

REVIEW PROTOCOL FOR CHARTER SCHOOLS VERSION 3.0 (SEPTEMBER 2016)

This review-specific protocol guides the review of research that informs the What Works Clearinghouse (WWC) Charter Schools topic area. The review-specific protocol is used in conjunction with the [WWC Procedures and Standards Handbook \(version 3.0\)](#).

PURPOSE STATEMENT

This review focuses on charter school interventions in the United States with students in grades PK–12 with a primary focus on student academic achievement and other outcomes.

The following research questions guide this review:

- What is the impact of charter schools on their students' academic achievement in language arts, mathematics, and other subjects, as measured by test scores?
- What is the impact of charter schools on their students' other outcomes, including attendance, experience of disciplinary actions (suspensions or expulsions), social-emotional competence, educational attainment (for example, high school graduation, college entry, and degree completion) and earnings in adulthood?
- What is the systemic impact of the expansion of the charter sector in a community on student achievement and educational attainment outcomes for students in traditional public schools?

KEY DEFINITIONS

Charter school. A charter school is a publicly funded school that is operated autonomously, outside the direct control of the local school district, by a group or organization under a contract (or charter) with the charter's authorizer (which may be the local school district, a state agency, or another state-designated authorizing organization), and that is typically enrolled by student/parent choice rather than by neighborhood assignment. The charter exempts the school from certain state or local rules and regulations. In return for flexibility and autonomy, the charter school must meet the accountability standards outlined in its charter. A school's charter is reviewed periodically (typically every 3 to 5 years) by the group or jurisdiction that granted it and can be revoked if guidelines on curriculum and management are not followed or if the standards are not met.

Charter Management Organization (CMO). A CMO is an organization that manages two or more charter schools.

PROCEDURES FOR CONDUCTING THE LITERATURE SEARCH

The *WWC Procedures and Standards Handbook* discusses the procedures for conducting a literature search in Section II: Developing the Review Protocol and Identifying Relevant Literature (p. 4) and in Appendix B: Policies for Searching and Prioritizing Studies for Review.

Search Terms

The following table presents the search terms by category.

Category	Search Terms
Study Design	<ul style="list-style-type: none"> • ABAB design* • Alternating treatment* • Assignment • Baseline • Causal • Changing criterion design* • Comparison group* • Control group* • Control* • Counterfactual* • Effect* • Efficacy • Evaluat* • Experiment* • Field trial* • Growth • Impact* • Intrasubject replication design* • Matched group* • Meta analys* • Metaanalys* • Meta-analys* • Multi element design* • Multi-element design* • Multiple baseline design* • Outcome* • Posttest • Post-test • Pretest • Pre-test • QED • Quasi experimental design • Quasi experimental study • Quasiexperiment* • Quasi-experiment* • Random control* trial* • Random* • Random* assign* • Randomized controlled trial* • RCT • RDD • Regression discontinuity • Reversal design* • SCD • Simultaneous treatment* • Single case design* • Single subject design* • Studies • Study • Treatment* • Withdrawal design*

Category	Search Terms	
Intervention (Broadly)	<ul style="list-style-type: none"> • Charter conversion* • Charter Management Organization* • Charter school* • Charter sector* • Charter transformation* • Charter turnaround* • Charter Network • Recovery School District* 	
Population	<ul style="list-style-type: none"> • Elementary school* • High school* • Junior high • K–5 • K–6 • K–8 • K–12 • Middle school* • PK–5 • PK–6 • PK–8 • PK–12 • PreK–5 	<ul style="list-style-type: none"> • PreK–6 • PreK–8 • PreK–12 • Pre-K–5 • Pre-K–6 • Pre-K–8 • Pre-K–12 • Primary school* • School* • School-aged • School-based • Secondary school*
Outcomes	<ul style="list-style-type: none"> • Achiev* • Dropout • Graduation • Highest grade completed • School persistence • Suspension • Non-cognitive • Student retention • Earnings • Employment • College enroll* • Attendance • Expulsion • Social-emotional 	

Additional Sources

In addition to those listed in the *WWC Procedures and Standards Handbook*, Appendix B, this review searched the following electronic database:

- **Campbell Collaboration.** C2-SPECTR (Social, Psychological, Educational, and Criminological Trials Register) is a registry of over 10,000 randomized and possibly randomized trials in education, social work and welfare, and criminal justice.

In addition to those listed in the *WWC Procedures and Standards Handbook*, Appendix B, this review searched the following websites:

- Bill & Melinda Gates Foundation
- Consortium for Policy Research in Education
- CREDO at Stanford University
- Institute of Education Sciences

- National Bureau of Economic Research
- National Center for Analysis of Longitudinal Data in Education Research
- RTI International
- University of Chicago Consortium on Chicago School Research
- Westat
- WestEd

ELIGIBILITY CRITERIA

Eligible Populations

In this review, the following populations are of interest:

- **Grade range.** Schools must provide instruction to students in grades PK–12.
- **Location.** Schools must be located within the United States, its territories, or tribal entities.

Potential subgroups of interest for this review include:

- Characteristics of students:
 - Special education status
 - English learner status
 - Gender
 - Race
 - Ethnicity
 - Economically disadvantaged (e.g., free or reduced-price lunch status)
 - Grade
- Characteristics of school/classroom settings:
 - Location of teaching setting (e.g., urban, suburban, rural)
 - Charter type or operator
 - School level (e.g., elementary, middle, high)
 - School size
 - Economically disadvantaged (e.g., Title I status, percentage of students eligible for free or reduced-price lunch)

Eligible Interventions

In this review, the following types of interventions may be included:

- **Policies.** A *policy* is a named condition, system, or set of formal rules that affect schools and students. The policy must be commonly understood in the field and

literature. Policies may be set by federal, state, or local governments or by the organization providing services. Examples of policies of interest include:

- Legalization of charter schools, or
- Requirements for the granting or renewal of charter school authorizations.
- **Programs.** A *program* is a school or group of schools implementing a charter school model. Examples of interest include both independent charter schools and groups of charter schools organized into CMOs or charter networks. Individual charter schools are eligible interventions. However, a study of the effectiveness of an individual charter school is not eligible to be included in a review for evidence of the effectiveness of a named CMO or charter network because the named intervention is bundled with other programs and policies implemented within the individual school.

Eligible Research

The *WWC Procedures and Standards Handbook* discusses the types of research reviewed by the WWC in Section II: Developing the Review Protocol and Identifying Relevant Literature (p. 4). In this review, the following additional parameters define the scope of research studies to be included:

- **Topic.** The study must examine the effect of charter schools on an eligible outcome (student achievement, social-emotional competence, disciplinary experiences, student attendance, student progression, and earnings in adulthood).
- **Time frame.** Studies must have been released or made public in 1996 or later and be obtained by the WWC for review prior to the drafting of the intervention report.
- **Sample.** The study sample must meet the requirements described in the “Eligible Populations” section above.
- **Language.** The study must be available in English to be included in the review.
- **Location.** The study must include schools in the United States, its territories, or tribal entities.

Eligible Outcomes

This review includes outcomes in the following domains:

English language arts achievement. Includes outcomes in the following areas: foundational reading (word reading, fluency and/or accuracy in reading connected text, vocabulary, reading comprehension), general reading, measures of English language conventions (e.g., grammar),

writing, and general English language arts achievement (i.e., a standardized test covering an array of language arts topics).

Mathematics achievement. Includes outcomes in the following areas: understanding of different subjects within mathematics, including algebra, arithmetic, calculus, geometry, probability, statistics, and trigonometry; understanding of concepts and procedures; understanding of word problems and applications; and general math achievement (i.e., a standardized test covering an array of mathematics topics).

Science achievement. Includes outcomes in any of the physical or life science disciplines, such as biology, chemistry, earth science, general science, and physics.

Social studies achievement. Includes outcomes in social studies subdisciplines, such as civics, economics, geography, history, and world cultures.

General achievement. Includes a general measure of student academic achievement, only to be documented if study authors do not distinguish students' achievement in specific areas (e.g., math, reading). Examples include composite scores from state assessments that represent a combination of reading and math scores.

Course grades, teacher reports of proficiency in the different subject areas, and assessments unrelated to academic achievement are not eligible outcome measures.

Social-emotional competence. Involves (a) self-awareness of thoughts, feelings, and behaviors; (b) social awareness of context and others; (c) social skills needed to initiate, sustain, and respond to social interactions; (d) relationship skills needed to establish and sustain social connections; (e) self-regulation needed to make responsible decisions, and manage personal behavior to achieve goals; and (f) self-perceptions associated with functional behavior. This includes social engagement (with adults or peers), self-management, adaptive functioning, and well-being.

Disciplinary experiences. Includes measures of disciplinary action such as suspensions and expulsions.

Student attendance. Includes measures of daily student attendance at school.

Student progression. Includes measures of students' progression in school. Constructs include:

- **Student promotion** (e.g., students' advancement to next grade level)
- **Student graduation** (e.g., students' completion of the PK–12 education system). Measures of graduation that assess whether the student graduated from a school or group of schools defined by the intervention or comparison group only are not

eligible, but measures that define graduation more broadly (from a school in the district, for example) would be eligible.

- **Earning a GED**
- **College matriculation** (e.g., students' enrollment and attendance in the postsecondary education system)
- **College graduation**

Earnings in adulthood. Includes measures of students' wage rates or earnings following completion of schooling.

Eligible student outcomes may be measured at the student, classroom, or school level. For example, the percentage of grade 12 students in a school who graduate is an eligible school-level outcome in the student progression domain. Similarly, student gain scores in math aggregated to the classroom or school level both would fall in the mathematics achievement domain.

DESIGN STANDARDS

Eligible studies are assessed against WWC design standards, as described in the *WWC Procedures and Standards Handbook* Section III: Screening and Reviewing Studies (pp. 8–21).

Sample Attrition

The *WWC Procedures and Standards Handbook* discusses the sample attrition standards used by the WWC in Section III: Subsection B.2 Sample Attrition: Is the combination of overall and differential attrition high? (pp. 11–15).

This review uses the **conservative** boundary for attrition. The decision to use this boundary was based on recognizing that selective attrition and enrollment in charter schools is an important topic of methodological debate within the literature on charter school effectiveness and in policy debates around charter school effectiveness as well. For example, it is often claimed that charter schools will push out low-performing or disruptive students, giving them an advantage over traditional public schools that does not reflect their underlying effectiveness and biasing studies of student outcomes. The *WWC Procedures and Standards Handbook* contains a figure illustrating the attrition boundary and an associated table with attrition levels that define high and low attrition. Based on the choice of the boundary, the study review guide calculates whether the combination of overall and differential rates of attrition results in a high or low level of potential bias.

Baseline Equivalence

Quasi-experimental designs, randomized controlled trials with high levels of attrition or compromised random assignment, and regression discontinuity designs with high levels of attrition must demonstrate baseline equivalence of the intervention and comparison groups for the analytic sample. The onus for demonstrating equivalence in these studies rests with the authors. The *WWC Procedures and Standards Handbook* discusses how authors must demonstrate baseline equivalence in Section III: Subsection B.3 Baseline Equivalence: Is equivalence established at baseline for the groups in the analytic sample? (pp. 15 and 16).

Baseline equivalence must be demonstrated for the intervention and comparison groups in the analytic sample on the following pre-intervention (or baseline) characteristics:

- A pre-intervention measure of the outcome; or,
- If a pre-intervention measure is not available, an acceptable alternative pre-intervention measure can be used for some outcomes, as explained below.

Acceptable measures for student achievement outcomes. For outcomes in the student achievement domains, studies must show that the groups are equivalent on an acceptable pre-intervention measure of student achievement. A pretest measure in the same subject as the outcome is preferred; however, if a same-subject pretest is not available, a pretest measure of general achievement (e.g., a combined mathematics and reading score) would be acceptable. In addition, a pretest measure of mathematics achievement can be used to establish baseline equivalence for a science achievement outcome, and a pretest measure of English language arts achievement can be used for a social studies achievement outcome.

Acceptable measures for social-emotional outcomes. For outcomes in the social-emotional competence domain, studies must show that the groups are equivalent on a pre-intervention measure in the social-emotional competence domain.

Acceptable measures for some non-achievement outcomes (disciplinary experiences, student attendance, student progression, and earnings in adulthood outcomes). For outcomes in these domains, if a pre-intervention measure of the outcome is unavailable, studies must show that groups are equivalent on the following set of baseline characteristics.

- Grade level; **AND**
- One of the following measures of student academic performance: standardized test scores, whether behind in grade level (could be measured by age among students in the same grade), prevalence of school behavior or discipline issues, rate of school attendance, or grade point average (GPA); **AND**
- One of the following: student race/ethnicity or a measure of degree of disadvantage (i.e., free or reduced-price lunch status, poverty status, family income, English learner status, special education status, or disability status).

When equivalence is required and the study includes two or more baseline measures listed above for an outcome domain, this review protocol follows the default approach described in the *WWC Procedures and Standards Handbook*. First, a difference larger than 0.25 standard deviations for any required baseline measure in a domain for a given analytic sample means that all findings for outcomes in the domain based on the same analytic sample do not meet WWC group design standards. However, findings for outcomes in different domains or using different analytic samples are still eligible to meet WWC group design standards if equivalence is demonstrated.

Also, analyses of all outcomes in a domain using the same analytic sample must include statistical adjustments for all required baseline measures in the domain that require adjustment (i.e., have baseline differences that fall between 0.05 and 0.25 standard deviations for the same analytic sample). The statistical adjustments must be performed using an acceptable approach for the WWC. For example, consider outcomes A, B, and C in the mathematics achievement domain, each also measured at baseline. If A and C have baseline differences less than 0.05 standard deviations, while B has a baseline difference in the adjustment range, the analyses of outcomes A, B, and C must all adjust for the baseline measure of B (but not A or C).

Cluster designs: If the unit of assignment is a cluster, such as a school (as in studies of systemic effects of charter schools on students in conventional public schools), AND the study makes consistent cluster-level inferences, then a pre-intervention cluster-level measure of the outcome is required for establishing baseline equivalence. Otherwise, the requirements for individual-level assignment studies apply.

Level of assignment in QEDs: When the intervention is participation in a charter school or program, then the unit of assignment is the student; when the intervention is a systemic effect of charter competition, then the unit of assignment is a cluster (such as a school).

A review should clearly document if a study has a baseline difference in any of the following characteristics, since it could be evidence that the populations were drawn from different settings and that the intervention and comparison groups are not sufficiently comparable for the purposes of this review:

- Socioeconomic status (SES)
- Race
- Ethnicity
- School location (e.g., urban, rural)

The provision of such information, however, is not a requirement of the review.

Outcomes

The *WWC Procedures and Standards Handbook* discusses the types of outcomes, criteria the outcome must meet, and how outcomes are reported by the WWC in Section III: Subsection B.4 Outcome Eligibility and Reliability (pp. 16–19). This review follows the general guidance regarding reliability, outcomes measured at different points in time, impacts measured at different points in time, composite and subscale scores, subgroup findings, categorical ordinal measures, and findings that account for missing data.

Statistical Adjustments

The *WWC Procedures and Standards Handbook* discusses the types of adjustments made by the WWC in Section IV: Subsection B Statistical Significance of Findings (p. 24).

Other Study Designs

Studies that use regression discontinuity designs are eligible for review using the pilot standards for reviewing regression discontinuity design studies in Appendix D of the *WWC Procedures and Standards Handbook*.