

Executive Summary

Single-sex education refers most generally to education at the elementary, secondary, or postsecondary level in which males or females attend school exclusively with members of their own sex. This report deals primarily with single-sex education at the elementary and secondary levels. Research in the United States on the question of whether public single-sex education might be beneficial to males, females or a subset of either group (particularly disadvantaged youths) has been limited. However, because there has been a resurgence of single-sex schools in the public sector, it was deemed appropriate to conduct a systematic review of single-sex education research.

A number of theoretical advantages to both coeducational (CE) and single-sex (SS) schools have been advanced by their advocates, a subset of whom have focused specifically on the potential benefits of SS schooling for disadvantaged males who have poor success rates in the educational system. The interpretation of results of previous studies in the private sector or the public sectors of other countries has been hotly debated, resulting in varying policy recommendations based on the same evidence. However, no reviews on this topic have been conducted using a systematic approach similar to that of the Campbell Collaboration (CC) or the What Works Clearinghouse (WWC). Thus, the objective of this review is to document the outcome evidence for or against the efficacy of single-sex education as an alternative form of school organization using an unbiased, transparent, and objective selection process adapted from the standards of the CC and WWC to review quantitative studies.

Concurrently with this review of the quantitative literature, we conducted a review of the qualitative literature on the subject of single-sex schooling using parallel coding techniques. Unlike quantitative studies, qualitative studies are not viewed by WWC as appropriate methodology when determining causal relationships. Rather, they contribute to theory building and provide direction for hypothesis testing. Few qualitative studies satisfied the criteria for inclusion. Therefore, the primary focus of this paper is the systematic review of quantitative research.

The following are the major research questions addressed by the systematic quantitative review:

1. Are single-sex schools more or less effective than coeducational schools in terms of concurrent, quantifiable academic accomplishments?
2. Are single-sex schools more or less effective than coeducational schools in terms of long-term, quantifiable academic accomplishment?
3. Are single-sex schools more or less effective than coeducational schools in terms of concurrent, quantifiable indicators of individual student adaptation and socioemotional development?

4. Are single-sex schools more or less effective than coeducational schools in terms of long-term, quantifiable indicators of individual student adaptation and socioemotional development?
5. Are single-sex schools more or less effective than coeducational schools in terms of addressing issues of procedural (e.g., classroom treatment) and outcome measures of gender inequity?
6. Are single-sex schools more or less effective than coeducational schools in terms of perceptual measures of the school climate or culture that may have an impact on performance?

As in previous reviews, the results are equivocal. There is some support for the premise that single-sex schooling can be helpful, especially for certain outcomes related to academic achievement and more positive academic aspirations. For many outcomes, there is no evidence of either benefit or harm. There is limited support for the view that single-sex schooling may be harmful or that coeducational schooling is more beneficial for students.

THE SYSTEMATIC REVIEW PROCESS

The systematic review of the literature consisted of the following steps:

1. An exhaustive search of electronic databases for citations, supplemented by other sources. This search strategy yielded 2,221 studies.
2. An initial Phase I exclusion of sources whose subject matter falls outside the defined scope of the study. Criteria used for exclusion in Phase I included:

Population—To be included, the students had to be enrolled in a full-time school. They had to be in elementary, middle, or high school as opposed to college and beyond. Finally, the schools being studied had to be in English-speaking or Westernized countries somewhat comparable to American public-sector schools.

Intervention—The single-sex school had to be one in which students were either completely segregated by sex or were completely segregated for all classes, even if co-located in the same building (i.e., dual academies). Studies of single-sex classes in a coeducational school were excluded from review.

This initial screening yielded 379 publications that fit the initial inclusion criteria.

3. A Phase II exclusion based on obvious methodological considerations (e.g., nonstudy, weak study). On the basis of titles and abstracts, citations that appeared to be essays, reviews, opinion pieces, and similar items were excluded, and only qualitative and quantitative studies that were likely to be codable in Phase III were retained. During Phase II, 114 citations were culled from the 379 items and coded as appropriate for review as quantitative (88) or qualitative (26) studies. Of the 26 qualitative studies, 4 met the criteria for final inclusion and were reviewed separately.

4. A Phase III evaluation and coding of the remaining quantitative articles. According to the guidelines of the WWC, all studies other than randomized controlled trials, quasi-experimental designs (QED) with matching, or regression discontinuity designs would be excluded prior to Phase III. Under the WWC criteria for inclusion, virtually all single-sex studies would have been eliminated from the review process because of the lack of experimental research on this topic. Therefore, for this review, a conscious decision was made to relax these standards and include all correlational studies that employed statistical controls. By relaxing the WWC standards, the number of candidate studies to be screened in Phase III was greatly increased. A more streamlined and efficient checklist was developed requiring dichotomous responses rather than descriptive responses in order to facilitate rater decision making. To be included in the quantitative review, a study had to use appropriate measurement and statistical principles. A primary criticism of previous single-sex literature has been the confounding of single-sex effects with the effects of religious values, financial privilege, selective admissions, or other advantages associated with the single-sex school being studied. Therefore, in particular a study had to include statistical controls to account for individual differences (e.g., socioeconomic status [SES], individual ability, and age) as well as school and class differences that might account for the differences between single-sex and coeducational schools. Even so, many studies that included at least one covariate lacked other important covariates such as ethnic or racial minority status, socioeconomic status, and grade level or age. Also, it is important to note that the inclusion of covariates cannot control for important unobservable differences between the groups, such as motivation. Because correlational studies cannot adequately address the issue of differences in unobservables (or selection bias), the studies in this review may over or understate the true effects of SS schooling.

THE QUANTITATIVE REVIEW

Two reviewers coded each study independently, using a quantitative coding guide. A quantitative study was coded for its treatment of the following broad issues: sample characteristics, psychometric properties, internal validity, effect, and bias. Each of these categories had several criteria by which they were coded. To be retained, a study did not have to meet all criteria.

Of the 88 quantitative studies, 48 were eliminated after further review using the coding guide, and 40 studies met the inclusion criteria and were retained. The reasons for the exclusion of these articles were 1) failure to operationalize the intervention properly; 2) failure to apply statistical controls during the analyses; 3) work that was actually qualitative in nature rather than quantitative; 4) work performed in a non-Westernized country and therefore not comparable; 5) work written in a foreign language and therefore not codable by the researchers; 6) failure to draw comparisons between SS and CE schools; and 7) participants not of high school, middle, or

elementary school age. In all, 40 studies met the inclusion criteria and were retained in the quantitative review. The following table shows results of each study according to the seven broad questions listed above and is broken into specific criteria within each larger category. Because some studies addressed multiple criteria, the total number of findings is greater than 40. Specifically, there are 112 findings considered in the 40 quantitative studies.

A table summarizing the findings is below. In each row, one of the 32 outcome categories is listed, as well as the total number of studies related to that outcome category and the raw number and percent of findings that either support SS schooling, support CE schooling, are null, or mixed (supporting both CE and SS schooling). While eight of the outcome categories have four or more studies, others have as few as one or two studies. For any outcome category, the percentage of studies falling in any of the dispositions (supporting SS, supporting CE, null, or mixed) and the confidence with which one can use the findings will increase with the number of studies. Therefore, the percentages in the summary table should be treated with caution when only one or two studies appear for that outcome category.

As shown in the Summary Table, every study falls into one of four categories: Pro-SS, Pro-CE, Null, or Mixed. If a study's findings all supported SS schooling for a given outcome variable, it was coded as "Pro-SS". If the study's findings all supported CE for a given outcome variable, it would be coded "Pro-CE". A study was coded "Null" if for all findings regarding that outcome variable, there were no differences between the SS and CE schools. A study was coded "Mixed" if the study had significant findings in opposite directions for different subgroups on the same variable. For example, a study would be coded "Mixed" if on a specific outcome, support was found for single-sex schooling in the case of boys and support was found for coeducation in the case of girls. Another example would be a finding favoring single-sex in a 10th-grade sample and coeducation in a 12th-grade sample for the same outcome variable within a single study. If a study had findings that were both pro-SS and null, it was coded a pro-SS; if the study had findings that were both pro-CE and null, it was coded as pro-CE. Only studies with findings favoring both single-sex and coeducation were coded as mixed. It should also be kept in mind that some researchers evaluated multiple outcome variables in their research; therefore, it is possible that a single published study would yield information that appears in multiple rows of the Summary Table.

Table 1—Summary of Findings

Outcome	Total Number of Studies	Pro-SS		Pro-CE		Null		Mixed	
		Number of Studies	Percent	Number of Studies	Percent	Number of Studies	Percent	Number of Studies	Percent
Concurrent Academic Accomplishment									
1) All-Subject Achievement Test Scores	9	6	67%	1	11%	2	22%	0	0%
2) Mathematics Achievement Test Scores	14	3	22%	0	0%	8	56%	3	22%
3) Science Achievement Test Scores	8	2	25%	0	0%	5	62%	1	13%
4) Verbal/English Achievement Test Scores	10	3	30%	0	0%	7	70%	0	0%
5) Grades	1	0	0%	0	0%	1	100%	0	0%
6) Social Studies Achievement Test Scores	1	1	100%	0	0%	0	0%	0	0%
Subtotal	43	15	35%	1	2%	23	53%	4	10%
Long-Term Academic Accomplishment									
7) Postsecondary Test Scores	2	1	50%	0	0%	1	50%	0	0%
8) College Graduation	1	0	0%	0	0%	1	100%	0	0%
9) Graduate School Attendance	1	0	0%	0	0%	1	100%	0	0%
Subtotal	4	1	25%	0	0%	3	75%	0	0%
Concurrent Adaptation and Socioemotional Development									
10) Self-concept	7	4	57%	0	0%	3	43%	0	0%
11) Self-esteem	6	1	17%	2	33%	3	50%	0	0%
12) Locus of Control	5	3	60%	0	0%	2	40%	0	0%
13) School Track/Subject Preference	14	5	36%	2	14%	6	43%	1	7%
14) Educational Aspirations	3	2	67%	0	0%	1	33%	0	0%
15) Career Aspirations	2	2	100%	0	0%	0	0%	0	0%
16) Delinquency	4	2	50%	0	0%	2	50%	0	0%
17) Attitudes Toward School	5	1	20%	1	20%	1	20%	2	40%
18) Time Spent per Week on Homework	2	1	50%	0	0%	1	50%	0	0%
19) Attitudes Toward Working Women	1	1	100%	0	0%	0	0%	0	0%
Subtotal	49	22	45%	5	10%	19	39%	3	6%

Continued on next page

Table 1—Summary of Findings (cont'd)

Outcome	Total Number of Studies	Pro-SS		Pro-CE		Null		Mixed	
		Number of Studies	Percent	Number of Studies	Percent	Number of Studies	Percent	Number of Studies	Percent
Long-term Adaptation and Socioemotional Development									
20) School Completion	1	1	100%	0	0%	0	0%	0	0%
21) Postsecondary Success	1	0	0%	0	0%	1	100%	0	0%
22) Postsecondary Unemployment	2	1	50%	0	0%	1	50%	0	0%
23) Eating Disorders	1	0	0%	1	100%	0	0%	0	0%
24) Choice of College Major	1	1	100%	0	0%	0	0%	0	0%
25) Sex-Role Stereotyping	2	1	50%	1	50%	0	0%	0	0%
26) Political Involvement	1	1	100%	0	0%	0	0%	0	0%
27) Percent Married to First Spouse	1	0	0%	0	0%	1	100%	0	0%
Subtotal	10	5	50%	2	20%	3	30%	0	0%
Perceived School Culture									
28) Climate for Learning	1	1	100%	0	0%	0	0%	0	0%
29) Opportunities for Leadership Roles	2	1	50%	0	0%	1	50%	0	0%
30) School Environment	1	0	0%	0	0%	1	100%	0	0%
Subtotal	4	2	50%	0	0%	2	50%	0	0%
Subjective Satisfaction									
31) Satisfaction with School Environment	1	0	0%	1	100%	0	0%	0	0%
32) College Satisfaction	1	1	100%	0	0%	0	0%	0	0%
Subtotal	2	1	50%	1	50%	0	0%	0	0%
TOTALS	112	46		9		50		7	

Implications of Review

Summary of Findings in Each Domain

Concurrent, quantifiable academic accomplishments

In general, most studies reported positive effects for SS schools on all-subject achievement tests. Studies examining performance on mathematics, science, English, and social studies achievement tests found similar findings with one caveat. Within each of these subject-specific categories, roughly a third of all studies reported findings favoring SS schools, with the remainder of the studies split between null and mixed results. This minimal to medium support for SS schooling applies to both males and females and in studies pertaining to both elementary and high schools. The overall picture is split between positive findings for SS schooling and no differences or null findings, with little support for CE schooling. The one study that found advantages for CE schooling found advantages for white females but not for Asian or black females. Males continue to be underrepresented in this realm of research.

Long-term, quantifiable academic accomplishment

As opposed to concurrent indicators of academic achievement, any positive effects of SS schooling on longer-term indicators of academic achievement are not readily apparent. No differences were found for postsecondary test scores, college graduation rates, or graduate school attendance rates. However, all the findings in this domain came from a pair of studies, indicating the lack of high-quality research on these important criteria. Although some studies favor single-sex education in the case of postsecondary test scores, there is a dearth of recent studies using controls. There has been a similar lack of research on other potential criteria in this domain, such as college grade point average, meritorious scholarships or funding attained, postgraduate licensure test scores, and any career achievement that could ostensibly be tied to quality of schooling.

Concurrent, quantifiable indicators of individual student adaptation and socioemotional development

This category includes a range of outcomes that are not easily grouped together, and the results are mixed. Regarding self-concept and locus of control, the studies are split between those showing positive effects for SS schooling and those showing no differences. In the case of self-esteem, a third of the studies supported CE schooling while half found no difference. Given a recent extensive review concluding that self-esteem's relationship to school success, occupational success, better relationships, leadership, delinquent behavior, and other desirable outcomes is modest to nonexistent, the implications of findings regarding self-esteem appear complementary. Furthermore, CE schooling only had a positive impact on the self-esteem of males.

Findings regarding school track and subject preferences were mixed, with the overall weight of the findings lying somewhere between pro-SS findings and no differences. A majority of studies favored SS schools on the outcome of higher educational aspirations, as evidenced by SS students showing more interest in and taking more difficult courses. SS schools fostered higher educational and career aspirations for girls. More studies emphasized the positive effect of SS schools on career aspirations than CE schools for boys, but evidence regarding their educational aspirations was mixed.

A category called “attitudes toward school” showing mixed results was actually a combination of single studies using somewhat different outcome variables, thus reducing the meaningfulness of the category. In terms of actual behaviors, a few studies focused on delinquency, reporting differences in favor of SS schools that were moderated by individual developmental differences. What is lacking is a conceptual framework to tie together the myriad academic-attitude outcome measures used in this realm so that studies will be more directly comparable.

Long-term, quantifiable indicators of individual student adaptation and socioemotional development

The outcomes in this domain generally do not appear in more than one or two studies that made it to Phase III review. Therefore, one must be cautious in generalizing from these results. Having said that, the results still suggest the potential that SS schooling could be associated with a number of post-high school, long-term positive outcomes. These include postsecondary success or participation in collegiate activities while maintaining full-time enrollment for a four-year period, reduced unemployment (males and females), reduced propensity to drop out of high school (males and females), the choice of a nontraditional college major (for females), and political activism (for females). The sole exception is eating disorders; one study found more SS students to have eating disorders than CE students.

Procedural (e.g., classroom treatment) and outcome measures of gender inequity

This question could not be addressed because of a lack of any quantitative studies that used gender equity as an outcome variable at the school level. Any studies that compared SS and CE classrooms within a CE school were outside the purview of this study and were not reviewed.

Perceptual measures of the school climate or culture that may impact performance

This category includes a number of disparate, single-study results. One of the two studies addressing leadership opportunities found more opportunities for both males and females in SS schools; however, the statistical significance of this finding depended on what other variables had been controlled for. The other found that both males and females in SS schools put more value on grades and leadership and less on attractiveness and money. However, there remains a dearth of high-quality empirical studies using this class of outcome variables as criteria.

A final category of outcomes examined as a subset of culture was the realm of subjective satisfaction of students, parents, and teachers with the school environment. The one study in this review that found the social environment more appealing in CE schools is a good case in point in that the same study found that SS students are more interested in grades and leadership and less interested in money and looks. Some qualitative studies have looked at why certain parents prefer SS schooling, and studies in other cultures have found mixed results regarding teacher satisfaction with CE versus SS schooling. However, no empirical studies comparing current parental satisfaction in equivalent SS and CE schools were available for review using the stated guidelines. There remains a lack of research both on this class of criteria and on the relationship of subjective satisfaction to other more critical criteria.

Expected Outcomes Not Seen in the Review

Teenage pregnancy, college performance, differential treatment by teachers, parental satisfaction, bullying in school, and teacher satisfaction were among the many outcomes that we expected to see in the review or that should be addressed but were not found in any included study.

General Trends

A few trends are apparent across all outcomes. The preponderance of studies in areas such as academic accomplishment (both concurrent and long term) and adaptation or socioemotional development (both concurrent and long term) yields results lending support to SS schooling. A limited number of studies throughout the review provide evidence favoring CE schooling. It is more common to come across studies that report no differences between SS and CE schooling than to find outcomes with support for the superiority of CE. In terms of outcomes that may be of most interest to the primary stakeholders (students and their parents), such as academic achievement test scores, self-concept, and long-term indicators of success, there is a degree of support for SS schooling.

The overwhelming majority of studies employ high school students, with a small minority using elementary school students. The preponderance of SS research has been conducted in Catholic SS schools in which students are separated by sex only when entering adolescence. Therefore, opportunities to study SS elementary or middle schools in either the public or private sector have been limited.

There is also a pronounced tendency to study girls' schools more than boys' schools: 76 studies compared SS and CE girls, and 20 of those focused exclusively on girls. Of those 20, 18 were split evenly between support for SS schooling and no differences (nine pro-SS and nine no differences). The other two studies resulted in findings supporting CE schooling. SS and CE schooling for boys was compared in 55 studies, of which only three were studies exclusively devoted to boys' schools.

There is a dearth of quality studies (i.e., randomized experiments or correlational studies with adequate statistical controls) across all outcomes. Even using the more relaxed criterion of allowing correlational studies, each outcome has only limited candidate studies. Too few researchers report descriptive statistics or effect sizes. Mathematics achievement test scores, English achievement test scores, and school subject preference were the only outcomes to have 10 or more qualifying studies. Even within these three categories, the studies differ in the criteria they use and the statistical controls they use to compare SS and CE schooling. This somewhat limits the arguments that can be built and extended from this quantitative review and renders it nearly impossible to conduct a meta-analysis on any outcome area. Many of the remaining studies have other conceptual or interpretive flaws. Many of the studies lacked well-developed hypotheses, and the hypotheses were often not linked directly to the outcomes being studied.

The list of outcomes of interest needs to be expanded in future research and defined more clearly. For example, outcomes such as teenage pregnancy and bullying in school did not appear in a single study of sufficient quality to be reviewed. Other outcomes that are implicit in arguments for and against SS schooling need to be addressed explicitly. These include work-related long-term outcomes such as job performance, leadership performance, mixed-sex work team performance, performance and leadership in volunteer associations, job involvement, and organizational commitment. Few studies

address important moderators, that is, variables that may have differential effects for single-sex schooling. For example, a number of authors have proposed that SS schools are particularly effective for students of lower socioeconomic status and perhaps specifically for those who are members of minority or disadvantaged communities. Unfortunately, only three studies addressed this moderator.

This review should not be interpreted as a condemnation of the work of the dedicated researchers who have chosen to study SS-CE differences, as they may not have been in a position to conduct a randomized experiment on this topic. Such a study has yet to be conducted. However, it could be argued that instead of trying to conduct only all-or-nothing studies of whether SS schooling is better or worse than CE schooling, more careful specification of hypotheses and direct linkage of hypotheses to specific outcomes may show ways to also conduct smaller studies that prove whether certain aspects of SS or CE schooling are beneficial.

Finally, there are limits to what a systematic review can accomplish when an intervention is being judged by multiple criteria and all stakeholders do not share the hierarchy of these criteria. Some issues cannot be resolved by any type of research, even randomized experiments, because they involve issues of philosophy and worldview and represent the relative priorities of dueling stakeholders. There is no way to resolve whether an outcome that is important to one stakeholder group, such as parents, students, civil libertarians, and feminists on both sides of the issue, should be accorded more weight than an outcome valued by another group. What is possible is to separate out fact in the form of evidence from fiction by converting as many claims as possible to testable hypotheses and performing the necessary research. In this way, the two parallel debates can be separated from each other. “Does SS schooling benefit or harm the students, and in what ways?” can be separated from “Is it worth it for society regardless of the benefits or costs?” with each debated on its own merits.

These general implications of the review provide a stepping-stone for future research through the continuation of quality research on extant outcomes, the refinement of methodology, better statistical reporting, and the expansion of the theoretical domain. If heeded, these implications can improve the generalizations made about single-sex schooling and coeducation.