

# REVIEW OF INDIVIDUAL STUDIES PROTOCOL

## VERSION 3.1 (April 2020)

### PURPOSE STATEMENT

What Works Clearinghouse™ (WWC) reviews are designed to provide education practitioners and policymakers with timely and objective assessments of the quality of research evidence. WWC reviews focus primarily on studies of the effectiveness of education or school-based interventions serving students in early childhood education through 12th grade, as well as students in postsecondary settings. However, in some instances, they might focus on studies of interventions serving other groups, such as teachers or other school leaders.

All WWC reviews are governed by a review protocol that describes the procedures that will be followed when conducting the review. When a study has been identified through a literature search defined by a WWC protocol, such as a topic area protocol for preparing WWC intervention reports or a protocol for preparing a WWC practice guide, the study review is guided by that protocol. However, the WWC also reviews studies that have been identified through other means and that may not fall under the umbrella of a more specific review protocol. This protocol guides those reviews of individual studies. This review-specific protocol is used in conjunction with the *WWC Procedures and Standards Handbook* (version 3.0).

### ELIGIBILITY CRITERIA

#### Eligible Populations

To be eligible for review under this protocol, the study must examine the effectiveness of an intervention administered to:

- students and other learners in early childhood education through postsecondary education, including learners of any age from infancy through adulthood;
- teachers and related educators; or
- school leaders.

#### Eligible Interventions

The study must examine an educationally relevant or school-based intervention. The WWC defines the term “intervention” broadly, and this term can include education practices, products, programs, and policies. Therefore, the following types of interventions may be included:

- **Practices.** Education practices are discrete, clearly defined activities focused on improving student learning and related outcomes. Practices may be used with a wide range of participants to address a wide range of learning goals – or they may be targeted to address a specific learning goal, skill, or population.
- **Products.** Education products are “branded” or commercial interventions such as curricula or software. Products may be used as the primary instructional tool in the classroom or to supplement classroom material with differentiated instruction, remediation, or enrichment. Products may possess a trademark or copyright and may be

supported by a developer who provides technical assistance or sells or distributes the intervention.

- **Programs.** Education programs are combinations of practices or products that are intended to either directly or indirectly improve the outcomes of students. These can include, for example, after-school programs and mentoring programs.
- **Policies.** Education policies involve structural changes that are intended to either directly or indirectly improve the outcomes of students. Examples of education policies include modifying the academic calendar and changing the number of credits required for graduation.

### Eligible Research

The *WWC Procedures and Standards Handbook (version 3.0)* lists and describes the types of research reviewed by the WWC. Under this protocol, the following additional parameters define the scope of research studies to be included:

- **Timeframe.** Studies must have been released or made public within the 20 years preceding the year of the review—for example, in 2000 or later for reviews occurring in 2020.
- **Language.** The study must be available in English to be included in the review.
- **Location.** The study must include students in the United States, in its territories or tribal entities, at U.S. military bases overseas, or in Organisation for Economic Co-operation and Development (OECD) member countries in which English is the primary or most commonly used language; that is, Australia, Canada, Ireland, New Zealand, and the United Kingdom.

### Eligible Outcomes

Using this protocol, the WWC reviews research findings that could be valuable to policy makers, administrators, practitioners, or parents who are deciding whether a specific intervention should be adopted or scaled up. Outcomes drawn from the table below are considered eligible outcomes for reviews conducted under this protocol.

When the study reports both immediate and longer-term measures of an outcome, the WWC selects one measure as the primary finding that will contribute to the rating for the intervention; findings for the other outcomes will be included in supplemental tables. The preference is determined by the review team leadership, with exceptions denoted in the table below.

**Exhibit 1—Outcome Domains**

Domain	Description
<b>Standardized Measures of Academic Readiness, Knowledge, or Skills</b>	
Cognition	Memory, problem solving, cognitive processing and flexibility, general knowledge, and IQ

<b>Domain</b>	<b>Description</b>
Communication/ language competencies	Ability to communicate with other people; ability to engage in conversation, request help, or follow prompts or instructions. The use of communication devices is allowed for these competencies. Note that outcomes for English Learners are included in the “English language proficiency” domain.
English language proficiency	Vocabulary including receptive vocabulary, expressive vocabulary, and academic language; oral language including listening and speaking skills; listening comprehension; and grammar including syntax and morphology in spoken or written language. Note that this domain is for majority English Learner samples.
Alphabetics	Letter identification, phonemic awareness, phonics, phonological awareness, print awareness
Comprehension	Vocabulary development and reading comprehension. Note that the “English language proficiency” domain should be used for majority English Learner samples.
Reading fluency	Reading text accurately, automatically, and with expression
Writing conventions	Using rules of standard English language, such as word usage, sentence structure, or grammar
Writing productivity	Measures of writing quantity, such as counts of written words, sentences, or ideas
Writing quality	Writing effective, clear, well-organized text
General literacy achievement	Content in two or more of the above literacy domains
Algebra	Ability to solve, graph, or write equations, systems of equations, and inequalities, as well as functions, exponents, polynomials, factoring, and quadratic equations
Data analysis, statistics, and probability	Collecting, organizing, and displaying data to answer questions, as well as statistical methods to analyze data, making inferences and predictions based on data, and probability
Geometry and measurement	Two-dimensional and three-dimensional geometric shapes and understanding properties, composition, and geometric relationships, including visualization, spatial reasoning, and geometric modeling, as well as understanding the attributes, units, systems, and processes of measurement, and applying techniques, tools, and formulas to determine measurements
Number and operations	Understanding numbers and integers, such as subitizing, estimation, number order, number combinations, counting, comparisons, operations, computing fluently, representing fractions and ratios and understanding the base-ten number system and fractions
Calculus	Differential calculus (concerning rates of change and slopes of curves) or integral calculus (concerning accumulation of quantities and the areas under and between curves)

<b>Domain</b>	<b>Description</b>
Trigonometry/ precalculus	Trigonometry topics such as relationships involving lengths and angles of triangles, and/or precalculus topics such as functions, complex numbers, vectors, and matrices
General mathematics achievement	Content in two or more of the above mathematics domains, as well as tests of mathematical understanding, procedures, and problem solving designed to measure more than one content area
Life sciences	The structures and functions of living things at different scales; growth, development, and reproduction of organisms; information processing and behavior in organisms; matter and energy transfer in living things and ecosystems; inheritance of and variation in traits; natural selection and adaptation; evidence of common ancestry; biodiversity
Physical sciences	The properties of matter and changes in matter; force, motion, and interactions of forces; energy and energy transfer and conservation; relationship between energy and forces; properties of waves; electromagnetic radiation
Earth/space sciences	The structures, properties and materials of Earth; tectonics; Earth's place in the solar system and the universe; changes in Earth over time; water, weather and climate; energy in Earth systems; bio-geology
General science achievement	Content in two or more of the above science domains, as well as knowledge of science practices, such as forming hypotheses and making predictions, control of variables, and planning and conducting experiments and observations
General social studies achievement	Outcomes in social studies disciplines, such as one or more of civics, economics, geography, history, and world cultures
Primary school academic achievement <sup>1</sup>	Academic measures based on student test scores across multiple subjects, including at least two of literacy, mathematics, science, and social studies
Secondary school academic achievement <sup>2</sup>	Mastering academic content, as measured by standardized achievement tests such as the ACT and SAT and state-mandated tests, or by secondary school grade point averages (if the formula is specified). Note that individual course grades or exam scores from secondary school courses are not eligible under this domain.

---

<sup>1</sup> "Primary school academic achievement" pertains to students in grades K-5, or to learners ages 5 through 11. This domain should only be selected for test score measures that *do not* correspond with a more specific domain.

<sup>2</sup> "Secondary school academic achievement" pertains to students in grades 6-12, or to learners age 12 and above who have not yet completed a high school diploma or a recognized equivalent, or who have completed a high school diploma or a recognized equivalent, but either are English Learners or lack basic skills such as literacy. This domain should only be selected for test score measures that *do not* correspond with a more specific domain.

<b>Domain</b>	<b>Description</b>
College academic achievement	Mastering academic content, as measured by the final grade in a single college-level course, the final exam score in a single college-level course, grade point average in college-level courses (if the formula is specified), or scores on professional exams such as the GRE. Note that technical skills proficiency is a separate domain.
Technical skills proficiency	Technical skills at the occupation level, measured by assessments aligned with industry-recognized standards
<b>Educational Attainment<sup>3</sup></b>	
Completing high school in the short term	Whether the student has earned a high school diploma or a recognized equivalent, such as obtaining passing scores on a state-approved high school equivalency test, by the time period closest to the end of the intervention
Completing high school in the long term	Whether the student has earned a high school diploma or a recognized equivalent, such as obtaining passing scores on a state-approved high school equivalency test, by the last time period of data collection at least 9 months following the end of the intervention
College degree attainment in the short term	Completion of a postsecondary degree, certificate, or program by the time period closest to the end of the intervention. Note that industry-recognized certificates are in a separate domain.
College degree attainment in the long term	Completion of a postsecondary degree, certificate, or program by the last time period of data collection at least 9 months following the end of the intervention. Note that industry-recognized certificates are in a separate domain.
Industry-recognized credential, certificate, or license in the short term	Completion of an industry-recognized credential, certificate, or license by the time period closest to the end of the intervention
Industry-recognized credential, certificate, or license in the long term	Completion of an industry-recognized credential, certificate, or license by the last time period of data collection at least 9 months following the end of the intervention
<b>School Attendance and Progress</b>	
School attendance	Attendance rates, absenteeism, or tardiness at school, such as the number or proportion of days in attendance, the proportion of students with excessive absences, or referrals for truancy
Staying in secondary school <sup>4</sup>	Whether the student has dropped out of school or is still enrolled

<sup>3</sup> Separate domains are distinguished for short-term educational attainment, measured closest to the end of the receipt of the intervention but not afterwards, and long-term attainment, measured at the longest follow-up after the end of the intervention. Educational attainment outcomes at all other time points are considered supplementary findings.

<b>Domain</b>	<b>Description</b>
Progressing in secondary school <sup>4</sup> or adult education	Number of Carnegie units or credits the student has earned, whether the student was promoted to the next grade, or highest grade completed
College readiness	Student progress on preparedness to enter postsecondary education, such as meeting specific minimal coursework requirements for entry into postsecondary institutions <sup>4</sup> , completing other required key milestones such as applying for college or for student aid, and completing college requirements in high school
College enrollment	Applying to, enrolling, and attending a postsecondary institution <sup>5</sup> , such as actual enrollment in college, number or selectivity of admitted institutions, enrollment by institution type, full-time vs. part-time enrollment, and immediate vs. delayed enrollment
College attendance	Attendance rates or absenteeism at a postsecondary institution <sup>5</sup> , such as the number or proportion of days in attendance, or the proportion of students with excessive absences
Progressing in developmental education	Complete required development coursework, grades earned in developmental courses, or completed vs. did not complete the first college-level course in which remediation was needed
Progressing in college	Progress toward the completion of a postsecondary degree, certificate, or program, such as number of college-level credits earned, number of terms of continuous enrollment, enrolled vs. did not enroll the next semester, or completed a course that was the focus of the intervention. Note that non-college level credits, such as developmental credits, are not eligible under this domain.
<b>Social-Emotional Learning and Behavior</b>	
Independent living skills	Ability to independently perform daily activities in the home, community, or vocational/ occupational settings, such as cooking, dressing, eating, using the toilet, cleaning, organizing office supplies, using office equipment, crossing the street, using a bank, and purchasing items. Note that this domain includes financial literacy, health literacy, and computer literacy.
Self-determination	Abilities that help learners set goals and take actions to achieve goals, such as goal setting, problem solving, decision making, self-advocacy, and choice making
Student social functioning	Student behaviors and skills that primarily involve interactions with others, such as students or teachers, or reflect attempts at social interactions or underlying skills for social functioning, such as consideration of others. Examples of outcomes in this domain include peer rejection, isolation, victimization, impulsivity, adaptive functioning, self-control, and social awareness.

<sup>4</sup> "Postsecondary institution" includes universities, colleges, community colleges, and trade schools.

<b>Domain</b>	<b>Description</b>
Compliant student behavior	Includes observed or recordable student behaviors that primarily reflect individual choices to follow behavioral rules or expectations, and have individual consequences for the student. Examples of outcomes in this domain include disruptive behavior, delinquent behaviors, physical aggression, bullying, harassment, arrests, substance abuse, and school suspension or expulsion.
Student emotional status	Includes student behaviors and self-ratings that are primarily focused inward and reflect a student's emotional state, sometimes referred to as emotional or internalizing behaviors; measures may be based on a diagnosis or classification, student self-report, teacher observation, or results from an assessment scale. Examples of outcomes in this domain include self-awareness, thought disorders, emotional regulation, depression, anxiety, and overall adjustment.
Student engagement in school	Behaviors that are typically only observed during school and often reflect school connectedness, such as coming to class prepared, paying attention in class, staying on task during class, completing assignments, and participating in classroom or other school activities.
<b>School Environment Outcomes</b>	
School climate	Observations or assessments of the school environment, the quality of interpersonal relationships within the school, and other factors that describe the character of a school
School equity	Measures of the degree of equity or assessments of equity within classrooms or schools, such as disproportionality assessments or student or staff surveys of their perceptions of equity
<b>Labor Market Outcomes<sup>5</sup></b>	
Short-term employment	Indicator, at the time period closest to the end of the intervention, of any paid employment, number of months or quarters employed, or number of hours worked in an average week. Note that these outcomes must be defined over a specific time period.
Long-term employment	Indicator, at the last time period of data collection at least 9 months following the end of the intervention, of any paid employment, number of months or quarters employed, or number of hours worked in an average week. Note that these outcomes must be defined over a specific time period.
Short-term earnings	Income received from work over a specific period closest to the end of the intervention. Note that earnings must be defined for those not employed as well those employed.

<sup>5</sup> Separate domains are distinguished for short-term labor market outcomes, measured closest to the end of the receipt of the intervention but not afterwards, and long-term labor market outcomes, measured at the longest follow-up after the end of the intervention. Labor market outcomes at all other time points are considered supplementary findings.

<b>Domain</b>	<b>Description</b>
Long-term earnings	Income received from work at the last time period of data collection at least 9 months following the end of the intervention. Note that earnings must be defined for those not employed as well those employed.
<b>Teacher Outcomes</b>	
Educator discipline practice	Attempts to influence problematic student behavior by responding to students' actions with consequences or rewards, such as office discipline referrals or discipline practices measured by teacher self-reports, classroom observations, or student surveys
Instructional practice	Quality of instruction provided by teachers and their application of content knowledge or pedagogical content knowledge as demonstrated by their actions in the classroom, based on rubrics assessed by school principals, supervisors, or trained evaluators, or based on surveys administered to students
Teacher attendance	Includes outcomes that indicate the number or percentage of eligible work days for which the teacher is present
Teacher retention at the school	Includes outcomes that measure the percentage of teachers who return to work as a teacher in the same school from year to year
Teacher retention in the school district	Includes outcomes that measure the percentage of teachers who return to work as a teacher in the same school district from year to year
Teacher retention in the state	Includes outcomes that measure the percentage of teachers who return to work as a teacher in the same state from year to year, or who return teach in certain kinds of settings, such as in special education or in economically disadvantaged districts
Teacher retention in the profession	Includes outcomes that measure the percentage of teachers who return to work as a teacher from year to year, regardless of location
<b>School Leader Outcomes</b>	
Leadership practice	Includes outcomes that measure the quality of leadership ability demonstrated by the school leader, as measured from rubrics assessed by supervisor or from surveys of school staff, parents, or students
School leader retention at the school	Includes outcomes that measure the percentage of school leaders who return to work as a school leader in the same school from year to year
School leader retention in the school district	Includes outcomes that measure the percentage of school leaders who return to work as a school leader in the same school district from year to year

<b>Domain</b>	<b>Description</b>
School leader retention in the state	Includes outcomes that measure the percentage of school leaders who return to work as a school leader in the same state from year to year, or who return to lead schools in certain kinds of settings, such as in economically disadvantaged districts
School leader retention in the profession	Includes outcomes that measure the percentage of school leaders who return to work as a school leader from year to year, regardless of location

The *WWC Procedures and Standards Handbook (version 3.0)* discusses the types of outcomes, criteria the outcome must meet, and how outcomes are reported. See Section III.B.4: Outcome Eligibility and Reporting.

### Findings to Review

**Exhibit 2—Distinguishing Main Findings and Supplementary Findings**

	<b>Main Findings</b>	<b>Supplementary Findings</b>
<b>Research question</b>	Confirmatory	Exploratory
<b>Sample</b>	Full sample	Subgroups
<b>Measure</b>	Composite	Subtests, Subscales
<b>Time period</b>	Measurement closest to the end of intervention <sup>6</sup> , except educational attainment and labor market outcomes are also reported in separate domains at the longest follow-up	Other time points
<b>Analytic methods</b>	Benchmark analyses	Sensitivity analyses

A study’s rating will be determined by the highest rating achieved across all findings eligible for review. Using Exhibits 1 and 2, review team leadership may apply restrictions on eligible *supplementary* findings to narrow the focus of reviews for a given effort. For example, when conducting reviews for a grant competition, review team leadership can indicate that no supplementary findings should be examined, or only a subset of them.

### DESIGN STANDARDS

Eligible studies are assessed, as applicable, against WWC group design standards, pilot regression discontinuity design standards, or pilot single-case design standards, as described in the *WWC Procedures and Standards Handbook (version 3.0)*; see Section III.B, Appendix D, and Appendix E. In addition, the *WWC Procedures and Standards Handbook (version 3.0)* requires additional input on sample attrition, baseline equivalence, and statistical adjustments for baseline equivalence.

---

<sup>6</sup> If the intervention had not ended when study data were collected, the time period closest to the end of the intervention is defined as the most recent time period reported in the study.

### **Sample Attrition**

The default attrition boundary for a study reviewed under this protocol is the *liberal* boundary. This boundary was selected because it is the most common boundary used across WWC review protocols, suggesting that in typical education studies, attrition is not thought to be strongly related to intervention status. The lead methodologist, in consultation with a content expert, may choose the conservative boundary if it is justified for a particular review, as in the case of a study with suspected endogenous attrition. If this decision is made, the rationale for it will be documented in the study review guide.

### **Baseline Equivalence**

For student outcomes, baseline equivalence of the intervention and comparison groups in the analytic sample should be demonstrated on a direct pretest in the same domain as the outcome. If a pretest in the same domain as the outcome was not given, or does not exist for certain outcomes, then the baseline equivalence of the intervention and comparison groups in the analytic sample should be demonstrated on at least two of the following:

- (a) a standardized measure of academic readiness, knowledge, or skills;
- (b) student GPA (grades 6-12 or at the postsecondary or adult education level only); and
- (c) socio-economic status; for example, student free- and reduced-price lunch status, family income, or parent's education.

For teacher outcomes, baseline equivalence should be established based on Exhibit 4 from the *Teacher Excellence* protocol, version 4.0. For school leader outcomes, baseline equivalence should be established based on Exhibit 5 from the *School Leadership* protocol, version 4.0.

### **Statistical Adjustment for Baseline Equivalence**

If a pretest is available for an outcome and the difference between conditions is shown to be within the range that requires statistical adjustment, the adjustment is needed only for that outcome. For outcomes that do not have a pretest, an adjustment is necessary only for required covariates, such as measures of academic achievement and socio-economic status, that are shown to be within the range that requires statistical adjustment.

### **WWC ADJUSTMENTS**

The *WWC Procedures and Standards Handbook (version 3.0)* discusses the types of adjustments made by the WWC in Section IV: Reporting on Findings (pp. 22-26). For characterizing the statistical significance of findings from a study that meets WWC standards, the WWC will conduct the multiple comparison correction separately for main findings within the same outcome domain, but not across domains or for any supplementary findings.