Routledge
Taylor & Francis Group

Research Papers in Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rred20

Are single-sex schools more effective than the coed ones? The effect of single-sex schooling on achievement among female adolescents in Catholic schools

Maciej Koniewski & Anna Hawrot

To cite this article: Maciej Koniewski & Anna Hawrot (2021): Are single-sex schools more effective than the coed ones? The effect of single-sex schooling on achievement among female adolescents in Catholic schools, Research Papers in Education

To link to this article: https://doi.org/10.1080/02671522.2021.1886318

RESEARCH PAPERS IN EDUCATION
https://doi.org/10.1080/02671522.2021.1886318

Routledge
Taylor & Francis Group

Are single-sex schools more effective than the coed ones? The effect of single-sex schooling on achievement among female adolescents in Catholic schools

Maciej Koniewski [a] and Anna Hawrot [b]

[a]Department of Philosophy, Institute of Sociology, Jagiellonian University, Cracow, Poland; [b]Leibniz Institute for Educational Trajectories, Bamberg, Germany

ABSTRACT

This study inquired into the effect of single-sex schooling on achievement of female adolescents. It used national examination data of 4,787 Polish female students (10 cohorts) attending single-sex and co-educational Catholic schools. We tested two sets of partially contradictory predictions derived from two different theoretical models explaining how and why achievement of female students in the two types of school may differ. The results of cross-classified random-effects regression models showed that after controlling for initial student and school differences female adolescents attending all-girls schools scored higher on the lower-secondary school exam in science and the humanities in comparison to those who attended co-educational schools. However, the difference was statistically significant only for science. The results were fully consistent with neither of the two adopted theoretical models, although provided more support for the one drawing upon peer effects. Although the effect of 17% of the exam scores standard deviation could be considered small, it appeared in the results of a high-stakes exam. Since the examination results were the main criterion for admission to the next-stage school, attending an all-girls school might significantly affect future educational career and job opportunities of young women.

ARTICLE HISTORY
Received 29 May 2020
Accepted 12 January 2021

KEYWORDS
academic achievement; same-sex education; religious education; junior high school students; female adolescents

Introduction

Single-sex schooling, as an alternative to co-educational (coed) one, has been a topic of heated debate for several decades. Its presence in public education systems has been discussed as a matter of school effectiveness, free educational choice, and gender equality of educational opportunities (e.g., Halpern et al. 2011; Liben 2015; Pahlke and Hyde 2016; Wiseman 2008).

According to its proponents, single-sex schools increase student motivation and academic achievement by tailoring instruction and pedagogical practice to the unique developmental needs of each gender, differences in brain functioning, learning styles, aptitudes, interests or participatory styles in the classroom. Single-sex environment is

CONTACT Maciej Koniewski maciej.koniewski@uj.edu.pl Department of Philosophy, Institute of Sociology, Jagiellonian University, Ul. Grodzka 52, 31-044 Cracow, Poland
Supplemental data for this article can be accessed here.

supposed to reduce gender-stereotyping, sexism in teacher-students interactions, prevent distraction caused in adolescence by the other sex, and protect female students from sexual harassment (e.g., Riordan 2015; see Bigler, Hayes, and Liben 2014; Liben 2015 for reviews).

According to the opponents, single-sex education is based on pseudo-science (Halpern et al. 2011), its promises lack sufficient empirical support, and research on the effect is often flawed by not accounting for selection bias (Bigler, Hayes, and Liben 2014; Signorella, Hayes, and Li 2013). Moreover, it yields multiple individual and social consequences, including widening education inequalities, strengthening gender stereotypes, limiting opportunities to learn cooperation with the other sex, as well as diverting time and money from more productive educational reforms (e.g., Halpern et al. 2011; Liben 2015). One's position in the debate is often influenced by anecdotal evidence and ideological assumptions (Liben 2015), which limits focus on robust empirical evidence.

The debate has resulted in numerous studies on the advantages and disadvantages of single-sex education versus coeducation, which however have yielded inconsistent results. Some did not find any effect (e.g., LePore and Warren 1997; Marsh 1989), whereas others indicated mixed effects, with one of the sexes benefitting and the other being disadvantaged (e.g., Daly and Defty 2004; Van de Gaer et al. 2004), or reported advantages for at least one of the sexes (e.g., Jackson 2016; Park, Behrman, and Choi 2018). The positive effects of single-sex schools were more often observed for young males, than for females (Hahn and Wang 2019; Jackson 2016) and if observed for both sexes, they were usually stronger for males (e.g., Park, Behrman, and Choi 2018).

This study inquiries into the effect of single-sex schooling among female adolescents in Poland. Unlike many past studies (see Pahlke, Hyde, and Allison 2014 for a review), we used a large dataset of 4,787 female students in 10 cohorts. The study contributes to the literature in the following ways. First, it provides information on the effect of single-sex schooling in the part of the world where it has not been studied extensively (see Eisenkopf et al. 2015; Garcia-Gracia & Donoso Vázquez, 2016; Kessels and Hannover 2008 for the only studies in Continental Europe). Second, it uses the results of a national high-stakes exam as a measure of academic achievement instead of school grades (e.g., Eisenkopf et al. 2015) or binary variables such as passing an exam (e.g., Jackson 2012).

Third, by restricting the analyses to all-girls and coed schools run by Roman Catholic sisterhoods we ruled out student- and school-level confounders, which were difficult to control in previous non-experimental studies. All-girls and coed schools in our study were similar. They attracted similar students, had similar recruitment criteria and provided comparable learning environments in terms of basic school rules, promoted values, educational practices, and teaching staff. Moreover, schools in our sample emphasised religious formation rather than merely academic achievement. The similarities they shared allowed an implicit control of student- and school-related covariates, which was additionally strengthened through statistical control of various student- and school-level characteristics, including prior academic achievement. In contrast, most of the previous studies on single-sex education have been run in the U.S., UK, Australia, Canada, or New Zealand (see Pahlke, Hyde, and Allison 2014), where single sex-schools are strongly academically oriented, hence highly selective (e.g., Mayfield School, Eton College in the UK), or provide a pro-academic offer for underprivileged communities (e.g., Urban Prep Academies, Eagle Academy for Young Men in the U.S.). As a result

their student body systematically differed from the student body in the general population in coed schools, to which they were compared (e.g., Hayes, Pahlke, and Bigler 2011; Jackson 2012; Signorella, Hayes, and Li 2013). Hence, robust comparisons of single-sex and coed schools students were virtually impossible in previous non-experimental studies. Past studies have often not addressed the (self-)selection effects and between-school differences sufficiently (see Mael et al. 2005; Pahlke, Hyde, and Allison 2014; Signorella, Hayes, and Li 2013), which led to biased estimates of the effects of single-sex schooling.

Moreover, this study adds up-to-date information on the effect of single-sex education to a limited body of literature on Catholic schools, for which only few studies that compare single-sex and coed Catholic schools are available (LePore and Warren 1997; Marsh 1989, 1991; Nagengast, Marsh, and Hau 2013).

Theoretical framework

Although many explanations why single-sex schools may be more effective than coed schools have been suggested (see Bigler, Hayes, and Liben 2014; Mael et al. 2005 for a review), researchers rarely put them to the test, focusing solely on the effect size instead. For the purpose of this study we chose two competing theoretical explanations. They provide partially contradictory predictions, and therefore offer an opportunity to verify which one gains more empirical support. Meanwhile, although we are aware of multiple other explanations, discussing them all is beyond the scope of this paper.

Pahlke, Hyde, and Allison (2014) suggested a theoretically sound explanation based on the expectancy-value theory (Eccles 2009). According to model, the behavioural choices people make depend on (a) their expectations for success and (b) perceived value of options, which they consider available. Both elements are shaped by a variety of personal and contextual factors, for example self-schema, motivations, aptitudes, past experience of success in a given domain, behaviour of key socialisers, or stereotypes and social-role systems embedded in the culture. The process is ongoing since current choices inform future development of views on oneself.

As noted by Pahlke and colleagues (2014), school and classroom environments also shape activity choices. The sole sex segregation in single-sex schools may highlight gender segregation of jobs (Pahlke, Hyde, and Allison 2014), leading to lower expectation of success and lower subjective value of subjects stereotypically associated with masculinity. This may result in less time and effort put by young women into studying them, and in poorer performance compared to young women in coed schools. The opposite effect may occur in case of humanities.

A partially different set of hypotheses can be derived from research on peer effects in the classroom. Although studies in this strand are rarely framed in any broader model of school learning or effectiveness, for the purpose of this study we placed them in the input-process-output model (Scheerens 2004). According to this model, various school- and classroom-level processes, for example school atmosphere or teaching quality, shape student outcomes. Those processes are affected by the context in which the school operates (e.g., educational policy or location) and the available inputs (e.g., teacher qualification or funding). However, the processes on the classroom level, as proximal to learning, are considered key for student outcomes (Seidel and Shavelson 2007).

How may peer effects in the classroom explain differences in achievement of female adolescents attending single-sex and coed schools? First, the higher share of female classmates may improve outcomes because girls and young women are less disruptive than their male counterparts (Bigler, Hayes, and Liben 2014; Hoxby 2000; Jackson 2016; Lavy and Schlosser 2011). Meanwhile, orderly and disruption-free classroom environment is a prerequisite of effective instruction and learning (e.g., Praetorius et al. 2018). Therefore, in all-girls classrooms less time may be wasted on disciplinary issues and more spent on learning activities, which translates to higher achievement in all subjects.

Second, students may achieve better results thanks to improved cooperative learning behaviour and communication with the same-sex peers (Jackson 2016; Lu and Anderson 2015). These in turn may promote feeling of relatedness among students that is necessary for building motivation (Ryan and Deci 2000), as well as create an environment in which teachers can more easily support the experience of social relatedness, which is an important element of high-quality instruction (Praetorius et al. 2018).

How large is the effect for females?

Despite the abundance of research investigating the impact of single-sex schooling on various outcomes, methodologically robust studies are sparse. Moreover, empirical evidence on the effect of single-sex schooling on female adolescents' achievement has been mixed, with some studies providing evidence for the effect and others not. Figure 1 presents a summary of the results obtained in one meta-analysis, which summed up mostly studies that used regression adjustment strategy to estimate the effect of single-sex education (Pahlke, Hyde, and Allison 2014) and several experimental studies that used large or multi-cohort data sets, high-quality measures of achievement, i.e. national

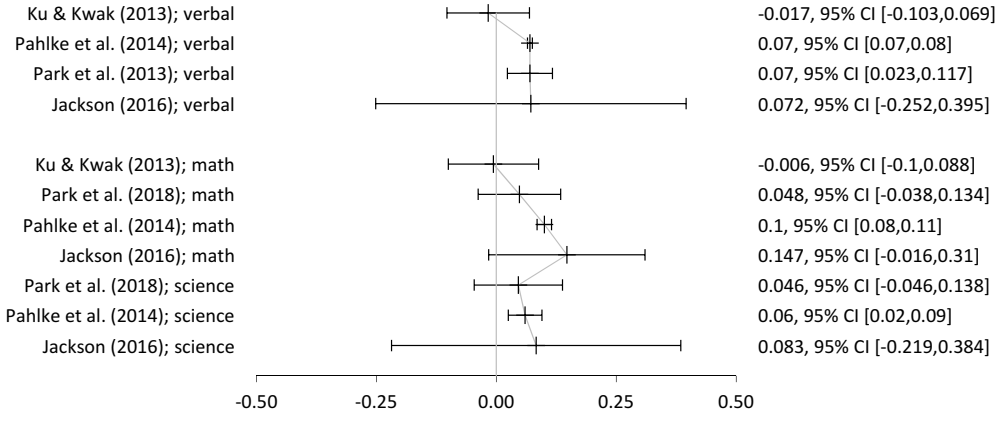

Figure 1. Selected results of robust studies on the effect of single-sex schooling on academic achievement among young women. If both overall and effect with controls were reported in one study, we presented only the effect with controlled variables. Verbal skills in mother tongue: Korean in Ku and Kwak (2013) and Park, Behrman, and Choi (2013); Spanish in Jackson (2016); various in Pahlke, Hyde, and Allison (2014). The estimates reported by Pahlke, Hyde, and Allison (2014) are Cohen's *d* mean effects from studies with random assignment or controls for selection effects weighted by sample sizes. The estimates from the other studies are differences between single-sex and coed female students achievement expressed in standard deviations as a unit.

standardised high-stakes exams, and accounted for clustering of students in districts, cohorts, schools (Jackson 2016; Ku and Kwak 2013; Park, Behrman, and Choi 2013, 2018). Some of them additionally controlled a set of important school-level characteristics (Ku and Kwak 2013; Park, Behrman, and Choi 2013). These effects lean towards positive values and fit the range of −0.219 to 0.384 with the unweighted mean of 0.068 in case of maths and science and −0.252 to 0.395 with the unweighted mean of 0.049 in case of verbal performance. Although, positive effects were reported in all but one study (Ku and Kwak 2013), only a few studies were able to prove the effect's statistical significance. One of them is the meta-analysis by Pahlke, Hyde, and Allison (2014), which included studies inquiring into the effects of two forms of single-sex education, i.e. single-sex schools and single-sex classrooms, showed a small positive effect of single-sex education among female students on verbal, science, and maths performance.

Researches run in Seoul that exploited a unique feature of the city's education system – the random assignment of students to single-sex and coed high schools, give a less clear picture. Park, Behrman, and Choi (2013) have demonstrated that while controlling for the set of school-level variables female students from single-sex schools outperformed their counterparts from coed schools in Korean and English by 0.070 and 0.072 of standard deviation, respectively. Sohn (2016) found the corresponding overall effect estimates (without additional controls) by 17% and 37% higher, while analysing the extended sample of 2002–2004 and 2009–2012 cohorts. However, Ku and Kwak (2013) did not confirm the positive overall effect on Korean and on English in 2001–2009 cohorts. Similarly, Ku and Kwak (2013) reported weak and not significant effect in maths. Also, nonsignificant effect in maths resulted from analyses by Park, Behrman, and Choi (2018) for female students in cohorts 2004–2011. It is notable to say that when the outcome variable was not a high- but a low-stakes test, no effects of attending private and public single-sex schools (separate analyses) on Korean, English, and maths were observed among female students (Hahn and Wang 2019). In sum, although the lottery-based admissions to high schools in Seoul offer a great opportunity to study the effects of single-sex schooling, discrepancies occurred depending on the cohorts studied, outcome and set of covariates used, and a method of data analysis applied.

The single-sex schooling effects have been studied in robust manners also outside South Korea. For example, in Trinidad and Tobago Jackson (2016) exploited the conversion of low-performing secondary schools from coed to single-sex, while holding other school inputs constant. He found no effects on behavioural and academic outcomes, including test scores among female adolescents. In Switzerland, a positive effect of learning in single-sex classes on maths grades was reported, which additionally increased if a male teacher taught the class. No significant effects were found for language skills (Eisenkopf et al. 2015). The authors benefited from a random assignment to all-girls and coed classes.

Quasi-experimental studies, which did not include random assignment to single-sex environments, but instead used various statistical techniques to address the (self-) selection bias, have not found any effect of attending single-sex schools by female adolescents on their school achievement (Jackson 2012; Marsh 1989, 1991; Nagengast, Marsh, and Hau 2013).

The estimates reported by (quasi-)experiments are considered free from the (self-) selection bias. However, it is not clear if and to what extent the observed differences in

student performance between single-sex and coed schools were a result of differences in school-level factors (e.g., teaching quality) instead of their sex composition. Hence, additional control of school-level factors relevant to student achievement is a must while investigating the single-sex schooling effect. According to the authors of various systematic reviews and meta-analyses (Mael et al. 2005; Pahlke, Hyde, and Allison 2014; Signorella, Hayes, and Li 2013), past studies using the regression-based adjustment strategy to control for confounders have often failed to deliver unbiased estimates of the single-sex schooling effect. Hence, in this short literature review we focused on (quasi-)experiments to report the most robust evidence on single-sex schooling effects. However, the regression-based adjustment is not doomed to failure. It can effectively approximate findings from randomised experiments (Pohl et al. 2009; Shadish, Clark, and Steiner 2008), provided that a set of covariates that are well measured and plausibly related to both the selection process and the outcome measures are included, with prior student achievement being indispensable. We believe this study meets this requirement. Additionally, by comparing Catholic all-girls and Catholic coed schools we were able to implicitly control for other not measured possible confounders at both school- and student-level.

Polish school system

During the period covered in this study the general education system in Poland included a six-year primary school (PS; ISCED level 1), followed by a three-year lower-secondary school (LSS; ISCED level 2) and a two- up to four-year upper-secondary school (ISCED level 3). Primary and lower-secondary schools followed a general curriculum, and the first selection threshold to different education tracks occurred between lower- and upper-secondary levels, relatively late compared to other European education systems (e.g., in Germany, Austria). At the end of PS and at the end of LSS students sat compulsory nationwide exams. The LSS one was high stake as its results were decisive for admission to an upper-secondary school. Between-school variance of its results oscillated around 20% (e.g. Dolata 2013).

The present study

This study inquired into the effect of single-sex schooling on academic achievement in Poland. It used data of 10 cohorts of all-girls and coed LSS female adolescents run by Roman Catholic sisterhoods and tested two partially contradictory sets of hypotheses derived from two theoretical explanations why single-sex and coed schools may differ in effectiveness. Based on the expectancy-value model (Eccles 2009), we expected female adolescents in single-sex schools to have higher school achievement in the humanities and lower in science compared to female adolescents attending coed schools. However, based on the research on peer effects framed in the input-process-output model of school effectiveness (e.g., Scheerens 2004), we expected female students in single-sex schools to outperform the ones from coed schools in science and the humanities.

The limitations of the data at our disposal prevented us from testing the mechanisms of the above-mentioned relationships, because the national examination databases we used contain only basic information about students and schools. Thus, we tested only

their predictions. However, we consider this paper as a first step that will allow us to answer the question of whether there is an effect of single-sex schools on academic achievement among Polish students. The existence of this effect is a necessary prerequisite for further tests of its mechanisms.

Method

Sample

This study used a nationwide secondary dataset containing full national examination results, which was developed and maintained by the Educational Research Institute with collaboration of Regional Examination Boards (zpd.ibe.edu.pl). We identified in this dataset all coed and all-girls LSSs run by Roman Catholic sisterhoods from which students graduated between 2007 and 2016 by searching for various key-words in the schools' names (e.g., the word 'sisterhood'). Facing the lack of a registry of single-sex schools in Poland, we identified all-girls schools by averaging information about student sex, to a school level in a given year. Since students' sex was coded 0 for male and 1 for female adolescents, the average of 1 indicated an all-girls school. Identification of a school as all-girls and/or Catholic was further cross-examined by visiting schools' websites and by contacting schools. To reduce heterogeneity, we excluded one sisterhood all-girls LSS that had only 21 graduates over the 10-year period, one sisterhood LSS with single-sex classes, and two all-girls LSSs not run by Catholic sisterhoods. We ended up with four all-girls (incl. three private) and 20 graduates in Catholic coed schools equalled 52–55%, depending on the cohort.

A total of 828 and 4,017 female adolescents graduated from the four all-girls and 20 coed schools in years 2007–2016, respectively. These students constituted 98% of all female graduates from these schools in the given time period; due to numerous administrative obstacles it was not possible to match the results of the PS and LSS exams for all students.

To further reduce heterogeneity we excluded students who repeated a grade in LSS (they were exposed to single-sex or coed environment for longer than students who did not repeat a grade). Also, students who were given an exemption from the PS or LSS exam in humanities or science were excluded from the analyses for a given outcome because these students, despite not taking the exam, received maximum scores as a prize for winning a subject Olympiad. Subjects, topics, and difficulty of Olympiads vary greatly, making the winners an immensely heterogeneous group; if they took the exam, they might not earn the highest possible score.

The final sample for studying the effect in the humanities included 803 female students from all-girls LSSs, and 3,934 female students from coed schools. The final sample for studying the effect in science included 815 female students from all-girls schools and 3,910 female students from coed schools. Less than 2% of the single-sex LSS graduates had attended single-sex PS and thus could not distort the final estimates. Detailed information of the sample in subsequent cohorts is available in the online supplement in Table S1.

Measures

Academic achievement

We used the results of the PS exam taken by students at the end of Grade 6 (just before transition to LSS) and the results of the LSS exam taken at the end of Grade 9 (at the end of LSS). The exams were nationwide and compulsory, and their content was based on the curriculum.

The PS exam, which the students took between 2004 and 2013, measured ability in general competence (mathematics and Polish). Its reliability as measured by Raju-Feldt's alpha (used for assessing reliability of tests comprising different item formats) varied between .82 and .86 (Szaleniec et al. 2015), depending on the year.

The LSS exam (taken between 2007 and 2016) included humanities and science – parts analysed separately. The humanities part covered Polish (reading and writing), history, and civic education, whereas the science part covered mathematics and natural sciences (biology, chemistry, geography, and physics). Reliabilities of the humanities and science parts as measured by Raju-Feldt's alpha equalled .84–.90 and .89–.92, respectively (Szaleniec et al. 2015). Detailed information about the exams is available in the online supplement.

Expected a posteriori (EAP) estimates, which were available in the used dataset, served as indicators of student ability. They are interpreted similarly to factor scores from factor analysis. These estimates were derived from Item Response Theory (IRT) models used to scale the exams: a two-parameter logistic model used in the case of dichotomous items and Samejima's graded response model in the case of polytomous items. The EAP estimation derives from Bayesian statistical principles. The EAP estimate is the expected value (i.e., mean) of the predicted probability distribution of possible scores of a given student, given the response pattern of that student, and the estimated model parameters (see Bock and Aitkin 1981 for a general description).

The models were estimated separately for each full cohort for the PS exam, the LSS exam in the humanities, and the LSS exam in science, 30 models in total. The sample size for scaling the PS and LSS exams varied between 460,679 and 324,675 depending on the year (in total 3.8 million graduates in 10 cohorts). The EAP estimates derived from each model were standardised and population centred (0; 1) within each full cohort separately. A detailed description of the scaling procedure is available in Żółtak (2015).

Controlled variables

Past studies have shown that single-sex schools are selective and tend to admit students who systematically differ from those attending coed schools in characteristics important for later school achievement, such as initial ability level (Hayes, Pahlke, and Bigler 2011; Jackson 2012; Signorella, Hayes, and Li 2013). Moreover, single-sex and coed schools may differ in the learning environments they provide (e.g., Choy 1997; Dronkers and Robert 2008). Thus, we controlled for the following student- and school-level characteristics: student past school achievement, student age, learning disabilities, average LSS class achievement at the moment of admission, proportion of peers from PS in the same LSS class, school size (number of students taking the LSS exam in a given year), and type of LSS school (public *vs.* private). This statistical control supplemented the implicit control achieved by comparing all-girls and coed schools run by Catholic sisterhoods.

Accounting for school achievement at the end of PS, just before the transition to LSS, allowed control of pre-existing differences in ability between female students in all-girls and coed schools. If they had differed and this had not been taken into account, differences in achievement at the end of LSS would have reflected initial differences instead of varied single-sex *vs*. coed school effectiveness (Jackson 2012; Signorella, Hayes, and Li 2013).

We controlled for learning disabilities because students who had them might score lower on exams (Swanson and Jerman 2006). The same negative impact of learning disability status in PS and in LSS on students' exam scores is expected. Single-sex schools, as more selective, might limit admissions of students with learning disabilities, because they did not want those students to lower the schools' average exam results in the future. However, in Poland some fully capable students might solicit learning disability status in LSS only to gain additional advantages, such as the extended time for taking LSS exam, in hope to increase their chances of being admitted to a better high school. This could produce contradictory results, i.e. students with learning disability status on LSS exam might in fact score higher than the ones without it, while controlling for PS exam and learning disability status in PS (Dolata et al. 2013). If a student was marked in our data as having a learning disability, she was diagnosed with a one (e.g., dyslexia) and presented the diagnosis certificate to the school authorities at a given school level. Students could present the certificate at one or both school levels.

Age (measured in months at the time of taking the LSS exam) was included because even small differences in age within the same cohort, a result of being born at various times of the year, postponing or bringing forward the age of school start, or skipping or repeating grades, may be important for school achievement. Older students usually perform better than younger ones (e.g., Lee and Fish 2010; Morrison, Alberts, and Griffith 1997), and the link between age and academic achievement is stronger for females (Smith 2009). Counterintuitively, the relationship between age and achievement while controlling for previous student outcomes is negative (Dolata et al. 2013); in other words, after controlling for past ability, younger students tend to score higher in later tests than older students. This is because the younger ones have already higher abilities during initial testing in comparison to the older students in their cohort.

We included school size because a vast body of literature has found that it might be related to student achievement, although the direction of the relationship is not entirely clear and may depend on various factors (e.g., Newman et al. 2006; Ready, Lee, and Welner 2004). Meanwhile, single-sex schools in Poland, because they are rare and non-mainstream, tend to be smaller than the coed ones. School size was operationalised as the number of all students who took the LSS exam in a given school in a given year.

Similarly, including average class achievement as a controlled variable is based on the notion that students show better performance if surrounded by peers with a similar ability level. The influence of peer ability level on individual academic performance has been found in numerous studies (e.g., Duflo, Dupas, and Kremer 2011; Hanushek et al. 2003). Average class achievement was calculated as the mean of the PS exam in a given LSS class and year.

We also controlled for the proportion of peers in a given student's class who attended the same PS. Evidence of negative effects of school transition on achievement has been reported in literature on student mobility (e.g., Pratt and George 2005). Having more

friends from a previous school in the current class prevents potential social isolation and positively affects individual school performance (Qualter et al. 2015; Wentzel et al. 2010). Students may feel more confident during the transition from PS to LSS when more friends are around. This could translate to more effective learning and higher LSS exam scores. On the other hand, this effect may diminish over time since LSS lasts three years; students not well adapted at the beginning of LSS may integrate into their class over time.

Controlling for school funding (public *vs*. private) provided an additional limit on the confounding influence of socioeconomic status. Students from wealthier families tend to have higher school achievement (e.g., Sirin 2005), and they attend private schools more often (e.g., Choy 1997), as they can afford the cost of the tuition fee. Moreover, private and public schools, due to differences in funding, may differ in various aspects potentially important for student achievement, such as teacher body, classroom equipment, or after-school opportunities (e.g., Choy 1997; Dronkers and Robert 2008).

Statistical analyses

The sample used in this study included female adolescents from 10 cohorts who attended 24 schools, which resulted in a cross-classified data structure. Each female student attended one and only one school and belonged to one and only one cohort. Meanwhile, each school occurred in multiple cohorts, and each cohort included data collected in multiple schools. In other words, students were nested in schools and years, but schools were not nested in years, and years were not nested in schools; they crossed instead. The data structure is presented in Table S1 in the online supplement.

In order to verify whether the year of taking the LSS exam (cohort) or the school accounted for variance in student achievement in the LSS exam, we estimated two empty (without covariates) cross-classified random intercept models with year of taking the LSS exam (cohort) and school as crossed levels. Student achievement in the humanities and science served as dependent variables. Next, we expanded the empty models by adding variables in focus. First, we added, as a fixed effect, information if a given school was single-sex. These models allowed us to see differences in average achievement in the LSS exam between female adolescents attending all-girls and coed schools. Second, in order to estimate the effect of single-sex schooling, we added a set of variables that controlled for initial differences between students attending those two types of schools and for potential differences between the schools. They were again fixed effects. At the student level, we included results of primary school exam, results of primary school exam squared, age in months, learning disabilities in primary school, learning disabilities in lower secondary school, per cent of peers from primary school in lower secondary school class, and mean result of primary school exam in lower secondary school class. Variables at the school level were: if school was public *vs*. private and school size (number of Grade 9 students in lower secondary school).

To verify the stability of the single-sex effect, we ran two-level mixed regression models for each year of taking the LSS exam (cohort) separately. The models included the same set of controlled variables. We did not include class as a separate level of analysis because initial analyses showed that adding this level did not improve model fit. Moreover, class level did not account for sizeable variance in the dependent variables, because 16 out of 24 schools had only one Grade 9 class in one or more cohort.

Individual pre-enrolment achievement in the PS exam, the LSS exam score in the humanities, and the LSS exam score in science were standardised and population centred (0; 1) within a given cohort. Student age was expressed in months at the moment of taking the LSS exam. Student sex, learning disability status, year of taking the LSS exam, school financing source, and school type variables were neither standardised nor centred.

We performed the analyses using the 'lmer' function in the 'lme4' package for R and maximum likelihood estimator (Bates et al. 2015). Standard errors were computed using the HC4 formula (Cribari-Neto 2004). In Cribari-Neto's (2004) simulation study HC4 outperformed previously developed formulas (e.g., HC3). Computations were double-checked in Stata/SE 14.2 statistical package using the 'mixed' command.

Results

Descriptive statistics

Both types of schools run by Catholic sisterhoods were selective. Young women admitted to those schools scored on average 0.873, 95% CI [0.851, 0.896] on the PS exam, which was by 0.775 SD above the national mean for young women in all ten cohorts. Young women admitted to Catholic schools also varied less in their prior school achievement (SD = 0.790, 95% CI [0.774, 0.806]) in comparison to the female population (SD = 0.963). Moreover, female students left those schools with mean achievement in science and the humanities almost one SD higher than female students in the population (mean result for the humanities 0.976, 95% CI [0.954, 0.998] and 0.978, 95% CI [0.953, 1.003] for science; that is respectively 0.823 and 0.951 SD higher than in the female population). This suggests that schools in our analyses might work with their students more effectively than other schools in the country. Population has been defined without students who repeated a grade in LSS or were given exemption from PS exam (a subject Olympiad laureates).

Descriptive statistics for the variables used in the study along with differences tests results between studied subsamples appear in (Table 1); details are available in Table S2 in the online supplement. The all-girls schools in the study were smaller than the coed ones. On average, twice as many students learned in a coed school in comparison to an all-girls one. In the coed schools, students were grouped on average in two classes, 95% CI [1.84, 2.172], whereas in all-girls schools in 1.25 classes, 95% CI [1.092, 1.408]. Moreover, classes in the coed schools were also larger, with on average 22.33, 95% CI [21.7, 22.957] students, while in all-girls schools 16.56, 95% CI [14.835, 18.284].

Other significant differences between the all-girls and coed schools included the number of students with learning disability status at LSS level, the number of laureates in science, mean score in LSS exam in the humanities, mean percentage of peers from the same PS that a student has in LSS class, and the number of all Grade 9 students in LSS. There were no significant differences at the confidence level of 95% between all-girls and coed schools in the number of students with learning disability status at PS level, mean score in PS exam, mean score in LSS exam in science, the number of laureates in humanities, and students' age.

In summary, the all-girls and coed schools selected for analyses differed from schools in the population in their students' average school achievement. However, they also

Table 1. Descriptive statistics of the variables used in the study, calculated on the final data subset[1].

Variable	Statistic	All-girls schools	Females in coed schools	Combined	p[4]
No. of schools (incl. private ones)	n	4 (3)	20 (9)	24 (12)	NA
No. of female students included in the analyses	n	816	3,971	4,787	NA
LD in PSs	n (%)	90 (11.0)	468 (11.8)	558 (11.7)	0.955
LD in LSSs	n (%)	139 (17.0)	507 (12.8)	646 (13.5)	.004
Laureate in humanities	n (%)	12 (1.5)	37 (0.9)	49 (1)	.126
Laureate in science	n (%)	1 (0.1)	60 (1.5)	61 (1.3)	.020
PS exam	M (SD)	0.866 (0.742)	0.875 (0.800)	.873 (0.790)	1
Humanities exam[2]	M (SD)	1.116 (0.733)	0.947 (0.774)	.976 (0.769)	.006
Science exam[2]	M (SD)	1.052 (0.832)	0.968 (0.899)	.978 (0.889)	.053
Age in months	M (SD)	189.30 (3.89)	189.23 (3.70)	189.24 (3.73)	0.997
Pct. of peers from PS attending the same class in LSS	M (SD)	8.89 (6.54)	31.92 (32.57)	27.98 (31.02)	.006
Avg. No. of all G9 students in LSS[3]	M (SD)	22.55 (9.12)	45.50 (33.32)	41.07 (31.51)	.006

Note. LD = learning disability; PS = primary school; LSS = lower secondary school; M = mean; SD = standard deviation; G9 = Grade 9, NA = not applicable.
[1]Only female students for whom it was possible to match PS and LSS exam results, without those who repeated a grade in LSS or who were given an exemption from the PS exam.
[2]A subject Olympiad winners (laureates) in LSS excluded.
[3]School-level variable; it reflects the number of all students who attended Grade 9 classes. For coed schools it includes both female and male adolescents. It includes also students for whom it was impossible to match PS and LSS exam results, those who repeated a grade in LSS, and those who were given exemption from the PS or LSS exam.
[a]With Bonferroni correction for multiple comparisons; for dichotomous variables: Chi[2], for continuous variables: t-test.

notably differed from each other in several factors potentially important for student achievement, which supports the decision to include various controlled variables in the models.

Single-sex schooling effect

The analysis of empty cross-classified random-effects models showed that 15.3%, 95% CI [7.5, 22.2] and 16.2%, 95% CI [8.4, 24.4] of the variation in the results of the LSS exam in the humanities and science, respectively, lay between schools. The share of variation explained by the year of taking the LSS exam (cohort) was below 1%.

Table 2 contains the results of regressing school achievement in the humanities and science on the type of school (all-girls *vs.* coed) with and without controlling for student and school characteristics. Detailed results are available in the online supplement Table S5. The analyses demonstrated that after controlling for initial student and school differences, female adolescents attending all-girls schools scored 0.166, 95% CI [−0.02, 0.352] of SD higher on the LSS exam in the humanities and by 0.165, 95% CI [0.01, 0.32] in science compared to young women from coed schools. This result does not support the hypotheses derived from the expectancy-value model. However, although it does not support the prediction on the humanities based on the input-process-output model, it does so for science.

Table 2 also reports the estimates of the effect obtained in two-level mixed effects models run for each cohort (year of taking the LSS exam) separately. Detailed results are presented in Tables S3 and S4 in the online supplement. Those models allowed checking the stability of the effect in subsequent cohorts. The effect of single-sex schooling proved

Table 2. Single-sex schooling effect on achievement in the humanities and science (regression coefficients from cross-classified multilevel models).

	LSS exam in humanities		LSS exam in science	
Model	Without controls variables	With controls variables	Without controls variables	With controls variables
All schools[a]				
	0.138 (0.165)	0.166[b] (0.095)	0.128 (0.198)	0.165*[b] (0.079)
All schools by year[c]				
2007		0.254 (0.130)		0.375* (0.159)
2008		0.165 (0.093)		0.191* (0.086)
2009		0.131 (0.096)		0.006 (0.114)
2010		−0.152 (0.170)		−0.094 (0.096)
2011		0.269 (0.143)		0.149 (0.080)
2012		0.243 (0.134)		0.327** (0.114)
2013		0.145 (0.088)		0.056 (0.110)
2014		0.071 (0.102)		0.078 (0.106)
2015		0.135 (0.098)		0.041 (0.107)
2016		0.121 (0.150)		0.062 (0.143)

Note. LSS = lower secondary school. All models with controls include the following variables: (i) student level: results of primary school exam, results of primary school exam squared, age in months, learning disabilities in primary school, learning disabilities in lower secondary school, percent of peers from primary school in lower secondary school class, mean result of primary school exam in lower secondary school class, (ii) school level: public *vs.* private school, number of Grade 9 students in lower secondary school. Full results are available in the online supplement. Heteroscedasticity consistent standard errors are presented in parentheses. [a]Three-level regression model.

[b]The bias-corrected estimates in bootstrap (1000 rep.) are for science 0.163 CI 95% [0.009, 0.318] and for humanities 0.163 CI 95% [−0.021, 0.346].

[c]Two-level regression model.

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

to be positive except for the year 2010 in the cases of the humanities and science. We do not have an explanation for this negative effect. Since both estimates are nonsignificant we treat this as a disturbance in the data.

The magnitude and signs of the regression coefficients estimated for controlled variables were as expected based on theory and previous research. Student pre-enrolment and class mean achievement were the strongest predictors of the results of the LSS exam in the humanities and science. However, the relationship between individual achievement and the LSS exam was not linear (see the results for the squared term of PS exam score). Age and learning disability status also proved important. Details are presented in Tables S3, S4, and S5 in the online supplement.

Discussion

The main goal of this study was to estimate the effect of single-sex schooling on student achievement. We used national examination results in the humanities and science of 10 cohorts of female adolescents attending Polish all-girls and coed schools run by Roman Catholic sisterhoods. Young women in single-sex schools scored on average 0.17 SDs higher in the humanities and science than those in coed schools. However, because the analyses included only four all-girls and 20 coed schools, resulting in relatively high standard errors of the school-level estimates and large ranges of the confidence intervals, only the effect for science was statistically significant. The large standard errors were mostly due to small number of schools and cross-classified data structure. Moreover, the study included all all-girls and coed Catholic LSSs in Poland and lacked only 2% of their

female students (due to difficulties in merging the examination results) and therefore was run on the almost whole population of interest. As a consequence, focusing solely on statistical significance of the effect might not give a full picture. The results suggest that the effect, although statistically significant only for science, is probably positive for both the humanities and science. Moreover, although the effect size was relatively small, it cannot be concluded that the impact of single-sex schooling is negligible or trivial, because the effect may accumulate over time and significantly affect academic achievement of those being exposed to it over period longer than three years.

Which theory helps interpret the results?

The observed pattern of differences stands in contrast with the expected lower achievement in science and higher in the humanities in all-girls schools according to the expectancy-value model (Eccles 2009). Meanwhile, significantly higher science achievement in all-girls schools agrees with the hypothesis based on peer effects framed in the input-process-output model (e.g., Scheerens 2004). However, similar in size but non-significant difference in the humanities does not fully support the model. Since the potential mechanism of the effect does not differ between subjects, female adolescents in all-girls schools should score significantly higher in both domains, and the fact that they did not, although does not disprove the model, does not warrant firm conclusions and suggests that further research is needed.

Although the results do not support the expectancy-value mechanism, they do not necessarily disprove it. For science, the mechanisms derived from the expectancy-value model and the input-process-output model are contradictory, therefore their effects should cancel each other out if both are present. However, the two effects may not be equally strong. It may be possible that female students in all-girls schools outperformed female students in coed schools in science because both mechanisms had their effects, but the effect of input-process-output mechanism was stronger.

Do the results fit previous findings?

Vast body of literature has not provided clear evidence on direction and size of single-sex schooling effect on female adolescents' achievement. Most discrepancies steam from insufficient control of (self-)selection processes to often specific and selective single-sex schools as well as differences between coed and single-sex schools' resources, most importantly – in teaching quality. However, methodologically robust studies presented in the theoretical part of this paper help to narrow down the scope of conclusions. As noted earlier (see also Figure 1), despite the fact that only a few studies reported statistically significant effects, the effects tend to lean towards positive values. The results of this study fit in this tendency. Moreover, point estimates reported in this study are 2.4 times higher in the case of verbal and 3.4 times higher in the case of maths and science, than point estimates reported in studies summarised in Figure 1 supplemented with the results of this study is available in the online supplement (see Figure 1s).

Methodological strengths

Studies that investigate the effect of single-sex schooling by using experimental designs, although highly informative, have several limitations. The studies that used data gathered in Seoul (Hahn and Wang 2019; Park, Behrman, and Choi 2013, 2018; Sohn 2016) provided results limited to a local education system in one urban area. The work by Jackson (2016) examined only two cohorts of students, who went through complete five-year secondary school programme as single-sex schools' students. Moreover, he estimated the effect on school achievement using the results of tests taken by students in Grades 6 and 8. The short-term gain reported by the author could have resulted from novelty-based enthusiasm of teachers, school staff, and students after transition. The results obtained by Eisenkopf et al. (2015) refer to single-sex classes, not schools. Our study, although not based on an experimental scheme, overcomes these drawbacks: it investigates schools from various parts of Poland, uses ten cohorts of students, and focuses on single-sex schools, not classrooms.

Two more methodological strengths of this study are worth to note. First, contrary to many previous studies, which relied on rather small samples (for a review see Pahlke, Hyde, and Allison 2014), we used large dataset of all eligible female students of LSSs run by Catholic sisterhoods in Poland. Second, by limiting our analysis to schools run by Catholic sisterhoods we have addressed potential bias due to (self-)selection and between-school differences. This was additionally strengthened through statistical control of various student- and school-level characteristics, including prior academic achievement.

Limitations

This study, although conducted with the utmost rigour, has several limitations. First, we acknowledge that the results should not be generalised to other groups, for example to all-girls schools not run by Roman Catholic sisterhoods.

Second, although the all-girls and coed schools included in the analyses were religion-based and thus shared important similarities, unknown differences between them might still influence the results. All-girls schools might attract better teachers, provide special infrastructure (e.g., laboratories), or offer more after-school learning programmes. In other words, the observed differences in effectiveness might stem from differences in teaching quality or learning opportunities between the two types of schools. However, the dataset we used, although unique, did not contain information that would allow us to control for such potential differences.

Third, we cannot exclude different selection mechanisms for the two types of schools. Those differences are suggested by differences in the average share of students in class who attended the same PS. In the coed LSSs in the analysis, female adolescents had in their class on average 32% of peers from the same PS, which is similar to the value observed in the population, including the trend towards a strong selection process to schools in urban areas (32% on average, 48% in rural areas, 14% in large cities, Humenny et al. 2014). Since male students are not allowed in all-girls LSSs, the share of peers from the same PS in an LSS class is expected to be lower by definition. The average sex ratio in PSs in Poland is approximately one. If there were no (self-)selection, the share of peers

from the same PS in a single-sex LSS class should be around half of the value observed in coed schools (i.e., 16%). However, it equalled 9%. In other words, students could choose coed schools run by Catholic sisterhoods and secular coed schools equally often, based solely on geographical proximity to residence, indicating that the former might be an equal alternative to secular coed schools. Meanwhile, all-girls schools were not chosen equally often based on their proximity and, as non-mainstream schools, may have attracted specific students. This specificity could affect later school achievement, leading most likely to inflated estimates of the effect of single-sex schooling. We tried to reduce this bias by controlling for the share of peers from PS in LSS; the coefficients oscillated around zero.

Fourth, although families of students from all-girls and coed schools are similar in many aspects, it is not clear to what extent the results were influenced by differences in the home learning environments of the students attending the two types of school. Parents of female adolescents attending all-girls schools could be more involved in their education, which resulted in choosing a non-mainstream school and later greater support. The role of parents as active actors in school choice (e.g., Dziemianowicz-Bąk, Dzierzgowski, and Wojciuk 2015) and the importance of parental involvement for student achievement are widely recognised in the literature (e.g., Castro et al. 2015). On the other hand, potential differences in parental involvement, as they affected the PS examination results, have been at least partially controlled for through including pre-enrolment achievement in the analyses.

Future research directions

The limited theoretical underpinnings of the majority of the research on single-sex schooling and unclear mechanism of the effect highlight the need to shift researchers' attention from estimating the effect itself to the development and thorough testing of comprehensive theoretical model(s) of the mechanism through which single-sex schooling affect student outcomes.

Large standard errors of effects observed even in multi-cohort or large-scale studies suggest that some students may benefit from attending them, whereas others may not, and that single-sex schools may differ in effectiveness. Therefore, future studies should go beyond the mean effect and search for its student- and school-level moderators, for example motivation (Jackson 2012), student perceptions of their school experiences, reasons for choosing a single-sex schools, student-teacher gender matching (Eisenkopf et al. 2015; Jackson 2012; Park, Behrman, and Choi 2018), or other classroom processes.

Moreover, majority of literature focused on the effect of single-sex schooling on academic achievement and did not address other important outcomes, such as interpersonal skills, self-concept, learning motivation, educational and occupational aspirations (Pahlke, Hyde, and Allison 2014; Signorella, Hayes, and Yidi 2013) Focusing scholars' effort on this area of research is of the most importance.

Furthermore, most of the studies compare achievement of students from single-sex and coed schools in academic tracks, i.e. high schools. As a consequence, they cover only high achievers who are more motivated, exposed to more challenging curricula, and better teaching quality. Future research should include less selective samples by for

example investigating the effect at earlier educational stages (i.e. before the first selection threshold) and among students from non-academic tracks.

Practice implications

Although this study showed a small effect of single-sex schooling, attending an all-girls school may have important long-term consequences for individuals, especially if it precedes a high-stakes exam, just as it did in this study. Even a minor increase in the test score may be decisive for admission to the chosen school at a higher educational level, or preferred education track and therefore affect future educational career and job opportunities. Moreover, small effects may accumulate over time and have major impact if a single-sex school is attended for a long time, for example at subsequent educational stages.

However, despite the fact that all-girls schools might contribute more to student achievement than coed schools, one should not conclude that school systems need to be redesigned to provide single-sex education for all. The available examples of mass single-sex education have proven that enforcing that type of schooling is ineffective and might yield various negative consequences (Ewing, 2006). However, the existence of single-sex schools in otherwise coed school systems may support diversity of educational options and serve the needs of some groups of students (Jackson 2012; Sohn 2016).

Better results of female adolescents attending single-sex schools raise the question if single-sex schooling can be a remedy for the underrepresentation of women in STEM. We believe that the answer is negative. First, the effect is small, and although it can potentially change one's educational trajectory, it is not strong enough to be a promising direction in terms of educational policy. Second, STEM-related outcomes other than achievement, for instance maths and science interest and self-efficacy, or choosing a STEM major, do not differ between female students of all-girls and coed schools (e.g., Park, Behrman, and Choi 2018; Sohn 2016 in case of public schools). Third, societal and psychological factors causing this underrepresentation are too complex and manifold to be compensated by all-girls schools. Moreover, the mechanism through which these factors may lead to such underrepresentation is not fully understood (Stoet and Geary 2018).

Conclusions

This study has shown that single-sex schools may increase achievement of female adolescents in science. However, the theoretical models we adopted were inconsistent with this pattern of results. Meanwhile, understanding the mechanism behind the effect of single-sex schooling is key for both proponents and opponents of single-sex schools, as well as for policymakers. If the mechanism is specific to single-sex schools and cannot occur in the coed ones, it will support their existence in educational systems. If it is generalisable and can be initiated also in coed schools (e.g., reducing sexual harassment of girls), it may help to improve their organisational and teaching-related processes. Nevertheless, investigating such mechanism should not be viewed as a path to refute claims of proponents or opponents of

single-sex schools, but as a way of bridging the debate between them, since their goal remains the same – providing high-quality education for all.

Acknowledgments

The authors would like to thank Mateusz Żółtak and Tomasz Żółtak for their help in gaining access to data used in this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Polish National Science Centre under Grant number 2014/13/N/HS6/02974. The article uses data collected and processed with support from the European Social Fund under Grant number UDA-POKL.03.02.00-00-001/13-00.

Notes on contributors

Maciej Koniewski is an assistant professor at the Institute of Sociology of the Jagiellonian University in Krakow, Poland. Among his research topics are educational measurement, school effectiveness and public policy.

Anna Hawrot is a postdoctoral researcher at the Leibniz Institute for Educational Trajectories in Bamberg, Germany. Her research interests include the role of various school and family factors for cognitive and non-cognitive outcomes.

ORCID

Maciej Koniewski http://orcid.org/0000-0001-9437-2822
Anna Hawrot http://orcid.org/0000-0002-2784-5455

References

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. "Fitting Linear Mixed-Effects Models Using Lme4." *Journal of Statistical Software* 67 (1): 1–48. doi:10.18637/jss.v067.i01.

Bigler, R. S., A. R. Hayes, and L. S. Liben. 2014. "Analysis and Evaluation of the Rationales for Single-Sex Schooling." In *Advances in Child Development and Behavior*, edited by L. S. Liben and R. S. Bigler, 225–260. Vol. 47. Burlington: Academic Press.

Castro, M., E. Expósito-Casas, E. López-Martín, L. Lizasoain, E. Navarro-Asencio, and J. L. Gaviria. 2015. "Parental Involvement on Student Academic Achievement: A Meta-Analysis." *Educational Research Review* 14 (February): 33–46. doi:10.1016/j.edurev.2015.01.002.

Choy, S. P. 1997. *Public and Private Schools: How Do They Differ? Findings from the Condition of Education, No. 12*. Washington, DC: National Center for Education Statistics.

Cribari-Neto, F. 2004. "Asymptotic Inference under Heteroskedasticity of Unknown Form." *Computational Statistics & Data Analysis* 45 (2): 215–233. doi:10.1016/S0167-9473(02)00366-3.

Daly, P., and N. Defty. 2004. "Extension of Single-Sex Public School Provision: Evidential Concerns." *Evaluation & Research in Education* 18 (1–2): 129–136. doi:10.1080/09500790408668313.

Darrell, B. R., and M. Aitkin. 1981. "Marginal Maximum Likelihood Estimation of Item Parameters: Application of an EM Algorithm." *Psychometrika* 46 (4): 443–459. doi:10.1007/BF02293801.

Dolata, R., A. Hawrot, G. Humenny, A. Jasińska, M. Koniewski, P. Majkut, and Ż. Tomasz. 2013. *Trafność Metody Edukacyjnej Wartości Dodanej Dla Gimnazjów*. Warszawa: Instytut Badań Edukacyjnych.

Dronkers, J., and P. Robert. 2008. "Differences in Scholastic Achievement of Public, Private Government-Dependent, and Private Independent Schools: A Cross-National Analysis." *Educational Policy* 22 (4): 541–577. doi:10.1177/0895904807307065.

Duflo, E., P. Dupas, and M. Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101 (5): 1739–1774. doi:10.1257/aer.101.5.1739.

Dziemianowicz-Bąk, A., J. Dzierzgowski, and A. Wojciuk. 2015. *Autoselekcja Na Progu Gimnazjum – Działania Rodziców W Kontekście Działań Szkół I Polityki Samorządu*. Warsaw: Instytut Badań Edukacyjnych.

Eccles, J. 2009. "Who Am I and What Am I Going to Do with My Life? Personal and Collective Identities as Motivators of Action." *Educational Psychologist* 44 (2): 78–89. doi:10.1080/00461520902832368.

Eisenkopf, G., Z. Hessami, U. Fischbacher, and H. W. Ursprung. 2015. "Academic Performance and Single-Sex Schooling: Evidence from a Natural Experiment in Switzerland." *Journal of Economic Behavior & Organization* 115 (July): 123–143. doi:10.1016/j.jebo.2014.08.004.

Ewing,, E. T. 2006. "The Repudiation of Single-Sex Education: Boys' Schools in the Soviet Union, 1943–1954." *American Educational Research Journal* 43 (4): 621–650. doi:10.3102/00028312043004621.

Garcia-Gracia, M., and T. D. Vázquez. 2016. "Mixed Schools versus Single-Sex Schools: Are There Differences in the Academic Results for Boys and Girls in Catalonia?" *International Journal of Inclusive Education* 20 (2): 149–167. doi:10.1080/13603116.2015.1079269.

Hahn, Y., and L. C. Wang. 2019. "The Effectiveness of Single-Sex Schools through out-of-School Activities: Evidence from South Korea." *Oxford Bulletin of Economics and Statistics* 81 (2): 369–393. doi:10.1111/obes.12266.

Halpern, D. F., L. Eliot, R. S. Bigler, R. A. Fabes, L. D. Hanish, J. Hyde, L. S. Liben, and C. L. Martin. 2011. "The Pseudoscience of Single-Sex Schooling." *Science* 333 (6050): 1706–1707. doi:10.1126/science.1205031.

Hanushek, E. A., J. F. Kain, J. M. Markman, and S. G. Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18 (5): 527–544. doi:10.1002/jae.741.

Hayes, A. R., E. E. Pahlke, and R. S. Bigler. 2011. "The Efficacy of Single-Sex Education: Testing for Selection and Peer Quality Effects." *Sex Roles* 65 (9–10): 693–703. doi:10.1007/s11199-010-9903-2.

Hoxby, C. 2000. "*Peer Effects in the Classroom: Learning Form* Gender *and* Race Variation." Working Paper 7876. NEBER Working Paper Series. Cambridge, MA: National Bureau of Economic Research. https://www.nber.org/papers/w7867.pdf.

Humenny, G., M. Koniewski, P. Majkut, and S. Paulina. 2014. "Migracje Uczniów Między Zespołami Klasowymi Przy Przejściu Ze Szkoły Podstawowej Do Gimnazjum." In *Diagnozy Edukacyjne: Dorobek I Nowe Zadania*, edited by B. Niemierko and M. K. Szmigel, 121–141. Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.

Jackson, C. K. 2012. "Single-Sex Schools, Student Achievement, and Course Selection: Evidence from Rule-Based Student Assignments in Trinidad and Tobago." *Journal of Public Economics* 96 (1–2): 173–187. doi:10.1016/j.jpubeco.2011.09.002.

Jackson, C. K. 2016. "The Effect of Single-Sex Education on Test Scores, School Completion, Arrests, and Teen Motherhood: Evidence from School Transitions." Working Paper 22222.

NEBER Working Paper Series. Cambridge, MA: National Bureau of Economic Research. https://www.nber.org/papers/w22222.pdf.

Kessels, U., and B. Hannover. 2008. "When Being a Girl Matters Less: Accessibility of Gender-Related Self-Knowledge in Single-Sex and Coeducational Classes and Its Impact on Students' Physics-Related Self-Concept of Ability." *British Journal of Educational Psychology* 78 (2): 273–289. doi:10.1348/000709907X215938.

Ku, H., and D. W. Kwak. 2013. "Together or Separate: Disentangling the Effects of Single-Sex Schooling from the Effects of Single-Sex Schools." 487. Discussion Paper Series. School of Economics, University of Queensland, Australia. http://www.uq.edu.au/economics/abstract/487.pdf.

Lavy, V., and A. Schlosser. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal. Applied Economics* 3 (2): 1–33. doi:10.1257/app.3.4.1.

Lee, J., and R. M. Fish. 2010. "International and Interstate Gaps in Value-Added Math Achievement: Multilevel Instrumental Variable Analysis of Age Effect and Grade Effect." *American Journal of Education* 117 (1): 109–137. doi:10.1086/656348.

Lee, S. H., and O. Jerman. 2006. "Math Disabilities: A Selective Meta-Analysis of the Literature." *Review of Educational Research* 76 (2): 249–274. doi:10.3102/00346543076002249.

LePore, P. C., and J. R. Warren. 1997. "A Comparison of Single-Sex and Coeducational Catholic Secondary Schooling: Evidence from the National Educational Longitudinal Study of 1988." *American Educational Research Journal* 34 (3): 485–511. doi:10.3102/00028312034003485.

Liben, L. S. 2015. "Probability Values and Human Values in Evaluating Single-Sex Education." *Sex Roles* 72 (9–10): 401–426. doi:10.1007/s11199-014-0438-9.

Lu, F., and M. L. Anderson. 2015. "Peer Effects in Microenvironments: The Benefits of Homogeneous Classroom Groups." *Journal of Labor Economics* 33 (1): 91–122. doi:10.1086/677392.

Mael, F., A. Alonso, D. Gibson, K. Rogers, and M. Smith. 2005. *Single-Sex versus Coeducational Schooling: A Systematic Review*. Washington, DC: U.S. Department of Education, Of- fice of Planning, Evaluation and Policy Development, Policy and Program Studies Service. doi:10.1037/e566992006-001.

Marsh, H. W. 1989. "Effects of Attending Single-Sex and Coeducational High Schools on Achievement, Attitudes, Behaviors, and Sex Differences." *Journal of Educational Psychology* 81 (1): 70–85. doi:10.1037/0022-0663.81.1.70.

Marsh, H. W. 1991. "Public, Catholic Single-Sex, and Catholic Coeducational High Schools: Their Effects on Achievement, Affect, and Behaviors." *American Journal of Education* 99 (3): 320–356. doi:10.1086/443985.

Morrison, F. J., D. M. Alberts, and E. M. Griffith. 1997. "Nature–Nurture in the Classroom: Entrance Age, School Readiness, and Learning in Children." *Developmental Psychology* 33 (2): 254–262. doi:10.1037/0012-1649.33.2.254.

Nagengast, B., H. W. Marsh, and K.-T. Hau. 2013. "Effects of Single-Sex Schooling in the Final Years of High School: A Comparison of Analysis of Covariance and Propensity Score Matching." *Sex Roles* 69 (7): 404–422. doi:10.1007/s11199-013-0261-8.

Newman, M., Z. Garrett, D. Elbourne, S. Bradley, P. Noden, J. Taylor, and A. West. 2006. "Does Secondary School Size Make A Difference? A Systematic Review." *Educational Research Review* 1 (1): 41–60. doi:10.1016/j.edurev.2006.03.001.

Pahlke, E., and J. S. Hyde. 2016. "The Debate over Single-Sex Schooling." *Child Development Perspectives* 10 (2): 81–86. doi:10.1111/cdep.12167.

Pahlke, E., J. S. Hyde, and C. M. Allison. 2014. "The Effects of Single-Sex Compared with Coeducational Schooling on Students' Performance and Attitudes: A Meta-Analysis." *Psychological Bulletin* 140 (4): 1042–1072. doi:10.1037/a0035740.

Park, H., J. R. Behrman, and J. Choi. 2013. "Causal Effects of Single-Sex Schools on College Entrance Exams and College Attendance: Random Assignment in Seoul High Schools." *Demography* 50 (2): 447–469. doi:10.1007/s13524-012-0157-1.

Park, H., J. R. Behrman, and J. Choi. 2018. "Do Single-Sex Schools Enhance Students' STEM (Science, Technology, Engineering, and Mathematics) Outcomes?" *Economics of Education Review* 62 (February): 35–47. doi:10.1016/j.econedurev.2017.10.007.

Pohl, S., P. M. Steiner, J. Eisermann, R. Soellner, and T. D. Cook. 2009. "Unbiased Causal Inference from an Observational Study: Results of a within-Study Comparison." *Educational Evaluation and Policy Analysis* 31 (4): 463–479. doi:10.3102/0162373709343964.

Praetorius, A.-K., E. Klieme, B. Herbert, and P. Pinger. 2018. "Generic Dimensions of Teaching Quality: The German Framework of Three Basic Dimensions." *ZDM* 50 (3): 407–426. doi:10.1007/s11858-018-0918-4.

Pratt, S., and R. George. 2005. "Transferring Friendship: Girls' and Boys' Friendships in the Transition from Primary to Secondary School." *Children & Society* 19 (1): 16–26. doi:10.1002/chi.830.

Qualter, P., J. Vanhalst, R. Harris, E. Van Roekel, G. Lodder, M. Bangee, M. Maes, and M. Verhagen. 2015. "Loneliness across the Life Span." *Perspectives on Psychological Science* 10 (2): 250–264. doi:10.1177/1745691615568999.

Ready, D. D., V. E. Lee, and K. G. Welner. 2004. "Educational Equity and School Structure: School Size, Overcrowding, and Schools-within-Schools." *Teachers College Record* 106 (10): 1989–2014. doi:10.1111/j.1467-9620.2004.00424.x.

Riordan, C. 2015. *Single-Sex Schools. A Place to Learn*. Lanham: Rowman & Littlefield Publishers.

Ryan, R. M., and E. L. Deci. 2000. "Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions." *Contemporary Educational Psychology* 25 (1): 54–67. doi:10.1006/ceps.1999.1020.

Scheerens, J. 2004. "Review of School and Instructional Effectiveness Research. Background Paper Prepared for the Education for All Global Monitoring Report 2005 the Quality Imperative." 2005/ED/EFA/MRT/PI/44. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000146695.

Seidel, T., and R. J. Shavelson. 2007. "Teaching Effectiveness Research in the past Decade: The Role of Theory and Research Design in Disentangling Meta-Analysis Results." *Review of Educational Research* 77 (4): 454–499. doi:10.3102/0034654307310317.

Shadish, W. R., M. H. Clark, and P. M. Steiner. 2008. "Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments." *Journal of the American Statistical Association* 103 (484): 1334–1344. doi:10.1198/016214508000000733.

Signorella, M. L., A. R. Hayes, and Y.Li. 2013. "A Meta-Analytic Critique of Mael Et Al.'s (2005) Review of Single-Sex Schooling." *Sex Roles* 69 (7): 423–441. doi:10.1007/s11199-013-0288-x.

Sirin, S. R. 2005. "Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research." *Review of Educational Research* 75 (3): 417–453. doi:10.3102/00346543075003417.

Smith, J. 2009. "Can Regression Discontinuity Help Answer an Age-Old Question in Education? The Effect of Age on Elementary and Secondary School Achievement." *The B.E. Journal of Economic Analysis & Policy* 9 (1): 1. doi:10.2202/1935-1682.2221.

Sohn, H. 2016. "Mean and Distributional Impact of Single-Sex High Schools on Students' Cognitive Achievement, Major Choice, and Test-Taking Behavior: Evidence from a Random Assignment Policy in Seoul, Korea." *Economics of Education Review* 52 (June): 155–175. doi:10.1016/j.econedurev.2016.02.007.

Stoet, G., and D. C. Geary. 2018. "The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education." *Psychological Science* 29 (4): 581–593. doi:10.1177/0956797617741719.

Szaleniec, H., B. Kondratek, F. Kulon, A. Pokropek, P. Skórska, K. Świst, T. Wołodźko, and Ż. Mateusz. 2015. *Porównywalne Wyniki Egzaminacyjne*. Warszawa: Instytut Badań Edukacyjnych.

Van de Gaer, E., H. Pustjens, J. Van Damme, and A. De Munter. 2004. "Effects of Single-sex versus Co-educational Classes and Schools on Gender Differences in Progress in Language and Mathematics Achievement'." *British Journal of Sociology of Education* 25 (3): 307–322. doi:10.1080/0142569042000216963.

Wentzel, K. R., A. Battle, S. L. Russell, and L. B. Looney. 2010. "Social Supports from Teachers and Peers as Predictors of Academic and Social Motivation." *Contemporary Educational Psychology* 35 (3): 193–202. doi:10.1016/j.cedpsych.2010.03.002.

Wiseman, A. W. 2008. "A Culture of (In)equality?: A Cross-National Study of Gender Parity and Gender Segregation in National School Systems." *Research in Comparative and International Education* 3 (2): 179–201. doi:10.2304/rcie.2008.3.2.179.

Żółtak, T. 2015. *Statystyczne Modelowanie Wskaźników Edukacyjnej Wartości Dodanej: Podsumowanie Polskich Doświadczeń 2005-2015*. Warszawa: Instytut Badań Edukacyjnych.