# PROTOCOL FOR REVIEWING INTERVENTIONS FOR CHILDREN IDENTIFIED WITH OR AT RISK FOR AN EMOTIONAL DISTURBANCE VERSION 3.0 (DECEMBER 31, 2015)

This review-specific protocol guides the review of research that informs the What Works Clearinghouse (WWC) intervention reports in the Children Identified With or at Risk for an Emotional Disturbance 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 interventions designed for use with K–12 students identified as having an emotional disturbance (ED) or at risk for classification of ED. This review will focus on the effect of interventions on the following types of outcomes: academic, problem behaviors, school engagement, self-determination, social-emotional competence, and transitional (i.e., the behavioral, social, functional skills that enable young adults with disabilities to obtain and hold meaningful employment, live independently, and obtain further training and education).

The following research questions guide this review:

- Which interventions are effective at improving outcomes of students identified with an ED, students at risk for classification of ED, students who have an emotional or behavior disorder consistent with ED, or students who are receiving an intervention to prevent classification? Are some interventions more effective than others for these students?

- Does the effectiveness of interventions for students with an ED vary by type of outcome?

- Are some interventions more effective for particular subgroups of students with an ED (for example, differences based on grade/age, English learner status, gender, race/ethnicity, socioeconomic status, school level) or in particular settings?

## KEY DEFINITIONS

***Emotional disturbance.*** This review adopts the definition used in the Individuals With Disabilities Education Act (IDEA), 20 U.S.C. § 1400 (2004) and accompanying federal regulations, 34 C.F.R. § 300.8 (2004). IDEA defines *emotional disturbance* as a "condition exhibiting one or more of the following characteristics over a long period of time and to a marked degree that adversely affects a student's educational performance:

(A) An inability to learn that cannot be explained by intellectual, sensory, or health factors.
(B) An inability to build or maintain satisfactory interpersonal relationships with peers and teachers.
(C) Inappropriate types of behavior or feelings under normal circumstances.
(D) A general pervasive mood of unhappiness or depression.
(E) A tendency to develop physical symptoms or fears associated with personal or school problems.

States will vary in how they define emotional disturbance. Note that the followings terms are also used to indicate an emotional disturbance: emotional or behavioral disorder, serious emotional disturbance, emotional handicap, behavioral disorder, or serious behavior disorder. For this review, these terms are considered synonymous with emotional disturbance.

# 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 | • 1-case design<br>• ABAB design<br>• Alternating treatment<br>• Baseline<br>• Causal<br>• Changing criterion<br>• Comparison group<br>• Control group<br>• Effectiveness<br>• Evaluation<br>• Experiment<br>• Impact<br>• Intrasubject replication<br>• Matched groups<br>• Meta-analysis<br>• Meta analysis<br>• Multi-element<br>• Multiple baseline<br>• Multiple probe design<br>• One-subject design | • Post-test<br>• Posttest<br>• Pre-test<br>• Pretest<br>• QED<br>• Quasi-experimental design<br>• Random*<br>• Randomized controlled trial<br>• RCT<br>• RDD<br>• Regression discontinuity<br>• Reversal design<br>• Simultaneous treatment<br>• Single case<br>• Single subject<br>• Treatment<br>• Withdrawal design |
| Intervention | • Approach*<br>• Curricul*<br>• Instruct*<br>• Intervention*<br>• Program* | • Strateg*<br>• Teach*<br>• Technique*<br>• Therap*<br>• Train* |
| Population | • Adolescen*<br>• Child*<br>• Student* | • Teen*<br>• Young adult*<br>• Youth |

| Disability | • Anorexia nervosa | • Emotionally handicapped |
| --- | --- | --- |
| | • Antisocial personality disorder* | • Externalized behavior* problem* |
| | • Anxiety disorder* | • Internalized behavior* problem* |
| | • Asperger syndrome* | |
| | • Bipolar disorder* | • Obsessive compulsive disorder* |
| | • Bulimia | |
| | • Conduct disorder* | • Oppositional defiant disorder* |
| | • Depress* | |
| | • Disruptive behavior disorder* | • Personality disorder* |
| | | • Post-traumatic stress disorder* |
| | • Dysthymia | |
| | • Eating disorder* | • Schizophrenia |
| | • EBD | • Selective mutism |
| | • Emotional disorder* | • Self-injurious behavior |
| | • Behavior* disorder* | • Somatic disorder |
| | • Emotional* disturb* | • Tourette's syndrome |

The asterisk (*) in a search term means that any word that begins with the specified letters is considered to be the search term (e.g., "student" and "students" are both search terms for student*).

**Additional Sources**

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

- Cambridge Center for Behavioral Studies
- Campbell Collaboration
- Council for Exceptional Children
- Kennedy Krieger Institute
- National Association of State Directors of Special Education
- National Center on Response to Intervention (RtI)
- National Dissemination Center for Children with Disabilities National Institute of Child Health and Human Development (NICHD)
- UC Davis M.I.N.D. Institute
- U.S. Department of Education

## ELIGIBILITY CRITERIA

### Eligible Populations

In this review, the following populations are of interest:

- *Grade range.* Students in grades K–12 (generally ages of 5–21) when the intervention is administered.

- *Location.* Students must reside in the United States, its territories, or tribal entities.

- *Disability classification.* Students have been identified as having an emotional disturbance under IDEA. Because some students who meet the classification criteria are not actually identified (for example, parents who are concerned about stigma do not want their children identified) and IDEA encourages prevention, the following sample populations are also eligible: (1) study reports the students are at risk of classification, (2) study reports the students have an emotional or behavior disorder consistent with an ED[1], or (3) study reports the intervention was administered to prevent students from being identified. Students cannot be identified with an emotional disturbance based solely on a parent reports.

- *Overlap with other WWC topic areas.* So that the study findings are relevant for the population of interest, at least 50% of the participants must be identified with an ED if a study uses a group design or a single-case design where the case is a group of students or the design is multiple probe/baseline across participants. If a study uses a reversal/withdrawal design, alternating treatment design, multiple baseline design across settings/conditions, or multiple probe design across settings/conditions, then at least one participant must be identified with an ED, and the review will focus on the identified participant(s).

Potential subgroups of interest for this review include:

- Characteristics of children and students:
  - Co-morbidity with other disabilities
  - Age
  - Gender
  - Race/ethnicity
  - English learner status
  - School level (elementary, middle, high)
  - Severity of disability
  - Socioeconomic status
  - Tracked for graduation

---

[1] The following are eligible conditions or disorders: antisocial personality disorder, anxiety disorder, bipolar disorder, conduct disorder, depression, disruptive behavior disorder, dysthymia, eating disorders, emotionally handicapped, externalized behavior problem, internalized behavior problem, obsessive-compulsive disorder, oppositional defiant disorder, personality disorder, post-traumatic stress disorder, schizophrenia, selective mutism, self-injurious behavior, Tourette's syndrome, and somatic disorder.

- Other characteristics:
    - Setting (such as general education classroom, special education classroom, school library)
    - Type of provider (teachers, paraprofessionals, peers, other)
    - Provider characteristics (staff credentials or training)
    - Other services received

**Eligible Interventions**

Interventions must have the acquisition of academic, cognitive, behavioral, school engagement, functional, self-determination, social-emotional competence, and transitional skills as their primary goals. Only interventions with a school-based component will be reviewed, including interventions partially delivered in a K–12 school site (as long as school-based staff members have a substantial role in implementation using resources at their disposal). This review will also include interventions with community and home-based components, if school staff members have responsibility for directly implementing an intervention component. These interventions will often involve a behavior plan, school-based counseling, and/or consultation.

Interventions administered outside of typical school settings (e.g., hospitals, residential treatment programs, or schools that only serve students with an ED or incarcerated youth) will not be reviewed. Simply collecting data in a school-based setting (for example, medication trials and therapy studies) suggest that the intervention does not have a substantial school-based component.

Only interventions that are replicable (i.e., can be reproduced) are eligible for review. The following characteristics of an intervention must be known to reliably reproduce the intervention with different participants, in other settings, and at other times:

- Intervention description: skills being targeted, approach to enhancing the skill(s) (e.g., strategies, activities, practices, and materials), unit of delivery of the intervention (for example, whole group, small group, individual), medium/media of delivery (for example, teacher-led instruction, paraprofessionals, peers, parents, software), and targeted population

- Intervention duration and intensity

- Description of individuals delivering or administering the intervention

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

- *Curricula.* A curriculum is a set of activities, materials, and/or guidance for working with students in classrooms. A curriculum has a clearly identified name, includes a description, and can be replicated by others based on written guidance, staff training, or technical assistance (for example, *The Me Book*). A curriculum may be (1) intended as the primary instructional tool designed to meet students' learning needs in multiple areas; or (2) designed to supplement the classroom material with

differentiated instruction or meet students' learning needs in specific areas. Both types of curricula will be included in this review.

- *Practices.* A practice is a named approach to promoting students' development that educators implement by interacting with students and materials in education settings (for example, functional behavior assessment). The review will include named practices that are clearly described, commonly understood, and used in published works by more than one investigator or team of investigators. Several terms may be used in the literature to refer to the same practice. A named practice may also refer to an array of specific procedures.

- *Policies.* A policy is a named condition under which programs operate. The policy must be commonly understood in the field and literature and directly affect services for students with an emotional disturbance (for example, inclusion). Policies may be set by federal, state, or local governments or by the organization providing services.

- *Programs.* A program is a service delivery model that may be associated with a funding stream and includes clear guidelines for implementation (for example, an early childhood special education class).

Both "branded" and "non-branded" interventions will be reviewed. Branded interventions are commercial or published programs and products that may possess any of the following characteristics:

- An external developer who provides technical assistance (e.g., instructions/guidance on the implementation of the intervention) or sells or distributes the intervention

- Trademark or copyright

**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 intervention must be focused on the acquisition of academic, cognitive, behavioral, communication/language, school engagement, functional, self-determination, social-emotional competence, and transitional skills.

- *Time frame.* The study must have been publicly released in 1989 or later (20 years prior to the start of the review in this area) and be obtained by the WWC for review prior to the drafting of the intervention report.

- *Sample.* The study sample must meet the requirements above in the "Eligible Populations" section. Follow-up findings are included when available and appropriate.

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

- *Location.* The study must include students in the United States, its territories, or tribal entities.

**Eligible Outcomes**

This review includes outcomes in the following domains:

- *Alphabetics.* Includes outcomes in the following areas: phonemic and phonological awareness, letter identification, print awareness, and phonics. Each is defined below.

  o **Letter identification** refers to knowledge of the names of the letters of the alphabet.

  o **Phonemic awareness (or phoneme awareness)** refers to the understanding that the sounds of spoken language—phonemes—work together to make words, and phonemes can be substituted and rearranged to create different words. Phonemic awareness includes the ability to identify, think about, and work with the individual sounds in spoken words. Phonemic awareness helps students learn how to read and spell by allowing them to combine or blend the separate sounds of a word to say the word (e.g., "/c/ /a/ /t/ – cat").

  o **Phonics** refers to the (a) knowledge that there is a predictable relationship between phonemes (the sounds in spoken language) and graphemes (the letters used to represent the sounds in written language), (b) ability to associate letters and letter combinations with sounds and blend them into syllables and words, and (c) understanding that this information can be used to decode or read words.

  o **Phonological awareness** is a more encompassing term than phoneme/phonemic awareness. It refers to phoneme awareness *and* to awareness of larger spoken units such as syllables and rhyming words. Tasks of phonological awareness might require students to generate words that rhyme, to segment sentences into words, to segment polysyllabic words into syllables, or to delete syllables from words (e.g., what is "cowboy" without "cow"?).

  o **Print awareness** refers to knowledge of concepts about print, such as (a) print carries a message; (b) print has conventions, such as directionality (left to right, top to bottom), differences between letters and words, distinctions between uppercase and lowercase, and punctuation; and (c) books have some common characteristics (e.g., author, title, front/back).

- *Comprehension.* Includes outcomes in the areas of vocabulary and comprehension development. Each is defined below.

  o **Reading comprehension** refers to the understanding of the meaning of a passage. Reading comprehension depends on various underlying components, including decoding (the ability to translate text into speech), knowledge of word meanings,

fluency (the ability to read text accurately and automatically), and the ability to understand and interpret spoken language. Struggling readers may have difficulty with any of these components of reading or with multiple components.

- o **Vocabulary development** refers to the development of knowledge about the meanings, uses, and pronunciation of words. The development of receptive vocabulary (words understood) and expressive vocabulary (words used) is critical for reading comprehension.

- *General reading achievement.* Includes outcomes that combine measures in two or more of the reading domains (alphabetics, reading fluency, and comprehension) or provide some other type of summary score across domains, such as a "total reading score" on a standardized reading test.

- *Math achievement.* Includes outcomes in the following areas: math facts, number sense, number and operations, fractions, measurement, data analysis, algebra, and geometry.

- *Problem behavior.* Includes actions that can function as a barrier to the social or academic development of a student or other students, such as behaviors caused by thought disorders or depression, aggression toward others (verbal, physical, and technological), disruption, lying/cheating/stealing, substance abuse, impulsivity, and lack of self-control or inhibition. (Outcomes focusing on time on task that incidentally measure lack of self-control, are classified as time on task.)

- *Reading fluency.* Includes outcomes that measure the ability to read text accurately, automatically, and with expression (including appropriate pausing, response to punctuation, etc.) while extracting meaning from it.

- *School engagement.* Includes outcomes in the following areas: educational attainment, involvement in individualized education plan process, staying in school, daily attendance, and time on task. (Outcomes focusing on problem behaviors that incidentally measure time off task, are classified as problem behavior.)

- *Science achievement.* Includes outcomes in the following areas: science facts, and the capacity to use the tools, procedures, inquiry, nature of science, argumentation in science, and reasoning processes of science. This includes subjects such as chemistry, earth science, biology.

- *Self-determination.* Includes outcomes that assess abilities that help youth set goals and take actions to achieve goals, such as goal setting, problem solving, decision making, self-advocacy, and choice making.

- *Social-emotional competence.* Social-emotional competence involves (a) self-awareness of thoughts, feelings, and behaviors; (b) social awareness of context and others; (c) social skills needed to interact and communicate with peers, teachers and other school staff; (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 domain includes social engagement (with adults or peers), self-management, adaptive functioning, and well-being.

- ***Social studies achievement.*** Includes, but is not limited to, areas included in the National Assessment of Educational Progress (NAEP), which measures knowledge of US history, world history, geography, economics, and civics.

- ***Writing achievement.*** Includes, but is not limited to, areas included in the NAEP, which measures narrative writing (the production of stories or personal essays), informative writing (communication of information), and persuasive writing (seeking to influence the reader to take action or bring about change). Spelling outcomes are considered part of writing achievement.

State alternative assessments are valid outcomes. When the student's first language is not English, outcomes that are not in the reading and writing domains can be administered in the student's first language.

## 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 in Section III: Subsection B.2 Sample Attrition: Is the combination of overall and differential attrition high? (pp. 11–15).

This review uses the ***liberal*** boundary for attrition. This boundary allows for higher levels of overall and differential attrition and was chosen based on the assumption that most attrition in studies of students with an ED was due to factors that were not strongly related to intervention status. For example, most attrition results from exogenous factors, such as parent mobility or absences on days that assessments are conducted. 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 attrition and whether it is high or low.

### Baseline Equivalence

If the study design is: 1) a randomized controlled trial with high levels of attrition; 2) a regression discontinuity design with high levels of attrition; or 3) a quasi-experimental design, the study 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. 15–16).

For studies that must demonstrate baseline equivalence, equivalence must be demonstrated for the intervention and comparison groups in the analytic sample on a baseline measure of the outcome or on another measure that is highly correlated with the outcome measure.

A study must also demonstrate baseline equivalence on problem behavior or social-emotional competence, even if neither is an outcome. If a study has a pre-intervention difference greater than 0.25 SD for *any* outcome in the problem behavior or social-emotional domain, the entire study is rated *Does Not Meet WWC Group Design Standards*. If a study requires an adjustment for *any* outcome in the problem behavior or social-emotional domain, all analyses in the study must adjust for this problem behavior or social-emotional outcome.

In a domain with an outcome that requires statistical adjustments, the adjustment is only required for that outcome, in addition to any necessary adjustments for problem behavior/social-emotional measures. For example, assuming there are no problem behavior or social-emotional measures that require an adjustment, if A, B, and C are available as pre- and post-intervention measures

and the pre-intervention difference in B requires statistical adjustment, only the analysis of outcome B must adjust for B.

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 very different settings and that the intervention and comparison groups are not sufficiently comparable for the purposes of this review:

- Percentage of sample identified as having an ED versus at risk for being identified
- Percentage of sample of each gender
- Percentage of sample speaking other languages
- Percentage of sample in each grade or of each age
- Percentage of sample from low-socioeconomic status families
- Percentage of sample from different racial/ethnic groups

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 outcomes 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 estimated effects using imputed 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).

Consistent with the *WWC Procedures and Standards Handbook*, the default intra-class correlations used for this review are 0.20 for reading, math, science, social studies, and writing outcomes, and 0.10 for problem behavior, school engagement, social-emotional, and self-determination outcomes.

**Other Study Designs**

Studies that use regression discontinuity or single-case designs are eligible for review using the appropriate design standards.

The *WWC Procedures and Standards Handbook* discusses the pilot standards for reviewing regression discontinuity design studies in Appendix D.

The *WWC Procedures and Standards Handbook* discusses the pilot standards for reviewing single-case design studies in Appendix E. The review of single-case design studies will follow the pilot standards with the following exceptions:

- There may be occasions when fewer than three data points in a phase would not automatically require the study to be rated as *Does Not Meet WWC Pilot Single-Case Design Standards*. In these circumstances, the content expert will consider the intervention and outcomes to determine whether one or two data points is required. Possible exceptions include the following:

  o Interventions for severe problem behavior, such as aggression and self-injury, in which extended initial baselines or reversal conditions pose serious ethical and procedural concerns.

  o Interventions on "zero baseline" behaviors or skills for which there is no logical reason to conceive that further assessment would yield other than zero baseline performance. An example of such a zero baseline performance is when a student is asked to provide a verbal label for an object ("what is this?") and consistently provides no response to the request because the student has little to no language and has never been observed to label the item or similar items.