WWC EVIDENCE REVIEW PROTOCOL FOR BEGINNING READING INTERVENTIONS

Topic Area Focus

The What Works Clearinghouse (WWC) review focuses on reading interventions for students in grades K–3 (or ages 5-8) that are intended to increase skills in **alphabetics** (phonemic awareness, phonological awareness, letter recognition, print awareness and phonics), **reading fluency, comprehension** (vocabulary and reading comprehension), or **general reading achievement** (see definitions below). Systematic reviews of evidence in this topic area address the following questions:

- Which interventions intended to provide basic literacy instruction improve reading skills (including alphabetics, fluency, comprehension or general reading achievement) among students in grades K–3 or ages 5-8?

- Are some interventions more effective than others at improving certain types of reading skills?

- Are some interventions more effective for certain types of students, particularly students who have historically lagged behind in reading achievement?

Individual intervention-level reports are released on a periodic basis; one topic-level report will subsequently be released.

Key Definitions

Alphabetics Domain

Phonemic awareness. 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").

Phonological awareness. Phonological awareness is a more encompassing term than phoneme/phonemic awareness (PA). Phonological awareness is a term referring to various types of awareness, which includes PA *and* also 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?). Tasks that require students to manipulate spoken units larger than phonemes are simpler for beginners than tasks requiring phoneme manipulation (Liberman, Shankweiler, Fischer, & Carter, 1974).

Letter Identification. Knowing the names of the letters of the alphabet supports reading acquisition. Letter-naming measures have been shown to be predictors of reading development especially when letter naming is taught in conjunction with other beginning reading skills.

Print Awareness. Print awareness refers to knowledge or concepts about print such as (a) print carries a message; (b) there are conventions of print such as directionality (left to right, top to bottom), differences between letters and words, distinctions between upper and lower case, punctuation; and that (c) books have some common characteristics (e.g. author, title, front/back). It has been shown that print awareness supports reading acquisition (e.g., decoding).

Phonics. Phonics[1] refers to (a) the 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) the ability to associate letters and letter combinations with sound and blending them into syllables and words; and (c) the understanding that this information can be used to decode or read words.

Reading Fluency Domain

Reading fluency. Fluency is the ability to read text accurately, automatically, and with expression, while still extracting meaning from it.

Comprehension Domain

Vocabulary development. This 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. Reading comprehension refers to the understanding of the meaning of a passage and the context in which the words occur. Reading comprehension is composed of two equally important components. Decoding, or the ability to translate text into speech, is only part of the process of reading comprehension. The other part is language comprehension, or the ability to understand spoken language. All struggling readers have difficulty with either language comprehension or decoding or both.

General Reading Achievement Domain

General reading achievement. Outcomes that fall in the general reading achievement domain are those that inextricably combine two or more of the previous domains (alphabetics, reading fluency, and comprehension) or provide some other type of summary score, such as a "total reading score" on a standardized reading tests, grades in reading or language arts class, or promotion to the next grade.

---

[1] "Phonics" also refers to an instructional approach that focuses on the correspondence between sounds and symbols and is often used in contrast to whole language instructional approaches. For the purposes of the Beginning Reading Evidence Reports, we use the term phonics as defined above, not as an instructional approach.

GENERAL INCLUSION CRITERIA

Populations to be Included

The beginning reading reviews address interventions designed for children in grades K–3 (or 5 to 8 years of age, which may include children beginning to read in preschool). Because students with learning disabilities, students living in high poverty, and minority children lag behind the population as a whole in reading achievement, studies involving these groups are of particular interest.

Types of Interventions to be Included

The specific interventions considered for inclusion are determined after an exhaustive search of the published and unpublished literature by the Evidence Report Team, as well as a review of nominations submitted to the WWC. Only research on interventions that are replicable (that is, the intervention can be reproduced) are reviewed. The types of interventions included are as follows:

- Programs/products, such as
  - Comprehensive, non-textbook based programs—some of which are curriculum-based, others focusing on staff development—intended to serve as a school's prime K–3 literacy instruction program, such as Exemplary Center for Reading Instruction, Direct Instruction (using SRA Reading Mastery materials), Renaissance Learning, Success for All, Voyager Universal Literacy Program, Waterford Early Reading program
  - Supplemental K–3 reading programs intended to enhance whole-school literacy
  - K–3 reading basals/textbooks intended for whole-school, whole-classroom use, such as Houghton-Mifflin, SRA/McGraw-Hill Reading Mastery, SRA/McGraw-Hill Open Court, Harcourt
  - Software designed to improve reading skills
- Practices (e.g., reading aloud; home literacy environments)
- Policies (e.g., mainstreaming students with learning disabilities)

Types of Research Studies to be Included

The beginning reading literature search focuses on studies involving reading programs, practices, policies, and products for children in grades K–3. To be included in the review, a study must meet several relevancy criteria:

- ***Topic relevance.*** The study has to be about reading or literacy, focusing on alphabetics, reading fluency, comprehension or general reading achievement. The study is required to focus on the effects of an intervention, not on

  - individual differences (e.g., correlational studies examining the covariance between reading speed and performance on a reading test; studies focusing on brain functions or structures), or

  - assessment (e.g., on properties of an instrument or approaches to identifying students with learning disabilities).

- ***Timeframe relevance.*** The study has to have been published in 1983 or later. Please note that because of the difficulty in reaching authors of conference papers, we altered our timeframe for those searches, and limited our review of conference proceedings to those that took place from 1998 to 2005.

- ***Sample relevance.*** The sample must include students in grades K, 1, 2, or 3[2] learning to read in English.

  - The intervention must have taken place in grades K, 1, 2, or 3; outcome may be measured in grades K–3 or later.

  - Studies that focus on literacy in other languages are excluded (e.g., phonemic awareness among early Spanish readers).

  - Studies that focus on early literacy among English language learners are excluded from consideration, to limit the scope of this Evidence Report .

- ***Study design relevance.*** Study design and focus are limited to manuscripts that

  - are empirical studies using quantitative methods and inferential statistical analysis, and

  - take the form of a randomized controlled trial, or use a regression-discontinuity design, a quasi-experimental design, or a single-subject[3] design.

- ***Outcome relevance.*** The study is required to

  - focus on student, not teacher, or other outcomes, and

  - include at least one relevant outcome that demonstrated adequate face validity or reliability.

---

[2] Interventions in PreK may be included if the children are 5 years old or older.

[3] At this time, the WWC has not developed standards for reviewing or reporting on regression-discontinuity or single-case design studies. Consequently, studies with regression-discontinuity or single-case designs are not addressed in this review.

SPECIFIC TOPIC PARAMETERS

The following parameters specify which studies are considered for analyses and which aspects of those studies are coded for the review.

1. Characteristics of beginning reading interventions.

We define reading interventions for students in grades K–3 as programs/products, practices, or policies that are intended to increase skills in alphabetics (phonemic awareness, phonological awareness, letter recognition, print awareness or phonics), reading fluency comprehension (vocabulary or reading comprehension), or general reading achievement.

Programs/products include:

- Comprehensive, non-textbook based programs—some of which are curriculum-based, others focusing on staff development—intended to serve as a school's prime K–3 literacy instruction program, such as Direct Instruction (using SRA Reading Mastery materials), Exemplary Center for Reading Instruction, Renaissance Learning, Success for All, Voyager Universal Literacy Program, Waterford Early Reading Program
- Supplemental K–3 reading programs intended to enhance whole-school literacy
- K–3 reading basals/textbooks intended for whole-school, whole-classroom use, such as Houghton-Mifflin, SRA/McGraw-Hill Reading Mastery, SRA/McGraw-Hill Open Court, Harcourt
- Software designed to improve reading skills

- *Practices* include reading aloud; home literacy environments.
- *Policies* include mainstreaming students with learning disabilities.
- Variations across programs, products, practices, and policies include:
  - Targeting to specific populations (e.g., readers below grade level; at-risk students)
  - Intention to be a school's primary literacy instruction program, versus a supplemental reading program
  - Relative emphasis on implementing a packaged curriculum versus provision of professional development
  - Relative emphasis on a phonics versus whole language versus a "balanced" approach
  - Relative emphasis on home-school connections

  - Different level of implementation (national; statewide; districtwide; schoolwide; individual grades; whole group; small group; one-on-one)
  - Different medium/media through which program is implemented (e.g., basal readers; computer software)
  - Relative emphasis on enhancing specific beginning reading outcomes (e.g., phonemic awareness and phonics skills versus text comprehension skills)

- We make the additional distinction between "branded" and "non-branded" interventions. Branded interventions are commercial programs and products that may possess any of the following characteristics:
  - Have an external developer who:
    - Provides technical assistance (e.g., provides instructions/guidance on the implementation of the intervention)
    - Sells or distributes the intervention
  - Replicable: packaged or otherwise available for distribution/use beyond a single site
  - Trademarked

2. Elements of intervention replicability.

The important characteristics of an intervention that must be documented in a study to reliably replicate the intervention with different participants, in other settings, at other times include:

- "Branded"
- Not "branded"
  - Intervention description such as
    - The skill(s) being targeted
    - The approach to enhancing the skill(s)
    - The targeted population
    - The unit of delivery of the intervention (i.e., whole group, small group, or individual student)
    - The medium/media of delivery of the intervention
  - Intervention duration

- The length of time the intervention took place

– Description of intervention deliverers

- Characteristics of the individuals administering the intervention

3. Outcomes relevant to beginning reading.

Measures of alphabetics (phonemic awareness, phonological awareness, letter recognition, print awareness and phonics), reading fluency and comprehension (vocabulary and reading comprehension), or general reading achievement are our primary outcomes of interest**.** In order to keep a focused project scope, we acknowledge the importance of emergent spelling and writing skills and have included these in the review as acceptable skills for pretests, but not as outcomes. Although we recognize the importance of motivation and attitudes toward reading, we have focused this review solely on achievement outcomes.

Reading skills may be measured by standardized achievement tests, by researcher- or teacher-developed materials, by post-intervention class grades, and indirectly, by grade promotion.

The alignment between the outcome and the intervention is another factor we include in our analyses. We would expect that interventions that are aligned or tailored to particular outcomes (e.g., phonics instruction as part of an intervention; phonics skills as a highly aligned outcome) to be more effective than those that are less aligned with one another.

4. Reliability of outcome measures.

Reliability (internal consistency, temporal stability/test-retest reliability, and inter-rater reliability) will be assessed using the following standards determined by the WWC Steering Committee and Technical Advisory Committee:

- Internal consistency: minimum of 0.60
- Temporal stability/test-retest reliability: minimum of 0.40
- Inter-rater reliability: minimum of 0.50

5. Timeframe of review.

The Beginning Reading Evidence Reports focus on a 22-year span, from 1983 to 2005 (a few 2006 studies were added during the report-writing stage, when developers submitted recently completed research). Recall that conference proceedings are limited to the 7-year span of 1998 – 2005. We believe this timeframe adequately represents the current status of the field as well as allows for a manageable project scope.

6. Defining characteristics of the target population.

The beginning reading population is defined as grades K–3 (approximate ages 5 to 8). We include studies on children that are typically developing as well as those at risk for reading difficulties, including children with learning disabilities and low-income or minority children.

7. Characteristics relevant to equating groups.

Important characteristics of participants that might be related to the intervention's effect and must be equated if a study does not employ random assignment include:

- Pretest measures of alphabetics (phonemic awareness, phonological awareness, letter recognition, print awareness and phonics), reading fluency and comprehension (vocabulary and reading comprehension), or general reading achievement.
- Individual characteristics such as IQ and demographic characteristics such as socioeconomic status, and other factors associated with sorting children into study groups

The issue of when the equating was done must also be considered, as well as whether the equating procedure may have resulted in groups with extreme scores in measurements (because upon repeated measurements, these scores tend to move toward the average, even without an intervention taking place).

In QED comparison studies, groups of children being compared must be drawn from the same population of children. Consequently, groups must be roughly equivalent with regard to the pretest of the outcome measure or its proxy (e.g., groups differ on the pretest by less than 1/2 a standard deviation or the difference is not statistically significant in an adequately powered test). Evidence that the groups in a QED comparison group study differ substantially on these dimensions can result in the failure of a study because substantial differences suggest that the groups represent distinct populations. Evidence that the groups come from distinctly different settings, or statistically significant pretest differences, or reported mean pretest differences between groups of more than 1/2 the sample standard deviation suggests that the groups represent different populations. The onus for demonstrating initial equivalence of groups rests with the investigator. Sufficient reporting of these factors should be included (or obtained) to establish the initial equivalence of the groups.

8. Effectiveness of the intervention across different groups.

An intervention's effectiveness will likely vary by subgroups in the population, and a study that claims to test the effectiveness of an intervention should attempt to examine the effects of the intervention within important subgroups. These important subgroups include:

- Students with disabilities

- Students of differing achievement levels (e.g., poor readers, underachievers)
- Students of different ages (i.e., different levels of cognitive development)
- Students from different socioeconomic backgrounds
- Students who are ethnic or racial minorities

9. Effectiveness of the intervention across different settings.

An intervention's effectiveness will likely vary by location, and a study that claims to test the effectiveness of an intervention should attempt to examine the effects of the intervention across different settings. Different settings include:

- Location types (urban, rural, suburban)
- School types (e.g., public, private, parochial)
- Classroom types (e.g., regular, special education, inclusion classrooms)

10. Measuring post-intervention effects.

Most reading studies do not involve much of a lag between the end of an intervention and the measurement of the outcome, and so typical intervals can range from days to several weeks. Few beginning reading studies involve a significant lag of several months or more between the end of the intervention and the measurement of the outcome. Accordingly, we define one (1) day or more post-intervention as an appropriate interval for measuring a reading intervention's effect. In addition, pretest, interim, and posttest measures also boost the ability of the researchers to evaluate the effectiveness of an intervention.

11. Defining differential attrition.

We examine the data presented by study authors and look for evidence that attrition may bias estimates of impacts. For example, if the authors present post-attrition evidence of group equivalence on pretest performance, we may eliminate our concerns about differential attrition. Post-attrition group equivalence on pretest data may be demonstrated by a well-powered (.80) test of equivalence that is non-significant, or a standardized mean difference between groups of less than d = .10.

12. Defining severe overall attrition.

We examine the data presented by study authors and look for evidence that overall attrition may bias estimates of impacts. For example, if the authors are able to present post-attrition evidence of group equivalence on pretest data, we may eliminate our concerns about severe overall attrition. Post-attrition group equivalence on pretest data may be demonstrated by a well-powered (.80) test of equivalence that is non-significant, or a standardized mean difference between groups of less than d = .10.

Note: In certain cases, attrition from the intervention may be an important outcome measure in its own right. Otherwise, only "success stories" would remain to be measured after the intervention, introducing significant bias into estimated effect sizes.

13. Statistical properties important for computing accurate effect sizes.

- For most statistics (including d-indexes), normal distribution and homogeneous variances are important properties.
- For odds-ratios there are no required desirable properties except the minimum of 5 observations per cell.
- In the case where a misaligned analysis is reported (i.e., unit of analysis is not the same as the unit of assignment) and the author is not able to provide a corrected analysis, the effect sizes computed by the WWC will incorporate a statistical adjustment for clustering. The default intraclass correlation used for beginning reading outcomes is 0.20. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch.
- In the case where multiple comparisons are made (i.e., multiple outcome measures are assessed within an outcome domain in one study), the WWC accounts for this multiplicity by adjusting the author reported statistical significance of the effect using the Benjamini-Hochberg correction. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance.

METHODOLOGY

Collecting and Screening Studies

The What Works Clearinghouse (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. The WWC searches core sources and additional topic-specific sources identified by the Principal Investigator and the Senior Content Advisor. The process is fully and publicly documented.

Sources for Studies

Trained WWC staff members use the following strategies in collecting studies:

Databases

This is the core list of electronic databases that are searched across topics:

1. ***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.

2. ***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.

3. ***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.

4. ***Dissertation Abstracts*.** As described by Dialog, Dissertation Abstracts is a definitive subject, title, and author guide to virtually every American dissertation accepted at an accredited institution since 1861. Selected Masters theses have been included since 1962. In addition, since 1988, the database includes citations for dissertations from 50 British universities that have been collected by and filmed at The British Document Supply Center. Beginning with DAIC Volume 49, Number 2 (Spring 1988), citations and abstracts from Section C, Worldwide Dissertations (formerly European Dissertations), have been included in the file. Abstracts are included for doctoral records from July 1980 (Dissertation Abstracts International, Volume 41, Number 1) to the present. Abstracts are included for Master's theses from Spring 1988 (Masters Abstracts, Volume 26, Number 1) to the present.

5. ***Sociological Collection*.** This database provides coverage of more than 500 full-text journals, including nearly 500 peer-reviewed titles. Sociological Collection offers information in all areas of sociology, including social behavior, human tendencies, interaction, relationships, community development, culture, and social structure. This database is updated daily via EBSCOhost.

6. ***Professional Development Collection*.** Designed for professional educators, this database provides a highly specialized collection of over 500 full-text journals, including nearly 350 peer-reviewed titles. Professional Development Collection is the most comprehensive collection of full-text education journals in the world.

7 ***Wilson Education Abstracts PlusText*.** Wilson Education Abstracts PlusText, also known as Education PlusText, combines abstracts and indexing from H.W. Wilson's Education Abstracts database with thousands of full-text and full-image articles. The database includes indexing and abstracts for articles published by more than 400 journals cited in H.W. Wilson's Education Abstracts database. It also includes full-text and full-image coverage for more than 175 of the sources. Overall dates of coverage: 1994–present. Special education, adult education, home schooling, and language and linguistics are just a few of the hundreds of topics users can research in the database.

Search Parameters

After the identification of the topics for review, the Project Coordinator and the librarians initiate

the search using keywords and search terms for each database. The Senior Content Advisor reviews and supplements the list with additional keywords and search terms. Table 1 displays the list of keywords used for the beginning reading electronic searches.

Table 1. Beginning Reading Keywords Used for Electronic Searches

Keywords	ERIC Thesaurus Term(s)	PsycINFO Thesaurus Term(s)	Sociological Collection	Professional Development Collection	Dissertation Abstracts
Literacy	(R) Basic skills, alphabet (letters), literacy education, reading skills	(R) Reading skills, reading education, reading development, literacy programs	(OT) information literacy	(OT) information literacy	Use keywords from Keyword column as needed. There is an Education, reading subject category (Descriptor code: 0535)
Reading skills	(B) Language skills, reading ability	(N) Reading comprehension, reading speed (R) sight vocabulary, work recognition	(SU) Reading comprehension, reading skills competency tests, ability testing	(SU) Reading comprehension, reading skills competency tests, ability testing	Use keywords from Keyword column as needed.
Literacy instruction	(UT) Literacy Education, (N) Basal reading, remedial reading (R) Reading instruction	Literacy programs (R) readings skills, reading education, literacy	(SU) Literacy instruction	(SU) Literacy instruction	Use keywords from Keyword column as needed.
K–3	Kindergarten, (B) Instructional program divisions, (NT) Kindergarten, Grade 1, Grade 2, Grade 3	Primary school students	(SU) Grades K–3, kindergarten, first grade (education), second grade (education), third grade (education)	(SU) Grades K–3, kindergarten, first grade (education), second grade (education), third grade (education)	Use keywords from Keyword column as needed.
Reading comprehension	Reading comprehension, (BT) Reading skills, comprehension (R) Reading strategies, Readability, Reading Rate	Reading comprehension, (B) Reading skills, verbal comprehension (R) Readability, reading	(SU) Reading comprehension	(SU) Reading comprehension	Use keywords from Keyword column as needed.
Vocabulary development	Vocabulary development, (R) Lexicography, verbal development (S) vocabulary building	(UT) Vocalization, (R) Communication, oral communication, verbal communication	(SU) Vocabulary, reading	(SU) Vocabulary, reading	Use keywords from Keyword column as needed.

Keywords	ERIC Thesaurus Term(s)	PsycINFO Thesaurus Term(s)	Sociological Collection	Professional Development Collection	Dissertation Abstracts
Phonemics	Phonemics, (BT) Phonology, (NT) phonemes	Found: Phonetics (B) Phonology (R) Articulation (speech), phonemes	(SU) Reading, spelling ability, phonemics	(SU) Reading, spelling ability, phonemics	Use keywords from Keyword column as needed.
Phonics	Phonics, (BT) Phonetics, (R) Aural learning, word study skills	Phonics (R) Reading education	(SU) Phonics, reading	(SU) Phonics, reading	Use keywords from Keyword column as needed.
Reading fluency (term not in our print copy of ERIC thesaurus - added online 6/20/02)	Reading fluency, (R) Reading comprehension	Found: Reading materials (R) Reading, readability	NOT FOUND (UT) readability	NOT FOUND (UT) readability	Use keywords from Keyword column as needed.
Intervention	Intervention, (N) Early intervention (R) Disabilities, disadvantaged, educational therapy	No specific term	No specific term	No specific term	Use keywords from Keyword column as needed.
Instructional effectiveness	(R) Instructional improvement, program effectiveness, administrator effectiveness, curriculum evaluation educational quality, outcomes of education	Found: Instructional Media	No specific term	No specific term	Use keywords from Keyword column as needed.
Elementary school children	(UT) Elementary school students	Found: Elementary school students (N) Primary school students	No specific term	No specific term	Use keywords from Keyword column as needed.
Assignments	(B) Instruction (R) Homework, reading assignments	No specific term	(SU) assignments	(SU) assignments	Use keywords from Keyword column as needed.
Reading achievement	Reading achievement (R) Reading failure, reading improvement, reading skills, achievement gains	Reading achievement (B) Academic achievement (R) Reading	(SU) Reading, literacy	(SU) Reading, literacy	Use keywords from Keyword column as needed.

Keywords	ERIC Thesaurus Term(s)	PsycINFO Thesaurus Term(s)	Sociological Collection	Professional Development Collection	Dissertation Abstracts
Improvement	(N) Reading improvement, speech improvement (R) Improvement programs, success	No specific term	No specific term	No specific term	Use keywords from Keyword column as needed.
Instructional strategies	(UT) Educational strategies	No specific term	(SU) Instructional strategies	(SU) Instructional strategies	Use keywords from Keyword column as needed.
Educational strategies	Educational strategies, (BT) Educational methods, (R) Instructional design, learning strategies	No specific term	(SU) Educational strategy	(SU) Educational strategy	Use keywords from Keyword column as needed.
Instructional programs (print copy deems “invalid descriptor”)	Use Programs?	No specific term	No specific term	No specific term	Use keywords from Keyword column as needed.
Instructional materials	(N) Courseware, Learning Modules, textbooks, workbooks, protocol materials (R) Reading materials, educational games, educational resources, instructional effectiveness, material development	Found: instructional media (N) Reading materials, textbooks	(SU) Educational tests & measurements, instructional materials industry	(SU) Educational tests & measurements, instructional materials industry	Use keywords from Keyword column as needed.
Schoolwork	(UT) Assignments	No specific term	(SU) Education, schools	(SU) Education, schools	Use keywords from Keyword column as needed.

Journals

The Cochrane Collaboration and the Campbell Collaboration have regarded hand searching of journals as the gold standard in retrieving studies. The yields obtained from hand searches are usually more than from electronic database searches. For a comprehensive review of the literature, each and every article in the journal is examined, even though this is a tedious and time-consuming process. Table 2 displays the list of 29 journals used for the beginning reading hand searches. Six of these journals are topic-specific and emphasize reading outcomes

(* denotes a topic-specific journal).

Table 2. List of Hand-Searched Journals

American Educational Research Journal	Journal of Research and Development in Education
American Journal of Education	**Journal of Special Education***
American Journal of Evaluation	**Learning Disabilities: Research and Practice***
Educational Evaluation and Policy Analysis	**Learning Disability Quarterly***
Educational Research and Evaluation	Peabody Journal of Education
Educational Researcher	Phi Delta Kappan
Effective School Practices	**Reading Research Quarterly***
Evaluation and Program Planning	Review of Research in Education
Evaluation Review	School Effectiveness and School Improvement
Harvard Educational Review	**Scientific Studies of Reading***
Journal of Education	Social Psychology of Education
Journal of Educational Psychology	Sociology of Education
Journal of Educational Research	Teachers College Record
Journal of Experimental Education	Urban Education
Journal of Literacy Research (aka Journal of Reading Behavior)*	

"Fugitive" or "Grey" Literature

Our search for fugitive or grey literature encompassed seven strategies:

1. Review of the 1999-2003 conference programs for the annual meetings of the
    - American Educational Research Association
    - American Psychological Association
    - American Sociological Association
    - National Reading Conference
    - Society for Research on Child Development
    - Society for the Scientific Studies of Reading
2. Public submissions
    - Materials submitted via the WWC website
    - Materials mailed or emailed directly to WWC staff
    - Materials given directly to WWC staff
3. Solicitations made to key researchers by the Evidence Report Team
4. Requests for research made to developers of early literacy-related programs and interventions

- Emailed or called approximately 130 developers to request voluntary participation
- Developers identified through SEDL and NWREL websites; Florida Center for Reading Research (FCRR)

5. Checking prior reviews and research syntheses (i.e., using the reference lists of prior reviews and research syntheses to make sure we have not omitted key studies)
    - NRC Report
    - Other reviews identified through searches
6. Requests for research made via the NRC and SSSR listservs
7. Searches of organizations' websites
    - AIR, CPRE, Mathematica, RAND, Urban Institute
    - AFT, NEA
    - NSF, NICHD
    - Federally funded research centers and programs (NIFL, CIERA)

REFERENCE

Liberman, I., Shankweiler, D., Fischer, F., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, *18*, 201-212.