REVIEW PROTOCOL FOR STUDIES OF INTERVENTIONS TO SUPPORT POSTSECONDARY SUCCESS VERSION 4.0

This protocol guides the review of research that informs What Works Clearinghouse (WWC) intervention reports in the Supporting Postsecondary Success topic area. The protocol is used in conjunction with the *WWC Procedures and Standards Handbooks (version 4.0).*

PURPOSE STATEMENT

Enrolling in postsecondary institutions and completing a degree or certificate is one of the primary pathways for economic success and is increasingly required for employment in a variety of fields. Yet, only 25% of students enrolled at public two-year institutions and 62% of students at private nonprofit two-year institutions complete their programs within three years (McFarland et al., 2019). At four-year institutions, 60% of students who start at public institutions and 66% of students who start at private nonprofit institutions complete their programs at any institution within six years (McFarland et al., 2019). The first year of college is critical, with about 19% of first-time full-time students in 4-year institutions and 38% of first-time full-time students in 2-year institutions failing to return to the same institution for their second year (McFarland et al., 2019). Supporting postsecondary students to promote academic performance, attendance, retention, and ultimately degree completion is a focus of many postsecondary initiatives.

This protocol focuses broadly on academic and nonacademic interventions that aim to support postsecondary success, including first year experience courses, credit recovery programs, linked learning communities, residential learning communities, interventions focused on financial aid, and student counseling and mentoring interventions. Systematic reviews of evidence in this topic area address the following research questions:

- Which interventions are effective at helping students improve access and enrollment in college?
- Which interventions are effective at helping students increase credit accumulation and persistence in college?
- Which interventions are effective at helping students improve academic achievement?
- Which interventions are effective at helping students complete college?
- Which interventions are effective at helping students improve their prospects in the labor market?
- Is the reviewed intervention more or less effective for certain subgroups of students (including first-generation college students, women, racial/ethnic minorities, academically underprepared students, students from low socioeconomic status backgrounds (e.g., Pell Grant recipients), and/or community college students)?

KEY DEFINITIONS

The descriptions below capture some of the many types of academic and nonacademic interventions that are covered by this protocol. However, it is not a comprehensive list of the programs, policies, and interventions that are covered.

Academic Tutoring

Academic tutoring refers to instruction in reading, writing, study skills, mathematics, science, and other subjects, typically conducted one-on-one or in small groups of students with a tutor. Tutors may be peers, staff, faculty, or external contributors and are not required to be affiliated with a school. Tutoring programs may be face-to-face or online, and may be tailored to individual students or to the needs of the group. Tutoring practices include coaching and support for academic tasks such as completing homework, understanding lecture material, or review of written work.

Advising

Advising at postsecondary institutions is designed to help students overcome barriers to persistence and completion, including: being academically unprepared; having difficulty paying for college; lacking familiarity with the college procedures and processes; and dealing with competing priorities, such as employment or family obligations. Advising interventions may include academic support, help in making sound decisions about financing college, assistance with course selection, and counseling to help students balance their academic and nonacademic obligations.

Credit Recovery Programs

Credit recovery programs for postsecondary students are programs that permit students to take high school, developmental, or transfer courses they may have missed or failed previously. In postsecondary settings, credit recovery programs are often billed as a method for transfer or returning adult students to maximize their transfer credits upon enrollment. Credit recovery programs may be offered in brick-and-mortar settings, but are increasingly offered online. As such, credit recovery programs may employ a variety of instructional formats, including traditional instruction (e.g., lectures), one-on-one instruction or tutoring, or independent study.

First Year Experience Programs

First Year Experience (FYE) programs, often referred to as college success courses or freshman seminars, may be a required or elective course for first-year students in 2-year or 4-year colleges. The general goal of such programs is to promote students' transition to college, academic performance, social development, persistence, and degree completion (Hunter & Linder, 2005; Pascarella & Terenzini, 2005). While courses vary in terms of content and focus, most First Year Experience programs are designed to introduce students to "campus resources, time management, study skills, career planning, cultural diversity, and student development issues" (Barefoot & Fidler, 1992, p. 2). First Year Experience programs may be offered as extended orientation seminars, academic seminars, or professional or discipline-linked seminars. All such programs are eligible for review by the WWC. However, college or university orientations alone are not considered First Year Experiences, nor are programs or seminars that focus exclusively on teaching students study or test-taking skills.

Linked Learning Communities

Linked learning communities are programs defined by having social and curricular linkages that provide undergraduate students with intentional integration of the themes and concepts that they are learning (Inkelas & Soldner, 2011). Linked learning communities are based on the theory that active learning in a community-based setting can improve academic outcomes by increasing social as well as academic integration (Smith, MacGregor, Matthews, & Gabelnick, 2004). To that end, linked learning communities tend to incorporate two characteristics: a shared intellectual theme with a linked or integrated curriculum and a community or common cohort of learners (Lenning & Ebbers, 1999).

Mentoring Programs

Mentoring programs involve a pairing between a more experienced, skilled, or knowledgeable person (e.g., upper-level student, instructor, or community member) and a less-experienced student. Mentoring may take place in educational or other community settings. Typically, the mentor receives training or support to maintain a reciprocal, personalized relationship with the student, and provides information, support, and guidance to the mentee (Cannata, Garringer, MacRae, & Wakeland, 2005; Crisp & Cruz, 2009). Mentoring programs may focus on interpersonal skills, self-esteem, career maturity and development, psychological and emotional support, as well as academic planning, troubleshooting academic assignments, or tutoring in problem areas (Cannata et al., 2005; Crisp & Cruz, 2009; DuBois, Holloway, Valentine, & Cooper, 2002; DuBois et al., 2013). Typical mentoring programs are characterized by a 1:1 relationship between the mentee and mentor and occur over a period of time (i.e., more than one session). However, group formats in which one or more mentors meets with a small group of students may also be used. Mentoring programs generally involve mentors in a non-professional helping capacity. For example, professional staff who counsel students at a counseling center would not be considered mentors, though these same individuals could serve as mentors outside the counseling center in a non-professional capacity.

Supplemental Instruction

Supplemental Instruction (SI) is a structured academic support program or model, which is typically led by more experienced peer leaders, and directly connected to a particular course (Dawson, van der Meer, Skalicky, & Cowley, 2014). Peer leaders facilitate SI sessions by asking leading questions for students to construct their own knowledge about a topic from class or deepen their understanding about a topic from class. SI Leaders prepare learning activities such as problem-solving exercises, mock exams, group work, or worksheets, and serve as role models by attending lectures, taking notes, reading assigned materials, and demonstrating effective study skills (Dawson, van der Meer, Skalicky, & Cowley, 2014).

ELIGIBILITY CRITERIA

Identifying Studies for Review

The *WWC Procedures and Standards Handbook (version 4.0)* discusses general procedures for conducting a literature search. For the supporting postsecondary success topic area, a broad electronic database search was conducted to identify potentially relevant studies. Once interventions have been identified as being targets for an intervention report, the WWC will update the electronic database search and supplement this with targeted searches of government and non-government

agency websites, relevant non-profit organizations that might fund research on the intervention, and by reviewing the bibliographies of literature reviews, meta-analyses, and primary studies of the intervention under review. The review team will also search the WWC database of previously reviewed studies to identify studies that have met standards in prior reviews. Those studies will be re-reviewed using the eligibility criteria and evidence standards described in this protocol. The team will also identify studies that have been rated as ineligible in prior reviews and will confirm that they are ineligible for this review based on the criteria described in this protocol. A broad search strategy for the interventions for students in supporting postsecondary success topic area is detailed in Appendix A.

Studies must meet several criteria to be eligible for review. These relate to the population that was sampled, the intervention that was studied, the study design that was used, outcomes that were measured, and when the study was conducted. Each of these is discussed below.

Eligible Populations

To be eligible for review under this protocol, a study must include postsecondary students in the United States or Canada.

In general, the WWC determines a study rating based on average intervention effects and will report subgroup analyses only for groups that are identified in the protocol as being of theoretical, policy, or practical interest. For studies reviewed under this protocol, eligible subgroups include:

- first-generation college students
- racial/ethnic minorities
- gender
- students from low socioeconomic status backgrounds, such as Pell Grant recipients,
- academically unprepared students (e.g., incoming students with developmental course requirements versus those with no developmental requirements)
- community college students.

Eligible Interventions

For this review, interventions that support postsecondary success for college students must be primarily focused on improving student outcomes during and after college. Thus, programs may be focused on improving the academic performance of students while attending college, increasing the number of students who transfer from 2- to 4-year colleges or attain a degree or certificate, or improving post-college labor market outcomes. This area may include instructional programs that occur in-person or online as part of postsecondary schooling, out-of-classroom practices such as mentoring, classroom practices such as strengths-based learning, adult education activities, and any other strategy focused on supporting student success in college. Within this framework, a number of broad intervention strategies are relevant, as follows:

- **Interventions to support students during the critical first year.** Many students arrive in higher education academically prepared, but may not have the study skills or other coping skills to deal with the learning and social environments in higher education. In addition, many students simply make poor academic and social decisions during their first year of college. Students may have difficulties understanding the consequences of these decisions for their eventual success in college. Many interventions are designed to ensure that first-year students who are otherwise academically qualified to succeed do so. These include interventions designed to simplify the advising process; interventions designed to increase academic momentum; "study skills" courses or resources; first-year or freshman seminars; programs that orient students to college life and the campus community; and various student learning community structures (e.g., residential colleges).

- **Interventions to support the retention, academic performance, and completion rates of postsecondary students.** Because degree completion rates have consistently been low for postsecondary institutions (McFarland et al., 2018), many institutions have implemented retention initiatives to help improve retention and completion rates. These initiatives have taken many forms, such as providing academic support, study skills, counseling, mentoring, and other institutional support mechanisms (Luna & Zienkewicz, 2014). These interventions can be delivered individually or in group settings (i.e., individual counseling, strengths-based education, etc.), and can occur in-person or online as part of postsecondary schooling.

- **Interventions relating to financing college.** The cost of higher education is a barrier to postsecondary success for many students, most especially those from low income families (Deming & Dynarski, 2009). Rapidly rising costs across all types of postsecondary institutions in recent years, as well as larger numbers of students enrolling in postsecondary institutions than ever before have increased the number of students for whom financial pressures are an obstacle to attending college and completing a degree (McFarland et al., 2018). A variety of strategies have been developed to reduce the price of college, with variants on the form of financial aid (e.g., need or merit), loan amounts, and different combinations and "packaging" of aid.

See the section "Key Definitions," above, for the operational definitions for interventions that are the subject (or potentially are the subject) of WWC reviews.

Only interventions that are replicable are eligible for review. The following characteristics of an intervention will be documented by the WWC, so that practitioners other than developers can reliably reproduce the intervention with different participants, in different settings, and at other times:

- Targeted population;
- Description of intervention provider or administrator, including their qualifications;
- Description of the intervention, including details of the services provided, unit of delivery (e.g., whole class, individual), medium/media of delivery (e.g., teacher-led instruction or software), and other activities that are part of the intervention;
- Length of calendar time and number of hours required to implement the intervention;

- Cost, which may include staff salaries to participate in training or provide the intervention; expenses for space, materials, and equipment needed for training and/or providing the intervention; travel and per diem expenses for training; price charged for intervention participants; and other intervention inputs; and
- Source of funding (when available).

Eligible Research

In order to be eligible for review a study must be a primary analysis of the effects of an intervention. If a study does not examine the effects of an intervention, or if it is not a primary analysis (e.g., if it is a meta-analysis or other literature review), then it is not eligible for review.

- ***Topic.*** The study must be focused the effects of an intervention, delivered in a postsecondary setting, on one or more measures of enrollment, credit accumulation, persistence, academic achievement, educational attainment, or labor market outcomes. Postsecondary education is any form of schooling occurring after the secondary level (i.e., after high school) and may include public or private technical colleges, community colleges, four-year institutions, and any other institution offering a postsecondary certificate or degree. Interventions may start as early as the summer immediately following high school and run through a student's entire postsecondary career.

- ***Time frame.*** The study must have been published within 20 years of the year of the review (for example, 1999 or later for reviews occurring in 2019). Rigorous evaluations of interventions implemented in this time frame test versions of interventions most likely to be available today and under conditions most likely to be current. For updated intervention reports, the study must have been released since the original intervention report's literature search start date (i.e., 1994 for reports released under version 3.0 of supporting postsecondary success protocol). Studies must be publicly available (accessible online or available through a publication, such as a journal) at the time of the original or updated literature search.

- ***Sample.*** The study sample must meet the requirements described in the "Eligible Populations" section above. Studies with samples that are comprised primarily of high school students, even those concurrently enrolled in postsecondary education (i.e., dual enrollment), are not eligible for review under this protocol; rather, they should be reviewed under the Transition to College protocol. Studies with samples that are comprised predominantly of postsecondary students in need of developmental coursework should be reviewed under the Review Protocol for Studies of Interventions for Developmental Students in Postsecondary Education.

- ***Design.*** The study must be empirical, using quantitative methods and inferential statistical analysis, and as described by the *WWC Procedures and Standards Handbook* (version 4.0), must take the form of a randomized controlled trial (RCT) or use a quasi-experimental design (QED), a regression discontinuity design (RDD), or a single-case design (SCD).

- ***Language.*** The study must be available in English.

- ***Location.*** The study must include students in the United States, in its territories or tribal entities, at U.S. military bases overseas, or in Canada.

Review team leadership should be notified when studies present counterfactuals other than business-as-usual (BAU), such as studies that compare two interventions to one another. These studies will be reviewed by review team leadership to determine whether their results can be reasonably combined with other studies without biasing WWC calculations. Review team leadership will advise reviewers on characteristics of the comparison condition to document so the counterfactual can be clearly documented in WWC reports.

Eligible Outcomes

To be eligible for review, a study must also report outcomes from a relevant postsecondary outcome domain. These include: (a) college enrollment, (b) college attendance, (c) progressing in college, (d) academic achievement, (e) postsecondary degree attainment, (f) credential attainment, (g) employment, and (h) earnings. Operational definitions for each outcome domain are provided below. When measures from an official and an unofficial source are available (e.g., grades reported by the institution vs. self-report) the WWC will focus on the official source.

- **College enrollment** refers to the process of applying to, actually enrolling, and attending a postsecondary institution. Examples of ways that enrollment might be operationally defined in studies include: (a) actual enrollment in college, (b) number and/or selectivity of admitted and/or enrolling institutions, (c) enrollment by institution type (2 year vs. 4 year), (d) intensity of enrollment (full time vs. part time), and (e) timing of enrollment (e.g., immediate vs. delayed enrollment after high school).
- **College attendance** refers to outcomes that measure attendance rates or absenteeism at school. Ways that attendance might be operationalized include (a) the number or proportion of days absent or in attendance during a school term and (b) proportion of students with excessive absences. Objective measures of attendance, such as those from school administrative records are preferred, but student reported measures are acceptable if a more objective measure is not available.
- **Progressing in college** refers to progress toward the completion of a degree, certificate, or program. Examples of ways that credit accumulation might be operationally defined in studies include: (a) number of college-level credits earned, (b) number of terms of continuous enrollment, (c) enrolled vs. did not enroll the next semester, and (d) completion of a single course that was the focus of the intervention. Completion of a single course will only be reported if other measures within this domain are not reported by the study. The number of non-college level credits earned (e.g., developmental credits) is not an eligible measure of credit accumulation.
- **Academic achievement** refers to the extent to which students master academic content. As such, eligible measures of academic achievement are those that arise naturally from student educational experiences. Examples of ways that academic achievement might be operationally defined in studies include (a) final grade in a single college-level course, (b) grade point average in college-level courses, and (c) the ratio of college-level courses passed vs. failed. Scores on department-wide exams, standardized tests, and professional or industry exams (e.g., the GRE and the NCLEX-RN) are also eligible. With the exception of

department-wide exams, measures that exist below the final course grade level are not eligible (e.g., average test score, score on a particular assignment or project).[1] Also ineligible are measures of academic achievement that do not directly contribute to student grades (e.g., a math test that is given after an experimental manipulation, the performance on which has no implications for a student's performance in a specific course).

- **Postsecondary degree attainment.** This domain refers to the completion of an associate or a baccalaureate degree. Outcomes pertaining to completing or progressing toward a graduate-level degree will not be included.
- **Credential attainment**. This domain refers to the completion of an industry-recognized credential, certificate, or license. Examples of ways completion might be operationally defined in a study include (a) certificate completion rates, (b) non-degree-award receipt rates, and (c) certifications from third-party licensing or credentialing bodies.
- **Employment** refers to outcomes related to employment after the postsecondary experience. Examples of ways that employment outcomes might be operationally defined in studies include (a) employed vs. not, (b) employed full-time vs. employed part-time, and (c) employed in field of study vs. not, as defined by the major groups of the Standard Occupational Classification (SOC) System.[2]
- **Earnings** refers to income from employment after the postsecondary experience. Examples of ways that earnings might be operationally defined in studies include (a) cumulative earnings over the previous six months or (b) earnings in a typical week in the previous month.[3] Earnings received during a student's course of study are ineligible for review.

Outcomes Measured at Different Points in Time

When outcomes are measured at multiple time points, the follow-up outcome measured closest to the end of the intervention on the full sample will be prioritized as the primary finding. This will allow for more clear attribution of the intervention to the outcome observed (especially in QEDs), relative to prioritizing the longest follow-up observation. Notable exceptions include:

- In the college enrollment domain, the first measure of enrollment (e.g., enrolled vs. not enrolled) is selected as the primary finding. Measures of continued enrollment that occur after the first semester or year of college would fall under the progressing in college domain and the follow-up outcome measured closest to the end of the intervention is selected as the primary measure.

[1] The "department-wide" exam criterion ensures that students are tested consistently across treatment and comparison groups, and those exam scores contribute meaningfully to overall academic achievement.

[2] https://www.bls.gov/soc/2018/major_groups.htm

[3] Total individual or household income is not an eligible outcome. An intervention that successfully increases individual earnings might decrease public benefit receipt, with the result that the participant's income might increase, decrease, or remain constant even though the intervention successfully increased earnings. Further, household income might include spousal earnings; an intervention that increases a participant's earnings might induce the spouse of the participant to reduce his/her hours worked, especially if the spouse was working an additional job to support the participant during training.

- In the attainment domain, the longest follow-up time point will be selected as the primary finding. In the employment and earnings domains, two primary findings will be presented when they are available: (1) the follow-up outcome period closest to the end of the intervention and (2) the longest follow-up time point.

EVIDENCE STANDARDS

Eligible studies are assessed against WWC evidence standards, as described in the *WWC Procedures Handbook*, Section IV: Screening Studies and Section V: Reviewing Studies, as well as the *WWC Standards Handbook*. Generally, these standards assess outcome reliability and validity, attrition, baseline equivalence, and similar methodological and statistical issues. This review determines the overall WWC study rating (see the *Procedures and Standards Handbook version 4.0* for further details). Details related to sample attrition in RCTs and baseline equivalence in QEDs and high-attrition RCTs are outlined below to highlight the way they are operationalized for this topic area.

Eligible Study Designs

Studies that use group designs (RCTs and QEDs), RDDs, or single-case designs (SCDs) are eligible for review using the appropriate standards.

Outcome Measure Requirements

In this review, the requirements for outcome measures are more stringent than those specified in the *WWC Standards Handbook* (in Section IV.A: Outcome Requirements and Reporting). Specifically, this review requires a minimum of 0.60 (as measured by, for example, Cronbach's alpha), for internal consistency to satisfy the reliability requirement for an outcome measure.

Sample Attrition

The *WWC Standards Handbook* discusses the sample attrition standards used by the WWC in the following sections:

- Section II.A—"Sample Attrition: Is the combination of overall and differential attrition high?" – in Step 2 of the WWC review process for individual-level group design studies.
- Section II.B—"Is the study a cluster RCT with low cluster-level attrition?" – in Step 1 of the WWC review process for cluster-level group design studies.
- Section II.B—"Is there a risk of bias due to non-response of individuals?" – in Step 3 of the WWC review process for cluster-level group design studies.
- Section II.D—"Calculating attrition when rating CACE estimates" – in Section 3 of the WWC standards for reviewing complier average causal effect estimates.
- Section III.C—in Standard 2 of the WWC standards for reviewing regression discontinuity designs.

In the *WWC Standards Handbook*, Figure II.2 illustrates the attrition boundary and Table II.1 reports attrition levels that define high and low attrition. Based on the choice of the boundary, the study

review guide calculates attrition and whether it is high or low. For most studies this review will entail use of the ***optimistic*** boundary for attrition based on the assumption that most attrition in studies of interventions focused on the promotion of postsecondary success would be due to factors that are not strongly related to intervention status. We assume that postsecondary students can have a range of life events that lead them to have missing outcome data that are unrelated to intervention status.

Joiners in Cluster Randomized Controlled Trials (RCTs)

According to the *WWC Standards Handbook* (page 23), to receive the highest rating a cluster RCT must limit the risk of bias due to individuals entering the cluster after the time of random assignment. This is because the presence of joiners in an analytic sample might introduce bias into estimates of an intervention's effectiveness. The WWC defines a joiner as any student who enters a cluster, such as a community college school or course, after the results of random assignment are known to any individual who could influence a student's placement into a cluster (for example, a college advisor). In some cases, joiners might enter clusters after random assignment, but before anyone outside of a study team could have known about cluster random assignment results. The WWC never considers these joiners to pose a risk of bias because the decisions that led these individuals to join clusters could not have been affected by the intervention. However, the burden for demonstrating that individuals could not have known about the intervention rests with the study authors.

In some cases, joiners who enter clusters relatively early in the study period have less potential to introduce bias than those who enter later. This is because late joiners might be more likely to do so because of the intervention. Therefore, the WWC differentiates between *early joiners* and *late joiners.* For this review protocol, we will consider college students to be *early joiners* if they enter a cluster within 6 weeks after the results of random assignment are publicly known. That is, the early period for joiners ends 6 weeks after the start of the school year if the results of random assignment were announced over the summer; otherwise, the early period ends 6 weeks after the results of random assignment were announced. *Late joiners* are those who enter clusters after 6 weeks.

With that background, the general default disposition in for this review is that ***all joiners in the analytic sample are expected to pose a risk of bias*** (there are exceptions for early joiners outlined below). Therefore, a study that includes at least one such joiner in the analytic sample does not limit the risk of bias from joiners. This is because study samples will be comprised of adults who can presumably choose their preferred education experiences, and might purposefully select or avoid a specific developmental education intervention.

An exception to the general default rule that all joiners in the analytic sample pose a risk of bias is when: (a) colleges/institutions, a group of colleges such as a coordinated group of community colleges, or blocks of courses within colleges represent the unit of assignment, and (b) the following conditions are in place:

- The intervention is not expected to directly affect joiners' enrollment or placement decisions. One example of an intervention that should not directly affect enrollment or placement decisions is when treatment and comparison groups are offered different types of potentially useful services, such as two competing mentoring interventions. In this case, we would not expect that individuals would be more likely to go out of their way to join one mentoring

intervention over the other unless there is time to closely investigate and consider the different options. In this scenario, ***only late joiners pose a risk of bias*.**

- Another example of an intervention that would likely not directly affect enrollment is when treatment group members receive a low-profile approach that is integrated into curricula of their community college, where individuals are unlikely to know about this add-on to the curricula even after the point of random assignment. This is consistent with aforementioned idea that joiners are likely to be unaware that a cluster is part of a study condition. In this scenario, ***only late joiners pose a risk of bias.***

Not all scenarios can be anticipated. When an intervention and unit of assignment in a cluster RCT do not fall into a category described above, the Review Team Leadership has discretion to make a decision about whether the joiners pose a risk of bias. Any time such discretion is exercised, the background and rationale of decisions will be documented in intervention reports

Baseline Equivalence

If the study design is an RCT or RDD with high levels of attrition or a QED, the study must satisfy the baseline equivalence requirement for the analytic intervention and comparison groups. The *WWC Standards Handbook* discusses how authors must satisfy the baseline equivalence requirement in:

- Section II.A—"Baseline Equivalence: Is equivalence established at baseline for the groups in the analytic sample?" – in Step 3 of the WWC review process for individual-level group design studies.
- Section II.B—"Does the study establish equivalence of individuals at baseline for groups in the analytic sample?" and "Does the study establish equivalence of clusters at baseline for groups in the analytic sample?" – in Steps 4 and 7 of the WWC review process for cluster-level group design studies, respectively.
- Section II.D—"Procedures for rating CACE estimates when attrition is high" – in Section 5 of the *WWC Standards* for reviewing complier average causal effect (CACE) estimates.
- Section III.C—in Standard 3 of the *WWC Standards* for reviewing RDDs.

This review assesses baseline equivalence within each domain and analytic sample. Elaborations follow:

- The outcome domains for this review cover multiple constructs, so an *outcome-by-outcome* approach to establishing equivalence is followed. The implication is that it is possible for a baseline difference to exceed 0.25 standard deviations on a given outcome, but this need not influence other outcomes within the domain. So, for example, a large baseline difference in mathematics will render all mathematics outcomes as not meeting WWC standards, but it would still be possible for a reading contrast from the same study to meet standards with reservations in the academic achievement domain. Furthermore, when the baseline difference for a pre-intervention measure is in the statistical adjustment range (that is, it is between 0.05 and 0.25 standard deviations), the adjustment must be made only in the analysis of the associated outcome measure. For example, if A, B, and C are available as pre- and post-intervention measures all within one domain, and the pre-intervention difference in B requires statistical adjustment, only the analysis of outcome B only must adjust for B.

- In cases where multiple baseline measures of SES and/or academic achievement are available, the Review Team Leadership is responsible for selecting the variable(s) to be used in the baseline equivalence assessment prior to the equivalence assessment being performed. For example, if both math and verbal scores on a college entrance exam are available, and the primary outcome is whether or not students passed their first college level math course, then the Review Team Leadership may decide that the score on the math portion of the entrance exam is the only achievement measure on which baseline equivalence will be assessed. However, if the primary outcome is attainment (and there is no natural pretest), then the Review Team Leadership might decide to assess balance on both the math subtest and the verbal subtest.

1. Baseline equivalence of individuals

For studies that must satisfy baseline equivalence of individuals, including cluster-level assignment studies being reviewed for evidence of effects on individuals, the baseline equivalence requirement must be satisfied for the analytic intervention and comparison groups. Pre-intervention measures of the outcome used in the analysis will be acceptable. However, in some cases it would be unusual to observe a meaningful baseline measure of some postsecondary outcomes, such as degree attainment. Within the supporting postsecondary success topic area, reasonable pre-intervention characteristics might encapsulate pre-college traits and events, such as high school GPA, college admission tests and work experience. These characteristics can be used as proxy pre-intervention variables to assess baseline equivalence within the domain to which they logically belong. For example, high school GPA can be used to establish baseline equivalence for outcomes within the academic achievement domain, and prior work experience can serve as a proxy baseline variable within the labor market domain.

With that background, reviewers should consider two options for baseline equating:

1. The first approach reviewers should take is to assess equivalence using a pre-intervention (baseline) measure of the outcome used in the analysis. If a pre-intervention measure of the outcome used in the analysis is not available, then baseline equivalence must be established on a pre-intervention measure of a proxy variable from within the same domain as the outcome used in the analysis. For example, high school GPA or SAT/ACT scores might be used to establish baseline equivalence for an academic achievement outcome measured in the freshman year.

2. If neither a pre-intervention measure of the outcome nor a proxy measure from the same domain are available, then baseline equivalence must be established on both of the following:
    - A continuously scaled pre-intervention measure of academic achievement. For example, high school GPA or SAT/ACT scores might be used to establish baseline equivalence.
    - A pre-intervention measure of student socio-economic status (e.g., FAFSA expected family contribution, family income, high school free- or reduced-price lunch status, parent education levels, Pell grant eligibility) is acceptable for establishing baseline equivalence.

2. Baseline equivalence of clusters

Assessing equivalence of clusters

In general, considerations for satisfying baseline equivalence of individuals also apply to satisfying baseline equivalence of clusters. In particular, baseline equivalence of clusters in the intervention and comparison groups must be satisfied by using the same baseline measures listed above for assessing baseline equivalence of individuals, and the same statistical adjustment requirements apply.

Acceptable samples for demonstrating baseline equivalence of clusters

For this review, any of the following three samples can be used to satisfy the baseline equivalence requirement for the analytic sample of clusters (provided the data are representative of the individuals who were in the clusters at the time the baseline data were collected).

(a) The analytic sample of the same individuals from any pre-intervention time period.

(b) Individuals from the same cohort as the individuals in the analytic sample, within the same clusters. The baseline data may be obtained at the time that clusters were assigned to conditions or during the year prior to when clusters were assigned to conditions.

(c) Individuals from the previous academic year cohort, in the same grade, and within the same clusters, as individuals in the analytic sample.

If authors provide baseline information at multiple time periods, a reviewer should assess baseline equivalence using the information collected at the latest period prior to the start of the intervention. If authors provide baseline information for multiple samples, a reviewer should assess baseline equivalence using the sample listed first in the list above—that is, (a) should be used if available, then (b), and then (c). If authors provide baseline information for multiple samples across multiple time periods, the reviewer should consult review team leadership to determine which information to prioritize.

When a study examines the effectiveness of an intervention in multiple time periods, the sample used to satisfy baseline equivalence of clusters in the base period (for example, the school year after random assignment) also satisfies baseline equivalence of clusters in the later time periods (for example, 2 years after random assignment), so long as the outcome data are representative of the individuals in the clusters.

Statistical Adjustments

The *WWC Procedures Handbook* discusses the types of adjustments made by the WWC in Section VI: Reporting on Findings. For "mismatched" analysis (that is, when a study assigns units at the cluster level but conducts analysis at the individual level), this topic area uses the WWC default intra-class correlation coefficients of 0.20 for all student achievement outcomes and 0.10 for all behavior outcomes, unless a study-reported intra-class correlation coefficient is available.

PROCEDURES FOR CONDUCTING THE LITERATURE SEARCH

The *WWC Procedures Handbook, version 4.0,* discusses the procedures for conducting a literature search in Section III: Identifying Relevant Literature and Appendix B: Policies for Searching Studies

for Review. For the supporting postsecondary success topic area, a broad search will be conducted to identify potentially relevant intervention studies, using the search terms identified in Exhibit 1. Content experts will also be asked to identify and recommend interventions with a large body of causal evidence likely to be of interest to decision makers.

The review team will also search the WWC database of previously reviewed studies to identify studies that have met standards in prior reviews. Those studies will be re-reviewed using the eligibility criteria and evidence standards described in this protocol. The team will also identify studies that have been rated as ineligible in prior reviews and will confirm that they are ineligible for this review based on the criteria described in this protocol.

After identifying relevant interventions, the next step is to conduct intervention-specific literature searches, using the intervention name, to identify all publications on each intervention. This review may include additional search terms. The WWC will also supplement the intervention-specific electronic database search with targeted searches of government and non-government agency websites, relevant non-profit organizations that might fund research on postsecondary success interventions, and via reviewing the bibliographies of literature reviews, meta-analyses, and primary studies of the intervention under review.

In the final step, each citation gathered through this search process will undergo a screening process to determine whether the study meets the eligibility criteria established in the review protocol. This screening process is described in Chapter IV of the *WWC Procedures Handbook*. Finally, the interventions are prioritized for review based on the quantity and quality of eligible studies of the intervention. This prioritization process is described in Appendix A of the *WWC Procedures Handbook*.

Exhibit 1: Search Terms Used for the Initial Electronic Search

“control group*” or random OR "comparison group*" OR "regression discontinuity" OR "matched group*" OR baseline OR treatment OR experiment OR intervention OR evaluation OR impact OR effectiveness OR causal OR posttest or post-test OR pretest or pre-test OR QED OR RCT OR "propensity score matching" or randomized or quasi-experiment*	AND	“financial aid” or “college admission*” or “college prep*” or “College plan*” or “college choice” or “college readiness” or “college counsel*” or “college completion” or “Federal student aid” or “college access” or “transition* from high school” or “Transition* to college” or “access to college*” or “educational advancement” or “ready for college” or “readiness for college” or “college ready” or “FAFSA” or “Pathway* to college” or “barrier* to college” or “postsecondary transition*” or “financing college” or “college knowledge” or “college pathway*” or “college pipeline” or “step* to college

Additional Sources

Literature reviews for this topic area involve searching the websites and electronic databases listed in Appendix B of the *WWC Procedures Handbook* as well as the following websites:

- Abt Associates
- American Educational Research Association (AERA)
- American Evaluation Association (AEA)
- American Federation of Teachers (AFT)
- Association for Public Policy Analysis and Management (APPAM)
- Center for Data-Driven Reform in Education (CDDRE) at Johns Hopkins University
- Center for Research in Educational Policy (CREP)
- Center for the Study of Higher Education at Berkeley (CSHE)
- Center on Education Policy
- Community College Research Center (CCRC)
- Consortium for Policy Research in Education (CPRE)
- Cornell Higher Education Research Institute working papers
- Iowa Reading Research Center
- Mathematica
- MDRC
- National Bureau of Economic Research (NBER)
- National Center for Postsecondary Research
- National Center for Postsecondary Improvement
- National Education Association (NEA)
- Pacific Resources for Education and Learning (PREL)
- Public Policy Research Institute at Texas A&M University
- RAND
- Society for Research on Educational Effectiveness (SREE)
- Stanford Center for Education Policy Analysis (CEPA)
- WISCAPE working papers

REFERENCES

Altman, D. G., & Bland, J. M. (2003). Interaction revisited: The difference between two estimates. *British Medical Journal, 326*(7382), 219.

American College Testing (2014). *College student retention and graduation rates from 2000 through 2013.* Iowa City, IA: ACT. Retrieved from http://www.act.org/research/policymakers/reports/graduation.html

Barefoot, B. O., & Fidler, P. P. (1992). *Helping students climb the ladder: 1991 national survey of freshman seminar programs.* (Monograph No. 10). Columbia, SC: National Resource Center for The Freshman Year Experience, University of South Carolina.

Cannata, A., Garringer, M., MacRae, P., & Wakeland, D. (2005). *Making the grade: A guide to incorporating academic achievement into mentoring programs and relationships*. Folsom, CA: Mentoring Resource Center. Retrieved from http://educationnorthwest.org/sites/default/files/making-the-grade.pdf

Crisp, G., & Cruz, I. (2009). Mentoring college students: A critical review of the literature between 1990 and 2007. *Research in Higher Education, 50*(6), 525–545.

Dawson, P., van der Meer, J., Skalicky, J., and Cowley, K. (2014). On the effectiveness of supplemental instruction: A systematic review of supplemental instruction and peer-assisted study sessions literature between 2001 and 2010. *Review of Educational Research, 84*(4), 609–639.

Deming, D., & Dynarski, S. (2009). *Into college, out of poverty? Policies to increase the postsecondary attainment of the poor.* National Bureau of Economic Research Working Paper 15387. Retrieved from: http://www.people.fas.harvard.edu/~deming/papers/Deming_Dynarski_RobinHood.pdf.

DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: A meta-analytic review. *American Journal of Community Psychology, 30*(2), 157–197.

DuBois, D. L., Portillo, N., Rhodes, J. E., Silverthorn, N., & Valentine, J. C. (2011). How effective are mentoring programs for youth? A systematic assessment of the evidence. *Psychological Science in the Public Interest, 12*(2), 57–91.

Hunter, M. A., & Linder, C. W. (2005). First-year seminars. In M. L. Upcraft, J. N. Gardner, B. O. Barefoot, & Associates, *Challenging and supporting the first-year student: A handbook for improving the first year of college* (pp. 275–291). San Francisco: Jossey-Bass.

Inkelas, K. K., & Soldner, M. (2011). Undergraduate living-learning programs and student outcomes. In J. Smart & M. Paulsen (Eds.), Handbook of theory and research (pp. 335–368). New York: Springer.

Lenning, O. T., & Ebbers, L. H. (1999). The powerful potential of learning communities: Improving education for the future. ASHE–ERIC Higher Education Report, 26(6). Washington, DC: Graduate School of Education and Human Development, George Washington University. http://eric.ed.gov/?id=ED428606

Luna, G., & Zienkewicz, L. (2014). Increasing postsecondary retention and graduation through strengths-based education. *GSTF International Journal on Education, 2*(1), 81.

McFarland, J., Hussar, B., Zhang, J., Wang, X., Wang, K., Hein, S., Diliberti, M., Forrest Cataldi, E., Bullock Mann, F., and Barmer, A. (2019). *The Condition of Education 2019* (NCES 2019-144). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved May 31, 2019 from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2019144.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students. Vol. 2: A third decade of research.* San Francisco: Jossey-Bass.

Ritter, G. W., Barnett, J. H., Denny, G. S., & Albin, G. R. (2009). The effectiveness of volunteer tutoring programs for elementary and middle school students: A meta-analysis. *Review of Educational Research, 79*(1), 3–38.

Smith, B. L., MacGregor, J., Matthews, R. S., & Gabelnick, F. (2004). Learning communities: Reforming undergraduate education. San Francisco, CA: Wiley.

Tinto, V. (1987). *Leaving college: rethinking the cause and cures of student attrition.* Chicago: University of Chicago Press.

Tobolowsky, B. F., Mamrick, M., & Cox, B. E. (2005). *The 2003 national survey of first-year seminars: Continuing innovations in the collegiate curriculum* (Monograph no. 41). Columbia, SC: National Resource Center for the First Year Experience & Students in Transition, University South Carolina.