and Régner 2007; Inzlicht and Ben-Zeev 2000, 2003; Picho and Stephens 2012). Some research indicates that gender-segregated education has a neutral impact on males' academic attainment (Sullivan et al. 2010), whereas other research suggests that males benefit more from being taught in co-educational settings (Jackson and Smith 2000; Schneider and Coutts 1982).

Opponents of such educational initiatives, however, argue that single-gender schooling may exacerbate gender stereotyping because students question why they have been separated from their other-sex peers (Bigler and Liben 2006, 2007; Halpern et al. 2011). Such environmental cues may, explicitly or implicitly, relay a message to students that gender is a fixed attribute of ability (Dweck 2008), which has been shown to have a deleterious impact on performance outcomes (Dar-Nimrod and Heine 2006; Pennington and Heim 2016). From a developmental perspective, research also suggests that single-gender schooling may come at a longer-term cost to successful gender-role socialization and intergroup cooperation once females and males are eventually re-integrated in

ensuing education and workplace settings (Fabes et al. 2015; Halpern et al. 2011; Martin and Fabes 2001).

Due to a number of substantial methodological weaknesses, researchers have argued that studies evaluating the potential efficacy of single-gender schooling need to be interpreted with caution (Halpern et al. 2011; Pahlke et al. 2013, 2014). The most pertinent issue is that many single-gender schools employ selective admissions procedures whereby students are recruited based on their previous ability and socio-economic background (Hayes et al. 2011; Marsh 1989; Signorella et al. 2013). However, many studies do not control for selection effects within their analyses (cf. Pahlke et al. 2014 for a meta-analysis). This greatly undermines the conclusions that can be drawn from research investigating the possible impact that single-gender schooling may have on educational outcomes because students who attend these schools may differ from those attending co-educational schools in important ways (Hayes et al. 2011). Demonstrating the significance of this problem, Pahlke et al. (2014) conducted a meta-analysis and found that studies which did not control for students' previous attainment showed a moderate positive effect of single-gender schooling for mathematics. On the other hand, their findings indicate that studies which controlled for prior achievement tend to show a negligible effect of single-gender classroom settings on attainment levels. They conclude that findings from high quality studies do not support the view that single-gender schooling provides benefits over and above co-educational schooling.

Presenting as an additional issue, research compares typically the effects of the school environment between single-gender and co-educational schools and generalizes these findings across nations (Baker et al. 1995). This creates a number of possible confounds, specifically with regard to the likelihood of differences emerging as a result of variations between school settings and the broader context in which learning takes place (Mael et al. 2005; Shapka 2009). Consequently, it is difficult to determine whether gains in academic attainment are the result of gender-segregation strategies or the product of other educational variables, such as the social and cultural environment in which students are taught (Pahlke et al. 2014).

On a more practical level, the creation of single-gender schools is influenced heavily by the organization of state education and broader economic factors. For example, the number of single-gender schools in the United Kingdom decreased by approximately 80% in the last three decades of the twentieth century because schools received considerable pressure to teach boys and girls jointly to sustain economic viability (Younger and Warrington 2006). As a consequence, it has been argued that this can make it challenging for teachers to tailor instructional strategies to the presumed different learning needs of females and males in certain subjects (Parker and Rennie 2002). For example, research indicates that teachers

are able to spend more time supporting boy's English performance in single-gender classrooms, as well as managing behavior more effectively (Parker and Rennie 2002).

The implementation of single-gender classrooms within co-educational schools therefore presents as a potentially viable option to bolster students' participation and performance. Empirical studies appear to show that single-gender classrooms increase females' long-term participation in counter-stereotypical domains such as science and mathematics (Gillibrand et al. 1999; Rosenthal et al. 2011), and bolster males' English proficiency (Parker and Rennie 2002). However, in their meta-analysis, Pahlke et al. (2014) failed to find a consistent advantage of single-gender classrooms over single-gender schooling, suggesting that selection effects may confound studies. Given the dearth of research in this area (Arnot et al. 1999; Warrington and Younger 2003), it is clear that additional research is required to examine reliably the potential effectiveness of single-gender classroom initiatives implemented within co-educational schools.

Building upon this review, the current research appraises the efficacy of a single-gender classroom initiative on students' academic attainment in a co-educational high school. This intervention was implemented due to a perceived gender-achievement gap in which teachers reported that girls were outperforming boys in the majority of school subjects, except for Mathematics and Science. Overcoming the limitations inherent in previous research, the current study controlled for students' prior attainment (pre-intervention), as well as variables relating to socio-economic status, special education needs (SEN), and native language. It was hypothesized that young women would achieve significantly higher grades in Language subjects (Young-Suk et al. 2015), whereas young men would outperform young women in STEM (Stoet and Geary 2013). Moreover, it was hypothesized that single-gender classrooms would show a positive effect on academic attainment when prior ability was not controlled for, but that these effects would be significantly reduced (if not disappear completely) when accounting for this (Halpern et al. 2011; Pahlke et al. 2013, 2014).

Method

Participants

Data analyses were performed on archived data for 266 students' academic attainment grades, which were obtained throughout their first year of high school (11–12 years of age) in a U.K comprehensive, co-educational school. Of this sample, 123 (46.2%) students were female and 143 (53.8%) were male. A total of 98 students (54 female, 44 male) were placed into single-gender classrooms, with the remaining 168

students (69 female, 99 male) taught in mixed-gender classrooms. Thirty-six percent of students ($n = 96$) were registered as having a diagnosis of Special Educational Needs (SEN) (i.e., moderate learning disabilities, attention deficit hyperactivity disorder, dyslexia, autism, hearing impairment, and dyscalculia). Sixty-six received free school meals (FSM; 24.8%), and eight did not speak English as their native language (EAL; 3.0%). None of these factors differed significantly as a function of classroom type or students' gender (all $p > .05$).

Procedure

The school implemented a single-gender classroom initiative with the aim of bolstering students' academic attainment. A letter was sent to the parents of each student explaining the initiative, and parents provided informed consent (through opt-out) for their children to be placed into a single-gender classroom from the start of secondary education. Across the entire sample analyzed, four parents vetoed the procedure and opted for their children to remain in co-educational classroom settings.

The school followed a specific selection criterion to assign students to single-gender or mixed-gender classrooms. Specifically, the school created an average score for each student, using aggregate predicted grades from primary school in English, mathematics, and science. They then assigned the highest achieving students ($n = 107$) to four classrooms of mixed-gender forms. The next 98 students were then placed into four single-gender forms of middle ability, with two all-male and two all-female classrooms. The remaining students were assigned to middle ability, mixed-gender classrooms. Students remained in either single-gender or mixed-gender classrooms for all school subjects, except for Physical Education in which they were taught in single-gender groups. Irrespective of classroom type (single/mixed-gender), students undertook the same standardized tests at the end of the academic year in the subjects of science, mathematics, information and communications technology (ICT), drama, music, English, and foreign language. Students completed an ongoing assessment in Art which was graded by teacher's professional judgement. Although different teachers taught each school subject, the same teachers taught students in both single-gender and mixed-gender classes in their respective subjects.

Analytic Strategy

Given that the school had not assigned randomly students to single-gender or mixed-gender classes, it was important to control for their prior attainment (Pahlke et al. 2014; Pahlke and Hyde 2016). First, we removed high attaining students (who were all assigned to mixed-gender classrooms,

$n = 107$) from the dataset so that we were left with only middle-attaining students (total $n = 266$; participant section reports this final number, after exclusions). We then computed a difference score by subtracting students' predicted grades (pre-intervention) from their obtained grades (post-intervention). Students' predicted grades were computed in line with their standardized test scores in primary school and were generated by an external organization. Students' obtained grades represent their standardized test scores in their first year of high school, which were graded in accordance with U.K National Curriculum guidelines (The National Curriculum 2010). They received a subject-specific attainment level between 1 and 8, with a higher level indicating better performance. Each of these levels was also split into three ability categories (e.g., Level 4: Lower, Middle, and Upper). For the purpose of statistical analyses, these grades were re-coded from categorical scores to continuous scores on a scale ranging from 1 (Level 2 L) to 21 (Level 8 U; see Table 1).

An average mean difference was computed for STEM subjects (Science, Math, ICT), non-STEM subjects (art, drama, music), and Languages (English, foreign language). This limited the number of analyses conducted and allowed greater control over Type 1 errors compared to analyzing each subject grade separately. Supporting Information File 1 (an online supplement) presents analyses for separate school subjects. Data analysis took the form of a 2 (Gender: male, female) \times 2 (Classroom type: single-gender, mixed-gender) between-participants Analysis of Variance (ANOVA). An Analysis of Covariance (ANCOVA) was also conducted to examine whether receiving free school meals (FSM), English as a native language (EAL), and special education needs (SEN) influenced these findings. An adjusted alpha level of $p < .01$ was utilized to elucidate any main effects and interactions. This decision was guided by the rationale that all p -values are uniformly distributed under the null hypothesis. As such, an alpha level of $p < .01$ provides stronger evidence against the null hypothesis relative to $p < .05$ and therefore provides more convincing findings (Cumming and Calin-Jageman 2017, pp. 130). Positive scores indicate that students' obtained grades were higher than their predicted grades were, whereas negative scores indicate that their obtained scores were lower than predicted.

Results

Languages

When controlling for prior attainment, there was no significant main effect of classroom type, $F(1, 256) = 1.26, p = .263, \eta^2 = .005, 99\% \text{ CI} [-.12, .31]$ (see Table 2a). There was no significant main effect of gender, $F(1, 256) = .61, p = .436, \eta^2 = .002, 99\% \text{ CI} [-.27, .15]$. There was also no significant

Table 1 Subject-specific attainment levels based on national curriculum guidelines, re-coded into ordinal classifications

Levels	Classification/Grouping		
	Lower	Middle	Upper
Level 2	1	2	3
Level 3	4	5	6
Level 4	7	8	9
Level 5	10	11	12
Level 6	13	14	15
Level 7	16	17	18
Level 8	19	20	21

interaction between gender and classroom type, $F(1, 256) = 4.41, p = .037, \eta^2 = .017$.

When prior performance was not controlled for, a main effect of classroom type was found, $F(1, 256) = 58.04, p < .001, \eta^2 = .19$, with students taught in single-gender classrooms ($M = 9.80, SD = 1.42$) appearing to outperform those in mixed-gender classrooms ($M = 7.97, SD = 1.98$), $p < .001, 99\% \text{ CI} [-2.26, -1.11]$. This highlights the confounding influence of selective admissions. Including FSM, EAL and SEN status as covariates did not significantly influence these findings.

STEM-Subjects (Mathematics, Science, ICT)

Controlling for prior attainment, there was no significant main effect of classroom type, $F(1, 258) = .25, p = .617, \eta^2 = .001, 99\% \text{ CI} [-.24, .36]$ (see Table 2b). There was a significant main effect of gender, $F(1, 258) = 7.31, p = .007, \eta^2 = .03$. Simple main effects indicated that young men ($M = -.58, SD = .89$) underperformed relative to their predicted grades compared to young women ($M = -.27, SD = .88$), $p = .007, 99\% \text{ CI} [.01, .61]$. There was no significant interaction between gender and classroom type, $F(1, 258) = .04, p = .850, \eta^2 < .001$.

When prior performance was not controlled for, a main effect of classroom type was found, $F(1, 258) = 76.53, p < .001, \eta^2 = .23$, with students taught in single-gender ($M = 9.45, SD = 1.15$) seemingly outperforming those in mixed-gender classrooms ($M = 7.75, SD = 1.45$), $p < .001, 99\% \text{ CI} [-2.07, -1.12]$. Including FSM, EAL and SEN status as covariates did not significantly influence these findings.

Non-STEM Subjects (art, Drama, Music)

Controlling for prior attainment, there was no significant main effect of classroom type, $F(1, 259) = .058, p = .809, \eta^2 < .001, 99\% \text{ CI} [-.14, .17]$ (see Table 2c). There was no significant main effect of gender, $F(1, 259) = 6.60, p = .011, \eta^2 = .025, 99\% \text{ CI} [-.31, .002]$. However, there was a significant interaction between gender and classroom

Table 2 Descriptive statistics for student's academic attainment (controlling for prior achievement) by gender and classroom type within subject areas

Students' Gender	Subject Areas								
	(a) Languages			(b) STEM			(c) Non-STEM		
	Classroom Type		Gender	Classroom Type		Gender	Classroom Type		Gender
	Single-gender <i>M (SD)</i>	Mixed-gender <i>M (SD)</i>	Main Effect <i>M (SD)</i>	Single-gender <i>M (SD)</i>	Mixed-gender <i>M (SD)</i>	Main Effect <i>M (SD)</i>	Single-gender <i>M (SD)</i>	Mixed-gender <i>M (SD)</i>	Main Effect <i>M (SD)</i>
Young Women	-.36 (.75)	-.10 (.65)	-.21 (.71)	-.32 (.95)	-.24 (.83)	-.27 (.88) _a	-.30 (.36) _a	-.07 (.39) _b	-.17 (.39)
Young Men	-.13 (.22)	-.21 (.68)	-.18 (.57)	-.61 (.72)	-.57 (.96)	-.58 (.89) _b	.07 (.56) _b	-.14 (.50)	-.07 (.53)
Classroom Main Effect	-.26 (.59)	-.16 (.67)		-.45 (.92)	-.44 (.92)		-.14 (.49)	-.11 (.46)	

Languages includes English and foreign languages; STEM includes Science, Mathematics and Information and Communication Technology; Non-STEM includes art, drama and music. Different subscripts comparing means for the main effects of classroom type and gender, as well as for the Classroom type*Gender interaction, indicate a statistically significant difference ($p < .01$)

type, $F(1, 259) = 13.62$, $p < .001$, $\eta^2 = .05$. Simple main effects indicated that young women underperformed relative to their predicted grades in single-gender compared to mixed-gender classrooms, $p = .006$, 99% CI [.02, .45]. However, there was no significant difference between young men in single-gender and mixed-gender classrooms, $p = .02$, 99% CI [- .42, .01]. Furthermore, when taught in single-gender classrooms, young women underperformed relative to their predicted grades compared to young men, who performed in line with their predicted grades, $p < .001$, 99% CI [- .61, - .13]. There was no difference between females and males in mixed-gender classrooms, $p = .36$, 99% CI [- .12, .26]. In sum, young women in single-gender classrooms performed significantly worse in Non-STEM subjects than both their female counterparts in mixed-gender classrooms and young men in single-gender classrooms.

When prior attainment was not controlled, there was a main effect of classroom type, $F(1, 259) = .6077$, $p < .001$, $\eta^2 = .19$, with students taught in single-gender classrooms ($M = 8.10$, $SD = .93$) seemingly outperforming those taught in mixed-gender classrooms ($M = 6.98$, $SD = 1.13$), $p < .001$, 99% CI [- 1.37, - .68]. Including FSM, EAL and SEN status did not influence these findings.

Discussion

The current study evaluated the efficacy of a single-gender classroom initiative implemented in a co-educational school in the United Kingdom. Such research is able to control for many extraneous environmental variables to a greater extent than research examining the impact of single-gender schooling in different contexts. Overcoming methodological issues within this literature, the current study also controlled for

selection effects by accounting for students' previous attainment grades, which were calculated prior to the intervention in line with national curriculum guidelines. In summary, the findings indicate that young women and young men's academic attainment in STEM-related (Mathematics, Science, ICT) and Language subjects (English, foreign language) did not differ significantly as a function of classroom type. These results are in line with recent meta-analytic findings (Pahlke et al. 2014), which reveal limited evidence for the effectiveness of single-gender classrooms on achievement when controlling for prior achievement.

Findings also indicate that young women underperformed relative to their predicted grades in Non-STEM subjects when they were taught in single-gender compared to those taught in mixed-gender classrooms. When taught in single-gender classrooms, young women underperformed relative to their predicted grades in Non-STEM subjects compared to young men, who performed in line with their predicted grades. This finding contrasts with previous research suggesting that female students may benefit more than males do when taught in single-gender compared to mixed-gender classrooms (Alon and Gelbgiser 2011; Lee and Bryk 1986; Mullholland et al. 2004).

In order to understand this finding, it may important to reflect on the nature of the school subject or pedagogic context. Specifically, subjects such as art, drama, and music are more open-ended by nature and often involve more peer observation and interaction than do STEM-related subjects. As a consequence, performance is perhaps more visible in these subjects, and females may respond differently to performance appraisal from other ingroup (i.e., a class of other young women) relative to outgroup others (i.e., a mixed-gender class). Furthermore, young women may be more self-aware or conscious when participating in performance-based subjects in

single-gender groups. In support of this speculation, research suggests that gender differences in self-concept emerge in adolescence, with young women becoming more self-conscious and aware of criticism than young men are (Rankin et al. 2004; Rosenberg and Simmons 1975), which may help to explain why we found an interaction between gender and the classroom intervention.

When the current results are analyzed without accounting for pre-existing ability, the single-gender classroom initiative appears to be highly efficacious. Such findings are simply a product of the school employing a selective admissions process to assign students to single-gender and mixed-gender classrooms. Our research therefore highlights the importance of controlling for selection effects in the evaluation of single-gender classroom initiatives. We argue that studies which do not control for students' prior ability may tell us little about the effectiveness of such interventions.

An additional unexpected finding was that young men taught in both single- and mixed-gender classrooms appeared to underperform relative to their predicted grades in STEM-related subjects compared to young women. Although we take caution in inferring explanations from these findings, they may be interpreted in numerous ways. For example, this finding suggests that males' predicted grades for STEM-related subjects might be overinflated relative to females, or that males might be underperforming (relative to their predicted grades) when they undertake standardized tests in exam settings. This suggestion appears to be supported because young women achieve higher predicted and actual grades compared to young men for all school subjects when analyzing predicted and obtained grades separately. However, when a difference score is calculated, males' grades in STEM-related subjects appear to be over predicted. This finding has important implications because students are informed typically about their predicted grades in order for teachers to set goals and encourage students to achieve these grades. However, it is plausible that, if predicted grades are set too high, this might have a paradoxical effect on motivation and subsequent exam performance because children feel that their predicted grades are unobtainable. We urge additional research to explore the factors that may explain this pattern of results and to elucidate whether these findings emerge in other educational settings. Such findings, if corroborated, could have major implications for policy and practice.

Limitations and Future Research Directions

The current research and many previous studies focus on the impact of gender-segregated educational initiatives on academic attainment. As such, there is a lack of research which examines other related psychosocial outcomes that may be influenced by single-gender schooling or classroom interventions. Moreover, studies that do examine additional factors

have presented somewhat mixed findings. Although some research indicates that single-gender classrooms may lessen the salience of gender-related stereotypes and performance expectations to bolster students' performance (Elizaga and Markman 2008; Huguet and Régner 2007; Inzlicht and Ben-Zeev 2000, 2003; Picho and Stephens 2012), other research suggests that gender saliency in single-gender classrooms may exacerbate intergroup biases (Fabes et al. 2015; Halpern et al. 2011; Martin and Fabes 2001). Accordingly, we recommend that future research examines how single-gender educational strategies may impact psychological factors such as mindset, competence beliefs, academic self-efficacy, self-esteem, gender stereotyping, and intergroup attitudes, in addition to academic attainment. The challenges that may arise when students subsequently rejoin the other gender in post-school settings also warrant further consideration.

Practice Implications

Our research proffers both pragmatic and methodological implications. First, we demonstrate how schools typically assign students to educational interventions using selective admissions criteria. Practically, it may be difficult for schools to assign students randomly to single- or mixed-gender classrooms because they are taught in ability settings in UK schools (e.g., grouping students into lower, middle and upper ability groups). In such cases, it is recommended that researchers account for prior achievement in order to elucidate reliably whether single-gender environments represent a practical strategy to bolster academic attainment over co-educational schooling.

Second, in the present case, the school had implemented a single-gender educational intervention to alleviate a perceived gender-achievement gap in scholastic achievement. Our research allowed us to inform the school whether there were indeed gender differences (separate analyses of obtained grades indicated that females were outperforming males in all school subjects), as well as whether the single-gender classroom initiative was successful in alleviating these. Given the findings, this evaluation enables the school to examine additional strategies, other than single-gender classroom instruction, that may be more effective in lessening achievement gaps. It also allows them to assess critically whether or not to continue this single-gender classroom initiative for students entering high school in the future.

Conclusion

Our research controlled for students' prior attainment to evaluate the effectiveness of a single-gender classroom initiative implemented in a co-educational, comprehensive UK school.

In summary, findings indicate that young women and men did not appear to benefit from being taught in single-gender relative to mixed-gender classrooms in Language and STEM-related subjects. Moreover, the single-gender intervention had a seemingly negative impact on young women taught in Non-STEM subjects, who underperformed compared to those taught in mixed-gender classrooms. When prior ability was not controlled for, the intervention appears to be highly efficacious, highlighting the confounding influence of selection effects. These findings therefore demonstrate how the observed advantages of single-gender educational initiatives are reduced greatly when accounting for students previous scholastic performance. They also provide empirical support for the notion that much of the reported success of gender-segregated education may be attributable to selection effects (Hayes et al. 2011; Signorella et al. 2013), with this methodological issue distorting the interpretations of research in this area.

Compliance with Ethical Standards The submitted research was ethically approved by both Edge Hill University (UK) and Lancaster University (UK) where the authors hold academic positions. We can confirm that the manuscript adheres to the following guidelines set out by *Sex Roles*:

- The manuscript has not been submitted to more than one journal for simultaneous consideration.
- The manuscript has not been published previously (partly or in full), unless the new work concerns an expansion of previous work (please provide transparency on the re-use of material to avoid the hint of text-recycling (“self-plagiarism”).
- A single study is not split up into several parts to increase the quantity of submissions and submitted to various journals or to one journal over time (e.g. “salami-publishing”).
- No data have been fabricated or manipulated (including images) to support your conclusions.
- No data, text, or theories by others are presented as if they were the author’s own (“plagiarism”). Proper acknowledgements to other works must be given (this includes material that is closely copied (near verbatim), summarized and/or paraphrased), quotation marks are used for verbatim copying of material, and permissions are secured for material that is copyrighted.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Alon, S., & Gelbgiser, D. (2011). The female advantage in college academic achievements and horizontal sex segregation. *Social Science Research, 40*, 107–119. doi:10.1016/j.ssresearch.2010.06.007.
- Amot, G., David, M., & Weiner, G. (1999). *Narrowing the gender gap*. Cambridge: Polity Press.
- Baker, D. P., Riordan, C., & Schaub, M. (1995). The effects of sex-grouped schooling on achievement: The role of national context. *Comparative Education Review, 39*, 468–482. doi:10.1086/447341.
- Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. *Advances in Child Development and Behavior, 34*, 39–89. doi:10.1016/S0065-2407(06)80004-2.
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children’s stereotyping and prejudice. *Current Directions in Psychological Science, 16*, 162–166. doi:10.1111/j.1467-8721.2007.00496.x.
- Bigler, R., & Signorella, M. L. (2011). Single-gender education: New perspectives and evidence on a continuing controversy. *Sex Roles, 65*, 659–669. doi:10.1007/s11199-011-0046-x.
- Bleeker, M. M., & Jacobs, J. E. (2004). Achievement in math and science: Do mothers’ beliefs matter 12 years later? *Journal of Educational Psychology, 96*, 97–109. doi:10.1037/0022-0663.96.1.97.
- Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions in Psychological Science, 19*, 275–279. doi:10.1177/0963721410383241.
- Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women’s underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*, 218–261. doi:10.1037/a0014412.
- Cumming, G., & Calin-Jageman, R. (2017). *Introduction to the new statistics: Estimation, open science, & beyond*. New York: Routledge.
- Dai, D. Y. (2001). A comparison of gender differences in academic self-concept and motivation between high-ability and average Chinese adolescents. *Journal of Advanced Academics, 13*, 22–32. doi:10.4219/jsge-2001-361.
- Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women’s math performance. *Science, 314*, 435. doi:10.1126/science.1131100.
- Dweck, C. (2008). *Mindset: How you can fulfill your potential*. New York: Random House.
- Eisenkopf, G., Hessami, Z., Fischbacher, U., & Ursprung, H. W. (2015). Academic performance and single-gender schooling: Evidence from a natural experiment in Switzerland. *Journal of Economic Behavior & Organization, 115*, 123–143. doi:10.1016/j.jebo.2014.08.004.
- Elizaga, R. A., & Markman, K. D. (2008). Peers and performance: How in-group and out-group comparisons moderate stereotype threat effects. *Current Psychology, 27*, 290–300. doi:10.1007/s12144-008-9041-y.
- Else-Quest, N. M., & Peterca, O. (2015). Academic attitudes and achievement in students of urban public single-gender and mixed-gender high schools. *American Educational Research Journal, 52*, 693–718. doi:10.3102/0002831215591660.
- 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*, 103–127. doi:10.1037/a0018053.
- Fabes, R. A., Lynn-Martin, C., Hanish, L. D., Galligan, K., & Pahlke, E. (2015). Gender-segregated schooling: A problem disguised as a solution. *Educational Policy, 29*, 431–447. doi:10.1177/0895904813492382.
- Gillibrand, E., Robinson, P., Brawn, R., & Osborn, A. (1999). Girls’ participation in physics in single sex classes in mixed schools in relation to confidence and achievement. *International Journal of Science Education, 21*, 349–362. doi:10.1080/095006999290589.
- Halpern, D. F., Eliot, L., Bigler, R. S., Fabes, R. A., Hanish, L. D., et al. (2011). The pseudoscience of single-gender schooling. *Science, 23*, 1706–1707. doi:10.1126/science.1205031.
- Hayes, A. R., Pahlke, E., & Bigler, R. (2011). The efficacy of single-gender education: Testing for selection and peer quality effects. *Sex Roles, 65*, 693–703. doi:10.1007/s11199-010-9903-2.
- Hoffman, L. (2002). Promoting girls’ interest and achievement in physics classes for beginners. *Learning & Instruction, 12*, 447–465. doi:10.1016/S0959-4752(01)00010-X.

- Huguet, P., & Régner, I. (2007). Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology, 99*, 545–560. doi:10.1037/0022-0663.99.3.545.
- Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*, 365–371. doi:10.1111/1467-9280.00272.
- Inzlicht, M., & Ben-Zeev, T. (2003). Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Psychology, 95*, 796–805. doi:10.1037/0022-0663.95.4.796.
- Jackson, C., & Smith, I. D. (2000). Poles apart? An exploration of single-gender and mixed-gender educational environments in Australia and England. *Educational Studies, 26*, 409–422. doi:10.1080/03055690020003610.
- Kimball, M. M. (1989). A new perspective on women's math achievement. *Psychological Bulletin, 105*, 198–214. doi:10.1037/0033-2909.105.2.198.
- Lee, V. E., & Bryk, A. S. (1986). Effects of single-gender secondary schools on student achievement and attitudes. *Journal of Educational Psychology, 78*, 381–395. doi:10.1037/0022-0663.78.5.381.
- Lee, V. E., & Marks, H. M. (1990). Sustained effects of the single-gender secondary school experience on attitudes, behaviors, and values in college. *Journal of Educational Psychology, 82*, 578–592. doi:10.1037/0022-0663.82.3.578.
- Mael, F., Alonso, A., Gibson, D., Rogers, K., & Smith, M. (2005). *Single-gender versus coeducational schooling: A systematic review* (No. 2005-01). Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. Retrieved from <https://www2.ed.gov/rschstat/eval/other/single-gender/single-gender.pdf>.
- Marsh, H. (1989). The effects of attending single-gender and Catholic coeducational high schools on achievement, attitudes and behaviours and on sex differences. *Journal of Educational Psychology, 81*, 70–85. doi:10.1037/0022-0663.81.1.70.
- Martin, C. L., & Fabes, R. A. (2001). The stability and consequences of young children's same-sex peer interactions. *Developmental Psychology, 37*, 431–446. doi:10.1037/0012-1649.37.3.431.
- Miller, D. I., Eagley, A. H., & Linn, M. C. (2015). Women's representation in science predicts national gender-science stereotypes: Evidence from 66 nations. *Journal of Educational Psychology, 107*, 631–644. doi:10.1037/edu0000005.
- Mullholland, J., Hansen, P., & Kaminski, E. (2004). Do single-gender classrooms in coeducational settings address boys' underachievement? An Australian study. *Educational Studies, 30*, 19–32. doi:10.1080/0305569032000159714.
- Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives, 24*, 129–144. doi:10.1257/jep.24.2.129.
- Pahlke, E., & Hyde, J. S. (2016). The debate over single-gender schooling. *Child Development Perspectives, 10*, 81–86. doi:10.1111/cdep.12167.
- Pahlke, E., Hyde, J. S., & Mertz, J. E. (2013). The effects of single-gender compared with coeducational schooling on mathematics and science achievement: Data from Korea. *Journal of Educational Psychology, 105*, 444–452. doi:10.1037/a0031857.
- Pahlke, E., Hyde, J. S., & Allison, C. M. (2014). The effects of single-gender compared with coeducational schooling on students' performance and attitudes: A meta-analysis. *Psychological Bulletin, 140*, 1042–1072. doi:10.1037/a0035740.
- 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*, 366–381. doi:10.1006/ceps.2000.1069.
- Park, H., Behrman, J. R., & Choi, J. (2013). Causal effects of single-gender schools on college entrance exams and college attendance: Random assignment in Seoul high schools. *Demography, 50*, 447–469. doi:10.1007/s13524-012-0157-1.
- Parker, L. H., & Rennie, L. J. (2002). Teachers' implementation of gender-inclusive instructional strategies in single-gender and mixed-gender classrooms. *International Journal of Science Education, 24*, 881–897. doi:10.1080/09500690110078860.
- Pennington, C. R., & Heim, D. (2016). Creating a critical mass eliminates the effects of stereotype threat on women's mathematical performance. *British Journal of Educational Psychology, 86*, 353–368. doi:10.1111/bjep.12110.
- Picho, K., & Stephens, J. M. (2012). Culture, context and stereotype threat: A comparative analysis of young Ugandan women in coed and single-gender schools. *The Journal of Educational Research, 105*, 52–63. doi:10.1080/00220671.2010.517576.
- Rankin, J. L., Lane, D. J., Gibbons, F. X., & Gerrard, M. (2004). Adolescent self-consciousness: Longitudinal age changes and gender differences in two cohorts. *Journal of Research on Adolescence, 14*, 1–21.
- Reilly, D., Neumann, D. L., & Andrews, G. (2015). Sex differences in mathematics and science achievement: A meta-analysis of National Assessment of educational progress assessments. *Journal of Educational Psychology, 107*, 645–662. doi:10.1037/edu0000012.
- Rosenberg, F. L., & Simmons, R. G. (1975). Sex differences in the self-concept in adolescence. *Sex Roles, 1*, 147–159.
- Rosenthal, L., London, B., Levy, S. R., & Lobel, M. (2011). The roles of perceived identity compatibility and social support for women in a single-gender STEM program at a co-educational university. *Sex Roles, 65*, 725–738. doi:10.1007/s11199-011-9945-0.
- Schneider, F. W., & Coutts, L. M. (1982). The high school environment: A comparison of coeducational and single-gender schools. *Journal of Educational Psychology, 74*, 898–906. doi:10.1037/0022-0663.74.6.898.
- Shapka, J. D. (2009). Trajectories of math achievement and perceived math competence over high school and postsecondary education: Effects on an all-girl curriculum in high school. *Educational Research and Evaluation: An International Journal on Theory and Practice, 15*, 527–541. doi:10.1080/13803610903354775.
- Shapka, J. D., & Keating, D. P. (2003). Effects of a girls-only curriculum during adolescence: Performance, persistence, and engagement in mathematics and science. *Educational Research Journal, 40*, 929–960. doi:10.3102/00028312040004929.
- Signorella, M. L., & Bigler, R. S. (2013). Single-gender schooling: Bridging science and school boards in educational policy. *Sex Roles, 69*(7–8), 349–355. doi:10.1007/s11199-013-0313-0.
- Signorella, M. L., Hayes, A. R., & Li, Y. (2013). A meta-analytic critique of Mael et al.'s (2005) review of single sex schooling. *Sex Roles, 69*, 423–441. doi:10.1007/s11199-013-0288-x.
- Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within- and across-nation assessment of 10 years of PISA data. *PLoS One, 8*, e57988. doi:10.1371/journal.pone.0057988.
- Sullivan, A. (2009). Academic self-concept, gender, and single-gender schooling. *British Educational Research Journal, 35*, 259–288. doi:10.1080/01411920802042960.
- Sullivan, A., Joshi, H., & Leonard, D. (2010). Single-gender schooling and academic attainment at school and through the lifecourse. *American Educational Research Journal, 47*, 6–36. doi:10.3102/0002831209350106.
- The National Curriculum. (2010). *The National Curriculum: Level descriptions for subjects*. Coventry: The Qualifications and Curriculum Authority (QCA). Retrieved from <http://www.empiribox.org/wp-content/uploads/2014/08/Level-Descriptors-for-Science.pdf>.

- Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics*, *50*, 49–62. doi:10.1023/A:1020518104346.
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*, 1174–1204. doi:10.1037/a0036620.
- Wai, J., Cacchio, M., Putallaz, M., & Makel, M. C. (2010). Sex differences in the right tail of cognitive abilities: A 30 year examination. *Intelligence*, *38*, 412–423. doi:10.1016/j.intell.2010.04.006.
- Warrington, M., & Younger, M. (2003). 'we decided to give it a twirl': Single-gender teaching in English comprehensive schools. *Gender and Education*, *15*, 339–350. doi:10.1080/09540250310001610553.
- Wood, D., Kurtz-Costes, B., Rowley, S. J., & Okeke-Adeyanju, N. (2010). Mothers' academic gender stereotypes and education-related beliefs about sons and daughters in African American families. *Journal of Educational Psychology*, *102*, 521–530. doi:10.1037/a0018481.
- Younger, M. R., & Warrington, M. (2006). Would Harry and Hermione have done better in single-gender classes? A review of single-gender teaching in coeducational secondary schools in the United Kingdom. *American Educational Research Journal*, *43*, 579–620. doi:10.1302/00028312043004579.
- Young-Suk, K., Al Otaiba, S., Wanzek, J., & Brandy, G. (2015). Toward and understanding of dimensions, predictors, and the gender gap in written composition. *Journal of Educational Psychology*, *107*, 79–95. doi:10.1037/a0037210.