WWC EVIDENCE REVIEW PROTOCOL FOR EARLY CHILDHOOD EDUCATION INTERVENTIONS, Version 3.0[1] (January 2014)

Topic Area Focus

What Works Clearinghouse (WWC) reviews in this topic area focus on early childhood education (ECE) interventions designed for use in school- or center-based settings with 3- to 6-year-old children who are not yet in kindergarten and are attending a center-based program. The primary focus for early childhood education interventions is on cognitive, language, and behavioral competencies associated with school readiness (specifically, language, cognitive, and social-emotional development, print knowledge, phonological processing, early reading and writing, and math; see definitions below).[2]

Systematic reviews of evidence in this topic area address the following questions:

- Which ECE interventions improve preschool children's school readiness (specifically, their language, cognitive, and social-emotional development, print knowledge, phonological processing, early reading and writing, and math; see definitions below)?

- Does the effectiveness of ECE interventions differ by type of outcome?

- Which ECE interventions are particularly effective for which children?

Key Definitions

Preschoolers. Preschoolers are 3- to 6-year-old children (i.e., children who are at least 3 years, 0 months old and not yet 6 years, 0 months old) who have not yet entered kindergarten, and children who are attending a program with a primary focus on cognitive, language, early literacy, and social-emotional competencies associated with school readiness.

Preschoolers with disabilities. Preschoolers with disabilities are children who meet the definition of preschoolers who are determined to be eligible for special education and related services under Part B of the Individuals with Disabilities Education Act (IDEA). This includes children who have been determined to have a disability or developmental delay based on state-defined eligibility criteria for preschool special education (Part B, Section 619 or extended Part C) under IDEA, and/or have scores on norm-referenced tests that are at least 2 standard deviations (SDs) below the mean in one developmental area or 1.5 SDs below the mean in two or more developmental areas, or have a delay of at least 25% in one developmental area or 20% in two or more developmental areas.

---

[1] This protocol is aligned with the *WWC Procedures and Standards Handbook* (version 3.0).
[2] Curricula or practices that have a primary focus on physical well-being and motor development are not included in this review.

English learners (ELs). ELs are students with a primary language spoken in the home other than English who have a limited range of speaking, reading, writing, and listening skills in English. This might include students who have been identified and determined by their program as having limited English proficiency (LEP) at the time of the study, or within the preceding 2 years.

Terms such as limited English proficiency (LEP), dual language learners (DLLs), English language learners (ELLs), non-English speakers, English as a second language (ESL), English for speakers of other languages (ESOL), language minority (LM), or second language learners (SLL) may also appear and should be brought to the attention of the lead methodologist for determination of eligibility.

ELIGIBILITY CRITERIA AND EVIDENCE STANDARDS

Populations to be Included

The Early Childhood Education topic area will review studies of interventions for 3- to 6-year-old children who are not yet in kindergarten and who are attending school- or center-based programs.[3] The children must attend a school- or center-based program in the United States, its territories or tribal entities, or in a country that is sufficiently similar to the United States that the study could be replicated in the United States (e.g., in which English is the societal language). To be included, the majority of the children in the study sample must speak English.

In cases where study authors provide aggregated data for children who are English speakers and ELs, and disaggregated data are unavailable, the ECE team will review the study as long as at least 50% of the children in the sample are *not* ELs, and the study meets the other research and intervention review requirements discussed below.

Similarly, if the authors provide aggregated data for preschool children with and without disabilities, the ECE team will review the study as long as at least 50% of the children in the sample do *not* have disabilities or delays, and the study meets the other research and intervention review requirements discussed below.[4]

In cases where the authors provide aggregated data for both preschool and kindergarten children and disaggregated data are unavailable, the ECE team will review the study as long as the majority of the children in the sample have not yet entered kindergarten and are attending a program that meets the requirements discussed in the section below, "Types of Interventions to be Included." If a study provides children's ages and not grade levels, the study will be reviewed in this area if at least 50% of the children are at least 3 years, 0 months and not yet 6 years, 0 months old. If the percentage of children in this age range is not reported, then the study will be reviewed if the mean age of the students in the sample is less than 5.0 years old.

---

[3] Studies with samples including at least 50% children in kindergarten may be reviewed by the Beginning Reading topic area.

[4] Studies with samples including at least 50% children with disabilities or delays would be reviewed by the Early Childhood Education for Children with Disabilities topic area.

Effectiveness of the intervention across different groups. An intervention's effectiveness could vary by subgroups defined by characteristics of children in the population or by characteristics of interventions' settings. In studies that present analyses of subgroups specified below, the subgroup analyses are also eligible for review in the ECE topic area. We will present findings for subgroups of interest in an appendix, provided the subgroup analyses meet topic area criteria and evidence standards with or without reservations. Potential subgroups of interest for this review include:

- Characteristics of children:
    - Age (for example, 3-year-old and 4-year-old subgroups);
    - Gender;
    - Socioeconomic status;
    - Race/ethnicity;
    - English language learner status;
    - Presence of a delay or disability.

- Characteristics of interventions' settings:
    - Location (e.g., urban, suburban, or rural);
    - Center type (e.g., child care center, school-based prekindergarten, Head Start, community-based preschool);
    - Staff education, qualifications (e.g., educational level, certification, years of experience), and training.

All WWC evidence standards apply to overall findings, as well as analyses of subsamples.

Types of Interventions to be Included

The interventions considered for inclusion are determined after a search of the published and unpublished literature by the Early Childhood Education Topic Area review team, as well as a review of nominations submitted to the WWC. The intervention must have a primary goal or goals of enhancing cognitive, language, literacy, math, or social-emotional competencies associated with school readiness, but it may additionally have goals that fall outside of these domains as well (though these will not be included in reviews). It must be implemented in a school- or center-based setting (e.g., child care center, school- or community-based preschool, or other center-based early childhood setting). The program may include other components (e.g., parent training, education), but only those interventions that are implemented primarily in the school- or center-based setting are included in the review.

Early childhood education interventions are curricula, practices, policies, and programs implemented in school- or center-based early childhood settings and designed to improve preschool children's cognitive, language, literacy, math, or social-emotional competencies associated with school readiness. Four broad types of interventions to be included in reviews are:

1. *Curricula.* A curriculum is a set of activities, materials, and/or guidance for working with children in classrooms that has a clearly identified name, includes a write-up/description,

and can be replicated by others based on written guidance, staff training, or technical assistance.

A curriculum may be (1) intended as the primary instructional tool designed to meet children's learning needs in multiple areas; or (2) designed to supplement the classroom material with differentiated instruction or meet children's learning needs in specific areas. Both types of curricula will be included in this review. Examples of early childhood education curricula include:

- A curriculum that fosters cognitive, language, social, physical, and emotional development of preschool children through a daily structure of thematic activities;
- A supplemental curriculum that features systematic, focused instruction in oral language, phonological and alphabetical awareness, and early reading concepts for preschool children and includes a teacher's guide and materials needed for the instruction; or
- A curriculum that consists of a set of guiding principles and practices that adults follow as they work with and care for preschool children. These principles are intended as an "open framework" that teams of adults are free to adapt to the special needs and conditions of their group, their setting, and their community.

2. *Practices.* The review will include both general and targeted practices. A general practice is a named approach to promoting children's development that program staff implements by interacting with children and materials in classrooms. A targeted practice is a named approach to promote the development of a subset of children in the classroom (e.g., ELs) or those with specific developmental issues such as giftedness, delays, or diagnosed disabilities. Both general and targeted practices must be clearly described and commonly understood in the field and literature. An example of an early childhood education practice is dialogic reading, a general practice that increases stimulation of children's language skills through interactive picture-book reading.

3. *Policies.* A policy is a named condition under which early childhood education programs operate. The policy must be commonly understood in the field and literature and directly affect services in preschool classrooms. Policies may be set by federal, state, or local governments or by the organization providing services. Examples of early childhood education policies include:

- Full-day or part-day program operation;
- Requirements for teachers to have a bachelor's degree or early childhood certification; or
- Class size limits or child-staff ratios.

4. *Programs.* A program is a service delivery model that may be associated with a funding stream and includes clear guidelines for implementation. Examples of early childhood education programs include Head Start or preschool programs established by states, such as New Jersey's Abbott preschool program or Oklahoma's state preschool program.

"Branded" and "non-branded" interventions
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.

Elements of intervention replicability
All reviewed curricula, practices, programs, and policies must be replicable (i.e., can be implemented by those other than the developers of the approach). The following characteristics of an intervention must be documented 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, and materials), unit of delivery of the intervention (for example, whole group, individual), medium/media of delivery (for example, teacher-led instruction or software), and targeted population;
- Intervention duration and intensity; and
- Description of individuals delivering or administering the intervention.

Types of Research Studies to be Included

To be included in the review, a study must meet several criteria for relevance:

- **Topic relevance.** The study must be focused on the effects of an approach to improving children's cognitive, language, print knowledge, phonological processing, early reading and writing, math, and social-emotional outcomes associated with school readiness.[5]
- **Time frame relevance.** The study has to have been published in 1985 or later. This time frame was established in order to define a realistic scope of work for the review. Rigorous evaluations of interventions implemented in this time frame test versions of interventions that are most likely to be available today and that were tested under conditions more likely to be similar to those existing today.
- **Sample relevance.** The study sample must meet the requirements described in the "Populations to be Included" section above. Outcomes may be measured later (e.g., when children are age 6 or older or attending a K–12 program).
- **Language relevance.** The study must be available in English to be included in the review. Also, studies examining competencies in other languages will not be included in the review.
- **Study location relevance.** The study must include children attending preschools a school- or center-based program in the United States, its territories or tribal entities, or in

---

[5] A main task for the WWC is to answer the question of intervention effectiveness. To this end, the WWC may use the data provided in studies differently than as presented by the study author.

a country that is sufficiently similar to the United States that the study could be replicated in the United States (e.g., in which English is the societal language).

- **Study design relevance.** The study must be empirical, using quantitative methods and inferential statistical analysis, and must take the form a randomized controlled trial (RCT) or use a quasi-experimental design (QED), a regression discontinuity design (RD), or a single-case design (SCD).
- **Outcome relevance.** The study must focus on child outcomes, rather than teacher outcomes, and it must include at least one outcome measure in at least one of the following domains with adequate face validity and reliability: cognitive, social-emotional, or language development, print knowledge, phonological processing, early reading and writing, or math (described in the next section "Outcome Measures").

In the overall rating of effectiveness, reviews in the Early Childhood Education topic area include any intervention comparison that permits estimation of the effects of the intervention. In some cases, this means that an entire intervention is compared to a no-treatment or business-as-usual comparison group (e.g., the typical preschool curriculum). In other cases, it means that the additive effects of a particular component of an intervention (e.g., adult interaction with shared book reading) will be examined in relation to the intervention in absence of that additive component (e.g., shared book reading). It can also mean that one intervention is compared to another.

Outcome Measures

Types of Outcomes to be Included. The primary outcome domains for the ECE topic area are cognition, mathematics, social-emotional development, language development, alphabetics, fluency, comprehension, and general reading achievement. Additional information about each of these domains is described below:

- *Cognition domain.* Includes outcomes in the following areas: memory, problem-solving, cognitive processing and flexibility, general knowledge, and IQ.

- *Mathematics domain.* Includes outcomes in the following areas: basic number concepts (e.g., number recognition, concepts of numerical order and one-to-one correspondence, counting, magnitude, number line estimation, and finding counting mistakes); number operations (e.g., age-appropriate addition or subtraction problems); patterns and classification (e.g., the ability to identify patterns in a set of objects, duplicate patterns provided by another person, create and replicate demonstrated patterns, and sort objects by similar and different characteristics); measurement; geometry (e.g., the identification of shapes and shape attributes, like angles and corners, as well as spatial relationships); and general numeracy (e.g., outcome measures that cover two or more of the previous content areas).

- *Social-emotional development domain.* Includes outcomes in the following areas: behavioral, social, and emotional competencies underlying school readiness, such as pro-social (or problem) behaviors, social interactions,

6

cooperation, self-concept, engagement, attention, persistence, impulsivity, self-control, and initiative.

- *Language development domain.* Includes outcomes that assess the ability to understand spoken language, communicate and understand thoughts or ideas through speech, use developmentally-appropriate discourse skills, and display grammatical knowledge or skill.

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

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

    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).

    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. Spelling is included in the review as an acceptable phonics outcome.

- *Fluency domain.* 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.

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

  - **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.

  - **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.

- *General reading achievement domain.* Includes outcomes that combine measures in two or more of the previous 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.

Overalignment of outcome measures. A study's rating will be based only on those measures that are not overaligned. Overalignment occurs when outcome measures are more closely aligned to one of the research groups (intervention or comparison) than the other and could bias a study's results. For instance, if the outcome measure assesses phoneme recognition using some of the materials included in the intervention (such as specific pictures used to practice learning phoneme sounds or words), it is considered to be overaligned with the intervention. In these situations, the intervention group may have an unfair advantage over the comparison group, and the effect size is not a fair indication of the intervention's effects.

Reliability and validity of outcome measures. Measures of the outcome of interest should demonstrate adequate reliability and face validity. Reliability (internal consistency, temporal stability/test-retest reliability, and inter-rater reliability) will be assessed using the following standards:

- Internal consistency: minimum of 0.60;

- Temporal stability/test-retest reliability: minimum of 0.40; or

- Inter-rater reliability: minimum of 0.50 (percent agreement, correlation, Kappa).

If the reliability of each outcome measure is not specified in the research article, data from the test or scale's publisher or other sources, including an author query, may be used to establish the reliability of an outcome measure.

In addition to meeting the reliability requirement above, each outcome measure must also have face validity. Also, there may be outcomes for which reliability cannot be assessed (e.g., number

of words read per minute, counting, etc.). For outcomes such as these, the lead methodologist will determine if the outcome is eligible for review.

For **single-case design research (SCD)**, the outcome variable must be measured systematically over time by more than one assessor, and the study needs to collect inter-assessor agreement in all phases and at least 20% of all sessions (total across phases) for each condition (e.g., baseline, intervention).

According to WWC standards, SCD studies must demonstrate reliability of outcomes through an inter-observer assessment (IOA)[6]. Inter-observer reliability is particularly relevant for measures that require an observer to score another person's behavior or complete a rating or checklist describing the behavior observed. Outcomes that involve written responses need not meet the SCD IOA requirement if the Lead Methodologist determines that the responses can be scored by a single coder with a high degree of reliability. An example of outcomes that would not require IOA is written responses to a reading test. An example that would require IOA is coding of a child's observed behavior.

The interval for measuring post-intervention effects. The benefits of an early childhood education intervention are intended to be retained well past the end of the intervention. Thus, measures at the end of an intervention, as well as any time thereafter, are admissible. The ECE team prioritizes immediate posttest findings for developing intervention ratings and improvement indices because these findings are most prevalent in ECE studies. Measures occurring several months or years after the intervention may provide strong evidence for an intervention's effectiveness. Therefore, the ECE team includes follow-up findings, when available and appropriate, in supplemental appendices to the intervention report.

Statistical and Analytic Issues

Attrition in RCTs

The WWC considers both the overall sample attrition rate and the difference in sample attrition between the intervention and comparison groups, as both contribute to the potential bias of the estimated effect of an intervention. The WWC has established conservative and liberal standards for acceptable levels of attrition. The conservative standards are applied in cases where the lead methodologist has reason to believe that much of the attrition can be attributed to the intervention reviewed—for example, high school students choosing whether or not to participate in a dropout prevention program. The liberal standards are applied in cases where the lead methodologist has reason to believe that little of the attrition is endogenous to the intervention reviewed. Attrition rates are based on the number of sample cases used in the analysis sample with measured, as opposed to imputed, values of the outcome measures.

The Early Childhood Education topic area uses the **liberal** standard. This reflects the assumption that most attrition in studies of early childhood interventions results from exogenous factors,

---

[6] See "http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_standards_handbook.pdf, Appendix E".

such as parent mobility or absences on days that assessments are conducted. Table 1 presents the maximum difference in the attrition rate for the intervention and comparison group that is acceptable for a given level of overall sample attrition. The empirical basis for these thresholds is described in *Assessing Attrition Bias.*[7]

Studies based on cluster random assignment designs must meet attrition standards for both the study sample units that were assigned to intervention or comparison group status (e.g., schools or districts) and the study sample units for analysis (e.g., typically, students). In applying the attrition standards to the subcluster level (e.g., students), the denominator for the attrition calculation includes only sample members in the clusters that remained in the study sample.

RCTs with combinations of overall and differential attrition rates that exceed the applicable threshold, based on the applicable standard, must demonstrate baseline equivalence of the analysis sample, or, if nonequivalence falls within the allowable range, statistically control for that nonequivalence, in order to receive the rating of *Meets WWC Group Design Standards With Reservations*. See the "Baseline Equivalence" section for more details.

---

[7] http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=243

Table 1: Attrition Standards for Randomized Controlled Trials

Highest Level of Differential Attrition Allowable to Meet the Attrition Standard Under the Liberal Attrition Standard

Overall Attrition	Allowable Differential Attrition		Overall Attrition	Allowable Differential Attrition
0	10.0		34	7.4
1	10.1		35	7.2
2	10.2		36	7.0
3	10.3		37	6.7
4	10.4		38	6.5
5	10.5		39	6.3
6	10.7		40	6.0
7	10.8		41	5.8
8	10.9		42	5.6
9	10.9		43	5.3
10	10.9		44	5.1
11	10.9		45	4.9
12	10.9		46	4.6
13	10.8		47	4.4
14	10.8		48	4.2
15	10.7		49	3.9
16	10.6		50	3.7
17	10.5		51	3.5
18	10.3		52	3.2
19	10.2		53	3.0
20	10.0		54	2.8
21	9.9		55	2.6
22	9.7		56	2.3
23	9.5		57	2.1
24	9.4		58	1.9
25	9.2		59	1.6
26	9.0		60	1.4
27	8.8		61	1.1
28	8.6		62	0.9
29	8.4		63	0.7
30	8.2		64	0.5
31	8.0		65	0.3
32	7.8		66	0.0
33	7.6		67	-

Baseline Equivalence

RCTs with high attrition and all QEDs must demonstrate baseline (that is, pre-intervention) equivalence between the intervention and comparison groups in the analytic sample in order to receive the rating of *Meets WWC Group Design Standards With Reservations*. The WWC assesses

baseline equivalence based on the magnitude of the difference between the intervention and comparison groups; statistical significance of the difference in means is not considered.

For the Early Childhood Education topic area, for the groups to be considered equivalent, the groups must be similar on the pretest score of each outcome measure, or on another measure that is highly related to the specific outcome measure. It is possible for a study to meet evidence standards in some domains and not in others. Thus, rules for establishing baseline equivalence are applied *within each domain*.

Groups are considered equivalent in a given domain if the reported differences between the groups are less than or equal to 0.05 SD in the sample on pre-intervention measures of each outcome measure in the domain. If this is the case, the equivalence standard is met, and the study can receive a rating of *Meets WWC Group Design Standards With Reservations* for that domain.

If differences are greater than 0.05 SD and less than or equal to 0.25 SD, the study findings must be based on analytic models that control for the individual-level baseline measure(s) on which the groups differ in order to receive a rating of *Meets WWC Group Design Standards With Reservations*. Otherwise, the domain is rated *Does Not Meet WWC Group Design Standards*.

If baseline differences are greater than 0.25 SD for *any* pre-intervention measure within a domain, then all outcomes within that domain, for that contrast, do not meet the baseline equivalence standard, regardless of whether or not the impacts are estimated using models that control for baseline measures. Therefore, for that analytic contrast, the domain is rated *Does Not Meet WWC Group Design Standards*.

For instance, if any of the outcomes in the language development domain are assessed with baseline differences greater than 25% of the pooled standard deviation (or 0.25 SD), all of the outcomes in that domain would be rated *Does Not Meet WWC Group Design Standards*.

Given the potential for selection bias in QEDs and RCTs with high attrition, the possibility that the intervention and comparison groups were drawn from different populations is also a concern. Fundamental differences in the settings from which the intervention and comparison groups in a QED study were drawn or baseline differences in the characteristics of the intervention and comparison groups in QEDs and RCTs with high attrition may indicate that the children in the two groups represent very different populations, even if they are equivalent on pretest measures. When there is evidence that the populations being compared are drawn from very different settings, the study will be referred to the review team leadership, who will determine whether the settings are too dissimilar to provide an adequate comparison. If the leadership decides that they are too dissimilar, the study is rated *Does Not Meet WWC Group Design Standards*. These characteristics include, but are not limited to, the percentage of children from:

- Low-SES families;
- Racial/ethnic groups;
- Special education classifications; and
- Different locations (e.g., urban, rural).

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

Other Statistical and Analytical Issues

RCT studies with low attrition do not need to use statistical controls in the analysis, although statistical adjustment for well-implemented RCTs is permissible and can help generate more precise effect size estimates. For RCTs that do not include statistical controls for the pretest in the analysis of effects of the intervention, the effect size estimates may be adjusted for differences in pre-intervention characteristics at baseline (if available) using a difference-in-differences method if the authors did not adjust for pretest (see Appendix B of the *Handbook*). Authors may make additional statistical adjustments that are not required by the WWC evidence standards.

For the WWC review, the preference is to report on and calculate effect sizes for post-intervention means adjusted for the pre-intervention measure. If a study reports both unadjusted and adjusted post-intervention means, the WWC review will report the adjusted means and unadjusted standard deviations. If effect sizes or the information required to calculate them are not reported, then the missing information will be requested from the author(s).

The statistical significance of group differences will be recalculated if (a) the study authors did not calculate statistical significance, (b) the study authors did not account for clustering when there is a mismatch between the unit of assignment and unit of analysis, or (c) the study authors did not account for multiple comparisons when appropriate. Otherwise, the review team will accept the *p*-values provided in the study.

When the unit of analysis in the study is not the same as the unit of assignment the effect sizes computed by the WWC will incorporate a statistical adjustment for clustering. The default intraclass correlation used for the Early Childhood Education topic area review is 0.10 for the outcomes in the social-emotional development domain and 0.20 for outcomes in all other domains.[8] For an explanation about the clustering correction, see Appendix C of the *Handbook*.

When multiple comparisons are made (that is, multiple outcome measures are assessed within an outcome domain in one study) and not accounted for by the authors, the WWC accounts for this multiplicity by adjusting the reported statistical significance of the effect using the Benjamini-Hochberg correction. See Appendix D of the *Handbook* for the formulas the WWC uses to adjust for multiple comparisons.

---

[8] For intraclass correlation estimates for preschool cognitive outcomes, see Schochet, P. Z. (2005). *Statistical power for random assignment evaluations of education programs*. Princeton, NJ: Mathematica Policy Research.

LITERATURE SEARCH METHODOLOGY

The WWC literature search is comprehensive and systematic. Detailed protocols guide the entire literature search process. At the beginning of the process, relevant journals, organizations, and experts are identified.[9] The WWC searches core sources and additional topic-specific sources identified by the lead content expert. For the Early Childhood Education topic area, different search strategies were used for the original literature search for the area conducted from 2002 to 2007, the subsequent searches to update literature on curricula originally identified from 2008 to 2011, and the prioritization search conducted in 2012. For details of the searches prior to 2012, please see ECE review protocol versions 1.0 and 2.0.

In 2012, the WWC staff members conducted a broad search of the literature for the ECE area in order to identify new curricula and practices that had emerged since the original prioritization search in 2003–04, and to include early childhood policies in the search. For this search, WWC staff used a comprehensive list of search terms and strategies designed to identify literature pertaining to a broad range of programs and practices in the ECE topic area. WWC staff used an extensive list of keywords to search electronic databases, conducted targeted searches of specific curricula identified through the keyword search and by the suggestion of content experts, and searched the websites of organizations that conduct research on early childhood education programs or advocate for early childhood programs.

1. KEYWORD SEARCH PARAMETERS AND DATABASES

The primary objective of the keyword search was to identify interventions with potentially eligible studies and assess the number of studies on each intervention, so that interventions could be prioritized for review. The WWC team identified keywords to capture the breadth of literature that falls within the scope of the protocol. Targeted outcomes and study design terms were included to focus the search on identifying literature that will support an intervention report.

Most of the literature search used the following search terms:

1. **Targeted Outcomes.** These potential outcomes were listed with OR between each word.

Academic	Geometry	Phonological
Achievement	Initiative	Print*
Alphabet*	Intellig*	Problem solving
Attention	IQ	Reading
Behavior*	Knowledge	Reasoning
Blend*	Language	Rhym*
Cognit*	Learning	School readiness
Communication	Letter*	Segment*
Comprehen*	Listening	Self-regulation

---

[9] Conference proceedings and master's theses are not included in literature searches.

Concept*	Literacy	Shapes
Cooperat*	Math*	Social*
Counting	Measure*	Sort*
Decod*	Memory	Sound-letter
Early literacy	Number*	Speech
Emergent literacy	Numer*	Syllab*
Emotion*	Oral	Vocabulary
Engagement	Pattern*	Writing
Executive function	Persistence	
Fluency	Phonem*	

2. **Target Ages and Programs.** These target age groups and programs were listed with OR between each word.

Child care	Pre-K*
Childcare	PreK*
Early intervention	Preschool
Early childhood	Pre-school

3. **Interventions.** These target intervention activities were listed with OR between each word.

Intervention*	Instruct*
Curricul*	Teach*
Program*	Train*
Strateg*	Technique*
Approach*	

4. **Study Design.** These target study design terms were listed with OR between each word.

ABAB design*	Pretest
Alternating treatment*	Pre-test
Causal	QED
Comparison group*	Quasi-experimental design*
Control group*	Random*
Effectiveness	Randomized control* trial*
Experiment*	RCT
Impact	Regression discontinuity design*
Matched group*	Simultaneous treatment*
Meta-analysis	Single case design*
Meta analysis	Single subject design*
Posttest	
Post-test	

Each of these four groups of keyword terms was joined by AND. These terms were used to search the electronic databases listed below.

Databases

WWC staff searched electronic databases using the search terms listed above. Staff used OVID to search Psychinfo; EBSCO to search Academic Search Premier, EconLit, Education Research Complete, ERIC, Socindex with full-text, and EJS E-journals; ProQuest to search Ph.D. dissertations; and WorldCat to search the holdings of 10,000 libraries.

a. ***Academic Search Premier.*** This multi-disciplinary database provides full text for more than 4,500 journals, including full text for more than 3,700 peer-reviewed titles. PDF backfiles to 1975 or further are available for well over one hundred journals, and searchable cited references are provided for more than 1,000 titles.

b. ***EconLit.*** EconLit, the American Economic Association's electronic database, is the world's foremost source of references to economic literature. The database contains more than 785,000 records from 1969–present. EconLit covers virtually every area related to economics.

c. ***Education Research Complete.*** Education Research Complete is the definitive online resource for education research. Topics covered include all levels of education from early childhood to higher education, and all educational specialties, such as multilingual education, health education, and testing. Education Research Complete provides indexing and abstracts for more than 1,840 journals, as well as full text for more than 950 journals, and includes full text for more than 81 books and monographs, and for numerous education-related conference papers.

d. ***EJS E-Journals.*** E-Journals from EBSCO host® provides article-level access for thousands of E-Journals available through EBSCO's Electronic Journal Service (EJS).

e. ***ERIC.*** Funded by the U.S. Department of Education (ED), ERIC is a nationwide information network that acquires catalogs, summarizes, and provides access to education information from all sources. All ED publications are included in its inventory.

f. ***ProQuest Dissertations.*** This database provides access to dissertations; the WWC searches are limited to Ph.D. dissertations.

g. ***PsycINFO.*** PsycINFO contains more than 1.8 million citations and summaries of journal articles, book chapters, books, dissertations and technical reports, all in the field of psychology. Journal coverage, which dates back to the 1800s, includes international material selected from more than 1,700 periodicals in over 30 languages. More than 60,000 records are added each year.

h. ***Sage Journals online.*** This database provides access to full-text articles published in over 500 SAGE journals. Articles published in 1999 or later are available as full-text reports. It includes all of the journals published by the American Educational Research Association (AERA), as well as many leading titles in psychology, early childhood, and survey methodology.

i. ***Scopus.*** The world's largest abstract and citation database of peer-reviewed literature and quality web sources in the scientific, technical, medical, and social sciences. It covers 19,000+ titles, articles in press, conference proceedings, and e-books. Subjects covered in the database include disability, health, nutrition, statistics, and survey. Scopus provides tools to track, analyze, and visualize research and conduct citation analysis.

j. ***SocINDEX with Full Text.*** SocINDEX with Full Text is the world's most comprehensive and highest quality sociology research database. The database features more than 1,986,000 records with subject headings from a 19,600+ term sociological thesaurus designed by subject experts and expert lexicographers. SocINDEX with Full Text contains full text for 708 journals dating back to 1908. This database also includes full text for more than 780 books and monographs, and full text for 9,333 conference papers.

k. ***WorldCat.*** WorldCat is the world's largest network of library content and services, and allows users to simultaneously search the catalogs of over 10,000 libraries, containing over 1.2 billion books, dissertations, articles, CDs, and other media.

In addition, WWC staff searched the Campbell Collaboration library using the target age terms (for example, preschool) from 2002 through 2012.

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

Finally, WWC staff used a custom Google Scholar search to search "grey" literature (papers and reports not published in peer-reviewed journals) across URLs. Search terms included preschool AND effectiveness; preschool AND random; and preschool AND "quasi-experimental design".

In addition to the database and Google searches, the review team identifies other relevant studies through the following approaches:

a. Public submissions:

   1) Materials submitted via the WWC website
   2) Materials submitted directly to WWC staff

b. Solicitations made to key researchers by the review team

c. Checking websites summarizing research on programs for children and youth, prior reviews, and research syntheses (i.e., using the reference lists of prior reviews and research syntheses to make sure key studies have not been omitted)

d. Searches of the websites of all the developers of relevant interventions or practices for any research or implementation reports

e. Searches of the websites of over 50 think tanks, research centers, and associations that conduct research in this topic area (see "Research Organizations" below)

References resulting from these searches are screened and sorted by intervention.

2. INTERVENTION SEARCH

Once a keyword search has been conducted and interventions prioritized, the next search is designed to identify all effectiveness studies conducted for a specific intervention.

Search Strategy:

- Conduct database search on the intervention name (e.g., *Ladders to Literacy*).[10]
- For practices, scan references to identify possible synonyms for the practice name in the literature (e.g., shared book reading). Conduct database searches of these terms.
- Request full text of potentially eligible studies and review the reference lists to cross-check search results. Similarly, review relevant literature reviews.
- Identify seminal researchers associated with the intervention. Conduct full text searches of the researcher name combined with the intervention name (e.g., Adele Diamond AND *Tools of the Mind*).

All references resulting from these searches will be screened for eligibility.

3. RESEARCH ORGANIZATIONS

The websites of the research organizations conducting studies related to early childhood education will be reviewed to identify studies for this review. Examples of these research organizations include:

Abt Associates
After-School Alliance
Alliance for Excellent Education
American Education Research Association
American Enterprise Institute
American Institutes of Research
American Speech-Language-Hearing Association (ASHA)
Best Evidence Encyclopedia
Broad Foundation (Education)
Brookings Institution
Campbell Collaboration
Carnegie Corporation for the Advancement of Teaching
Carnegie Corporation of New York
Center for Research and Reform in Education
Center for Research in Educational Policy (CREP)

---

[10] A standard library search consists of searching titles and abstracts in each of the databases described above.

Center on Education Policy
Center on Instruction
Chapin Hall Center for Children at the University of Chicago
Child Care and Early Education Research Connections
Congressional Research Service (via OpenCRS.org)
Council for Exceptional Children
Council for Learning Disabilities
Erikson Institute, University of Chicago
Florida Center for Reading Research (FCRR)
Frank Porter Graham Child Development Institute
Government Accountability Office (GAO)
Harvard Family Research Project
Harvard Graduate School of Education
Head Start Early Childhood Learning & Knowledge Center
Heritage Foundation
Hoover Institution
Institute for Public Policy and Social Research (IPPSR)
Johns Hopkins University School of Education
Learning Disabilities Association of America
Linguistic Society of America (LSA)
Mathematica Policy Research
MDRC
Mid-continent Research for Education and Learning
National Association for Bilingual Education (NABE)
National Association for the Education of Young Children
National Association of State Boards of Education
National Association of State Directors of Special Education
National Center for Children in Poverty
National Center for Education Research
National Center for Learning Disabilities
National Center for Research on Early Childhood Education (NCRECE)
National Center for Special Education Research
National Center on Response to Intervention (RTI)
National Child Care Information Center
National Dissemination Center for Children with Disabilities
National Early Childhood Technical Assistance Center (NECTAC)
National Early Literacy Panel
National Governors' Association
National Head Start Association
National Institute for Early Education Research (NIEER)
National Institute on Out-of-School Time at the Wellesley Centers for Women
National Reading Panel

NBER Working Papers

New America Foundation's Early Education Initiative

NHSA Dialog (Head Start Association Journal)

Office of Early Learning

Pacific Resources for Education and Learning

Policy Study Associates

PolicyArchive

Promising Practices Network

Public Education Network

Public Policy Research Institute at Texas A&M University

Public/Private Ventures (PPV)

RAND

REL Appalachia (contractor: CNA)

REL Central (contractor: Marzano Research Laboratory)

REL Mid-Atlantic (contractor: ICF International)

REL Midwest (contractor: American Institutes for Research [AIR])

REL Northeast and Islands (contractor: Education Development Center, Inc.)

REL Northwest (contractor: Education Northwest)

REL Pacific (contractor: Mid-continent Research for Education and Learning [McREL])

REL Southeast (contractor: Florida State University)

REL Southwest (contractor: REL Southwest at SEDL)

REL West (contractor: WestEd)

Society for Research in Child Development

Southwest Educational Development Laboratory (SEDL)

SRI

Teachers of English to Speakers of Other Languages (TESOL)

Technical Assistance Center on Social Emotional Intervention for Young Children

The Education Resources Institute

Thomas B. Fordham Institute

U.S. Department of Education (includes Institute for Education Sciences, National Center for Special Education Research, etc.)

U.S. Department of Health & Human Services (including: Office of Head Start [OHS], Administration for Children and Families [ACF], Office of the Assistant Secretary for Planning and Evaluation [ASPE], Office of Planning, Research, & Evaluation [OPRE], Office of Child Care [OCC], and the National Institutes of Health [NIH])

Urban Institute

WestEd

Curricula and Practice Names

Primary Curricula

A Beka
Bank Street Developmental Interaction Approach
Beyond Centers and Circle Time
Breakthrough to Literacy
Bright Beginnings
Building Blocks for Literacy
Core Knowledge Preschool Sequence
Creative Curriculum
Curiosity Corner
DLM Early Childhood Express
Doors to Discovery
Early Literacy and Learning Model (ELLM)
FunShine Express: Fireflies/Sprouts
Funsteps, Inc.
Growing Readers Early Literacy Curriculum
    (High/Scope)
High Reach
High/Scope Curriculum
Innovations Comprehensive Preschool
    Curriculum (Gryphon House Pub.)
Language-Focused Curriculum

Language for Learning
Learning Experiences: An Alternative Program
    for Preschoolers and Parents (LEAP)
Let's Begin with the Letter People
Literacy Express
Marazon system
Montessori Method
Opening the World of Learning
Pebble Soup
Primrose Schools
Project Construct
Read, Play, and Learn!
Ready, Set, Leap!
Reggio Emilia
Saxon Early Learning
Scholastic Early Childhood Program
School Readiness Express
S.P.A.R.K.
Tools of the Mind
We Can! Curriculum
Wee Learn

Supplemental Curricula

Active Learning
Big Math for Little Kids
Building Blocks for Math (SRA Real Math)
Building Early Literacy and Language Skills
    (BELLS)
Building Language for Literacy (BLL Scholastic)
Compass Learning Odyssey Pre-K/K
Daisy Quest
Direct Instruction
Direct Instruction Math
Fast ForWord Preschool
Headsprout Reading Basics
Houghton Mifflin PreK
Journeys into Early Literacy (precursor to
    Destination Reading)
Kaplan Planning Guide to the Preschool
    Curriculum
Ladders to Literacy: A Preschool Activity Book
LeapDesk Workstation
Learningames—Abecedarian

Lidcombe Program
Lindamood Phoneme Sequencing Program (LiPS)
Links to Literacy Curriculum Kit
My Magic Story Car
Open Court Reading (OCR) Pre-K
Phonemic Awareness in Young Children: A
    Classroom Curriculum
Phono-Graphix
Pre-K Mathematics
Project Approach
ReadingLine Kits
Rightstart/Numberworlds
ScienceStart!
Sing, Spell, Read, & Write
Sound Foundations
Sounds Abound
Spell, Read, PAT
Stepping Stones to Literacy
Waterford Early Reading Program Pre-K
Words and Concepts

General Practices

Dialogic Reading/Interactive Shared Picture-Book Reading

Letter Knowledge Training
Phonological Awareness Training
Shared Book Reading

Programs

Head Start

Targeted Practices

Classwide peer tutoring
Conversation-based language intervention
Conversational-recasting
Explicit attention to articulation
Functional communication training
Graphics-based software tools
Imitation-based language intervention
Peer-mediated intervention
Peer training
Pragmatic teaching
Redirects
Self-initiated augmentative communication
     treatment

Stimulus control procedure
Syntax program
Teaching phonological awareness
Teaching rhyming
Teaching-script
Teaching story grammar knowledge
Text-based software tools
Time delay
Verbal labeling responses
Video discourse intervention
Written text cueing