What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Early Childhood Education May 21, 2007

Direct Instruction, DISTAR, and Language for Learning

Program description

This report focuses on a family of interventions that includes all *Direct Instruction* products (*DISTAR* and *Language for Learning*). We use *Direct Instruction* to refer to this family of interventions, as well as to all versions past and present. *Direct Instruction* includes teaching techniques that are fast-paced, teacher-directed, and explicit with opportunities for student response and teacher reinforcement or correction. The What Works Clearinghouse (WWC) English Language Learners topic also reviewed the effects of a *Direct Instruction* program called *Reading Mastery* (SRA/McGraw-Hill) on the skills of children in elementary school; the findings are reported in the WWC English Language Learners intervention report *Reading Mastery/SRA/McGraw-Hill*.

Research

One study of *Direct Instruction* met the WWC evidence standards with reservations. This study included 164 special education preschool and kindergarten children from a Child Development and Mental Retardation Center at the University of Washington in Washington state.[1] This report focuses on immediate posttest findings to determine the effectiveness of the intervention.[2] The WWC considers the extent of evidence for *Direct Instruction* to be small for oral language, small for print knowledge, small for cognition, and small for math. No studies that met the WWC evidence standards with or without reservations addressed phonological processing or early reading/writing.

Effectiveness

Direct Instruction was found to have no discernible effects on the oral language, print knowledge, cognition, and math skills of special education students.

(continued)

1. The ECE topic includes studies with preschool and kindergarten children when the majority (60% or more) of children in the sample are in preschool. In this study, there were six preschool classes and two kindergarten classes and the mean age of the children was under five, indicating that the study meets this criterion for inclusion. The WWC was unable to obtain the exact sample sizes separated by age; however, the ratio of preschool to kindergarten classrooms meets the criterion described above. For further details on inclusion criteria please see the Early Childhood Education Protocol.
2. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

Effectiveness *(continued)*

	Oral language	Print knowledge	Phonological processing	Early reading/ writing	Cognition	Math
Rating of effectiveness	No discernible effects	No discernible effects	na	na	No discernible effects	No discernible effects
Improvement index[3]	Average: +1 percentile point Range: –7 to +6 percentile points	Average: –3 percentile points	na	na	Average: –1 percentile point	Average: +8 percentile points

na = not applicable

Additional program information

Developer and contact

Developed by Dr. Siegfried Engelmann, Dr. Wesley Becker, and colleagues. Address: The National Institute for Direct Instruction, P.O. Box 11248, Eugene, OR 97440. Email: zig@nifdi.org. Telephone: (877) 485-1973.

Distributed by SRA/McGraw-Hill. Address: 220 East Danieldale Road, Desoto, TX 75115. Email: SRA_CustomerService@mcgraw-hill.com. Web: www.sraonline.com. Telephone: (888) SRA-4543. Computer-based *Direct Instruction* materials are distributed by *Funnix.* Web: www.funnix.com or fun@funnix.com. Telephone (for orders): (866) 732-3963. Telephone (for support): (541) 284-2441.

Organizations that provide training and in-class coaching in the *Direct Instruction* methodology include:

- The National Institute for Direct Instruction
  P.O. Box 11248
  Eugene, OR 97440
  Telephone: (877) 485-1973
  Email: info@nifdi.org
  Web: www.nifdi.org
- Educational Resources, Inc.
  7105 Old Grant Creek Road
  Missoula, MT 59808
  Telephone: (406) 542-5010
  Web: www.erigroup.us
- J/P Associates, Inc.
  284 East Chester Street
  Valley Stream, NY 11580
  Telephone: (516) 561-7803
  Web: www.jponline.com

In addition, the Association for Direct Instruction sponsors workshops and conferences on *Direct Instruction:*

Association for Direct Instruction
P.O. Box 10252
Eugene, OR 97440
Telephone: (541) 485-1293
Web: www.adihome.org

Scope of use

Direct Instruction was developed in the 1960s and has been revised and renamed over the years. The *Direct Instruction* model was first branded as the *Direct Instruction System for Teaching Arithmetic and Reading (DISTAR),* which included *DISTAR Reading, DISTAR Arithmetic,* and *DISTAR Language I* (the programs used in the study reviewed for this WWC report). While *Direct Instruction* products have evolved over the years, the *Direct Instruction* approach remains the same, and the

3. These numbers show the average and range of student-level improvement indices for all findings across the study.

Additional program information *(continued)*

version of *Direct Instruction* (DISTAR) used in the studies in this report are still available and relevant to today's consumers. An overview of *Direct Instruction* programs for preschool and kindergarten children follows. *Direct Instruction* products are also available for older students, but they are not included in this report because they are not relevant to an early childhood population.

- *DISTAR Reading* became *Reading Mastery. Reading Mastery* includes these programs:
  Prekindergarten
  Reading Mastery Classic
  Reading Mastery Plus
  Kindergarten
  Horizons
  Journeys
  Funnix (computer-based)
- *DISTAR Arithmetic* retained that name and includes these programs:
  Prekindergarten (high performing students only)
  DISTAR Arithmetic
  Kindergarten
  Connecting Math Concepts
  DISTAR Arithmetic
- *DISTAR Language I* has become *Language for Learning.* (Student materials only are still available for the original *DISTAR Language I.*) There are currently two companion programs:
  Language for Learning
  Language for Thinking

Direct Instruction materials are currently used by thousands of schools across the nation and internationally; however, specific information is not available on the number or demographics of children or centers using the materials.

Teaching

Direct Instruction is a teaching technique based on extensive task analysis. Instruction is fast-paced, teacher-directed, prescribed, and explicit with all children receiving instruction on a pre-specified sequence of activities at the same time. In center-based settings, teachers typically use *Direct Instruction* interventions with small groups and with typically developing, disadvantaged, or at-risk children. *Funnix* is a computer-based program that is based on the principles of *Direct Instruction* but does not require substantial teacher training and is taught in one-on-one or small group settings with a teacher or tutor operating the computer. Ample opportunity is provided for cued student response, teacher reinforcement, and frequent correction. In all *Direct Instruction* based programs, the school staff takes responsibility for the performance of students, teachers, and instructional subgroups, and individual student progress is closely monitored. *Direct Instruction* is primarily used with K-12 students and adults; however, some *Direct Instruction* based programs have been developed for preschool children. At the time the studies the WWC reviewed were conducted, the programs used for preschool children were *DISTAR Language I, DISTAR Arithmetic, and DISTAR Reading.*

DISTAR Language I focuses on teaching syntactic, semantic, and pragmatic skills. Instruction is teacher-directed, with students responding to teacher instructions, questions, and visual cues and with teachers providing verbal praise for correct responses. *Language for Learning* (the updated version of *DISTAR Language I*) also uses an explicit instructional approach with scripted lessons to teach oral language skills to young children, but has larger print and pictures than *DISTAR Language I. Language for Thinking* (a companion to *Language for Learning*) is designed to teach children the basic language of instruction, concepts, and reasoning. *DISTAR Arithmetic* uses explicit strategies to teach children basic math skills necessary to learning difficult concepts later on (such as ratios, proportions, and data analysis). *DISTAR Reading* uses explicit strategies to teach children the skills they need to learn how to read (skills like blending and rhyming).

All available *Direct Instruction* programs have teacher materials available for purchase and customer service and customer support is available from SRA/McGraw-Hill via phone or email. Training is required to orient teachers and provide practice in executing

Additional program information *(continued)*

the various program techniques. Only *Reading Mastery, Language for Learning,* and *DISTAR Arithmetic* are recommended for at-risk, low-performing preschoolers. For average and above-average preschoolers, the other programs are appropriate.

Cost

DISTAR Language I (1987 version) workbooks I and II each cost $46 for a package of five; workbook III costs $43 for a package of five. Teacher materials are no longer available. *Language for Learning* workbooks A, B, C, and D each cost $31 for a package of five. Teacher materials cost $660 plus $27 for an additional teacher's guide. *Language for Thinking* children's picture books cost $37, and a package of five workbooks costs $60. The teacher materials cost $597 and an additional teacher's guide is available for $27. Additional resources for the programs are also available for purchase.

DISTAR Arithmetic and *Connecting Math Concepts* are both available for kindergarten children. *DISTAR Arithmetic* Level I take-home workbooks I, II, and III each cost $46 for a package of five. Teacher materials cost $600 plus $27 for an additional teacher's guide and $23 for a behavioral objectives book. *DISTAR Arithmetic* 1975 and 1976 materials are also available for purchase. *Connecting Math Concepts* workbooks I and II each cost $59 for a package of five. Teacher materials cost $299; additional teacher guides, answer keys, and other materials are available for $173.

DISTAR Reading is no longer available for purchase; however, *Reading Mastery Plus* and *Reading Mastery Classic* are available for preschool children and *Horizons, Journeys,* and *Funnix* are available for kindergarten children. The student storybook for *Reading Mastery Plus* costs $9, workbook A costs $32 for a package of five, and workbooks B and C each cost $52 for a package of five. Teacher materials for *Reading Mastery Plus* cost $798 plus $27 for an additional teacher's guide. *Reading Mastery Classic* storybooks I, II, and III each cost $17 and workbooks A, B, and C each cost $52 for a package of five. Teacher materials for *Reading Mastery Classic* cost $612 plus $27 for an additional teacher's guide. Other additional resources are also available for both *Reading Mastery* programs. *Funnix* Beginning Reading costs $129 and includes two instructional discs, a tutorial overview disc, and a student workbook.

Horizons literature collection level A (kindergarten) costs $138 plus $30 for an additional literature guide. Level A textbook I costs $29 and textbooks II and III each cost $33. Level A workbooks I, II, and III each cost $52 for a package of five. Level A teacher materials cost $597 plus $27 for an additional teacher's guide. Fast track materials are also available for purchase. *Journeys* kindergarten textbook costs $29 and workbooks I, II, and III each cost $52 for a package of five. Teacher materials cost $798 plus $27 for an additional teacher's guide.

Research

Six studies reviewed by the WWC investigated the effects of *Direct Instruction* in center-based settings. One study (Cole, Dale, Mills, & Jenkins, 1993) was a randomized controlled trial that met WWC evidence standards with reservations.[4] The remaining five studies did not meet WWC evidence screens.

Cole et al. (1993) included 164 three- to seven-year-old children from one experimental school in Washington state. They had special needs and no previous preschool experience. Cole et al. compared oral language, print knowledge, cognition, and math outcomes for children participating in a *Direct Instruction* group with outcomes for children participating in a comparison group that used Mediated Learning. Mediated Learning focuses on developing children's social and cognitive processing skills without emphasis on external reinforcement.

4. The study was classified as "meets evidence standards with reservations" due to severe overall attrition. Based on the number of classes and children in the original study, the sample size at assignment was 368 children with disabilities [Cole et al. (1993) stated that the full sample included just 206 children]. However, the analysis sample was 164 children. Based upon the inconsistency between the figures at assignment, the study was downgraded for severe overall attrition.

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or moderate to large (see the What Works Clearinghouse Extent of Evidence Categorization Scheme). The extent of evidence takes into account the number of studies and the total sample size across the studies that met WWC evidence standards with or without reservations.[5]

The WWC considers the extent of evidence for *Direct Instruction* to be small for oral language, small for print knowledge, small for cognition, and small for math. No studies that met WWC evidence standards with or without reservations addressed phonological processing or early reading/writing.

Effectiveness

Findings

The WWC review of interventions for early childhood education addresses children's outcomes in six domains: oral language, print knowledge, phonological processing, early reading/writing, cognition, and math.[6]

Oral language. Cole et al. (1993) analyzed findings for three measures in this outcome domain [McCarthy Scales of Children's Abilities (MSCA) verbal subtest; Peabody Picture Vocabulary Test-Revised (PPVT-R); Test of Early Language Development (TELD)] and found no statistically significant differences between the intervention and comparison groups for any of the measures.[7] The average effect size across the three oral language measures was neither statistically significant nor large enough to be considered substantively important according to WWC criteria (that is, at least 0.25). In the oral language domain, this study showed no discernible effects according to WWC criteria.

Print knowledge. Cole et al. (1993) analyzed findings for one measure in this outcome domain [Test of Early Reading Ability (TERA)], but did not find a statistically significant difference between the intervention and comparison groups. The effect was not large enough to be considered substantively important according to WWC critera (that is, at least 0.25). In the print knowledge domain, this study showed no discernible effects according to WWC criteria.

Cognition. Cole et al. (1993) analyzed findings for two measures in this outcome domain (MSCA Perceptual subtest; MSCA Memory subtest) and found no statistically significant differences between the intervention and comparison groups for either measure.[8] The average effect size across the two cognitive measures was neither statistically significant nor large enough to be considered substantively important according to WWC criteria (that is, at least 0.25). In the cognition domain, this study showed no discernible effects according to WWC criteria.

Math. Cole et al. (1993) analyzed findings for one measure in this outcome domain (MSCA Quantitative subtest) and did not find a statistically significant difference between the intervention and comparison groups. The effect was not large enough to be considered substantively important according to WWC criteria

5. The Extent of Evidence Categorization was developed to tell readers how much evidence was used to determine the intervention rating, focusing on the number and sizes of studies. Additional factors associated with a related concept, external validity, such as students' demographics and the types of settings in which studies took place, are not taken into account for the categorization.
6. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate the statistical significance. In the case of *Direct Instruction,* no corrections for clustering or multiple comparisons were needed.
7. The authors also analyzed findings for three additional measures, but the WWC does not include them in the overall rating of effectiveness because they were not administered to the full sample of children in the study, and the WWC was unable to obtain the analysis sample sizes for each of these measures.
8. The authors also analyzed findings for a composite measure, but the WWC does not include it in the overall rating of effectiveness because the WWC includes the individual measures used to develop the composite.

Effectiveness *(continued)*

(that is, at least 0.25). In the math domain, this study showed no discernible effects according to WWC criteria.

Rating of effectiveness

The WWC rates the effects of an intervention in a given outcome domain as: positive, potentially positive, mixed, no discernible effects, potentially negative, or negative. The rating of effectiveness takes into account four factors: the quality of the research design, the statistical significance of the findings,[6] the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies (see the WWC Intervention Rating Scheme).

The WWC found *Direct Instruction* to have no discernible effects for oral language, print knowledge, cognition, or math

Improvement index

The WWC computes an improvement index for each individual finding. In addition, within each outcome domain, the WWC computes an average improvement index for each study and an average improvement index across studies (see Technical Details of WWC-Conducted Computations). The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus the percentile rank of the average student in the comparison condition. Unlike the rating of effectiveness, the improvement index is based entirely on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analyses. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.

The average improvement index for oral language is +1 percentile point for the one study, with a range of –7 to +6 percentile points across findings. The improvement index for print knowledge is –3 percentile points for the one outcome in the study. The average improvement index for cognition is –1 percentile point for the one study, with no range across findings because the improvement index for each of the two measures is also –1. The improvement index for math is +8 percentile points for the one outcome in the study.

Summary

The WWC reviewed six studies on *Direct Instruction.* One study met WWC evidence standards with reservations. The remaining studies did not meet WWC evidence screens. Based on the study included in the overall rating of effectiveness, the WWC found no discernible effects for oral language, print knowledge, cognition, or math. The evidence presented in this report is limited and may change as new research emerges.

References

Met WWC evidence standards with reservations

Cole, K. N., Dale, P. S., Mills, P. E., & Jenkins, J. R. (1993). Interaction between early intervention curricula and student characteristics. *Exceptional Children, 60*(1), 17–28.

Additional source:

Cole, K. N., Dale, P. S., & Mills, P. E. (1991). Individual differences in language delayed children's responses to direct and interactive preschool instruction. *Topics in Early Childhood Special Education, 11*(1), 99–124.

Did not meet WWC evidence screens

Cole, K. N., & Dale, P. S. (1986). Direct language instruction and interactive language instruction with language delayed preschool children: A comparison study. *Journal of Speech and Hearing Research, 29*(2), 206–217.[9]

Dale, P. S., & Cole, K. N. (1988). Comparison of academic and cognitive programs for young handicapped children. *Exceptional Children, 54*(5), 439–447.[10]

9. Confound: there was only one teacher in the *Direct Instruction* condition (supported by her two teaching assistants) compared with two instructors (and their teaching assistants) who taught in the Interactive Instruction condition (one instructor taught two of the three classes), so the analysis could not separate the effects of the intervention from the effects of the teacher.
10. Complete data were not reported: the WWC could not compute effect sizes.

References *(continued)*

Additional sources:

Cole, K. N., Mills, P. E., & Dale, P. S. (1989). A comparison of the effects of academic and cognitive curricula for young handicapped children one and two years postprogram. *Topics in Early Childhood Special Education, 9*(3), 110–127.

Cole, K. N., Mills, P. E., Jenkins, J. R., & Dale, P. S. (2005). Early intervention curricula and subsequent adolescent social development: A longitudinal examination. *Journal of Early Intervention, 27*(2), 71–82.

Dale, P. S., Mills, P. E., Cole, K. N., & Jenkins, J. R. (2004). When paths diverge: "Errors of prediction" from preschool test scores to later cognitive and academic measures. *The Journal of Special Education, 37*(4), 237–248.

Dale, P. S., Jenkins, J. R., Mills, P. E., & Cole, K. N. (2005). Follow-up of children from academic and cognitive preschool curricula at 12 and 16. *Exceptional Children, 71*(3), 301–317.

Jenkins, J. R., Dale, P. S., Mills, P. E., Cole, K. N., Pious, C., & Ronk, J. (2006). How special education preschool graduates finish: Status at 19 years of age. *American Educational Research Journal, 43*(4), 737–781.

Mills, P. E., Cole, K. N., Jenkins, J. R., & Dale, P. S. (2002). Early exposure to direct instruction and subsequent juvenile delinquency: A prospective examination. *Exceptional Children, 69*(1), 85–96.

Mills, P. E., Dale, P. S., Cole, K. N., & Jenkins, J. R. (1995). Follow-up of children from academic and cognitive preschool curricula at age 9. *Exceptional Children, 61*(4), 378–393.

Seifert, H., & Schwarz, I. (1991). Treatment effectiveness of large group basic concept instruction with Head Start students. *Language, Speech, & Hearing Services in Schools, 22*(2), 60–64.[11]

Waldron-Soler, K. M., Martella, R. C., Marchand-Martella, N. E., Warner, D. A., Tso, M. E., Warner, D. A., & Miller, D. E. (2002). Effects of a 15-week Language for Learning implementation with children in an integrated preschool. *Journal of Direct Instruction, 2*(2), 75–86.[12]

Weisberg, P. (1988). Direct Instruction in the preschool. *Education & Treatment of Children, 11*(4), 349–363.[12]

For more information about specific studies and WWC calculations, please see the WWC *Direct Instruction, DISTAR, and Language for Learning* Technical Appendices.

11. Confound: the researcher taught the two intervention groups but had no or minimal contact with the children in the comparison groups; therefore, the effects of the intervention cannot be separated from the effects of the teacher.
12. Incomparable groups: the intervention and comparison groups cannot be considered equivalent at baseline, even with the use of covariates in the analysis.

Appendix

Appendix A1 Study characteristics: Cole, Dale, Mills, & Jenkins, 1993 (randomized controlled trial with attrition problems)

Characteristic	Description
Study citation	Cole, K. N., Dale, P. S., Mills, P. E., & Jenkins, J. R. (1993). Interaction between early intervention curricula and student characteristics. *Exceptional Children, 60*(1), 17–28.
Participants	This study included 164 preschool and kindergarten children from a larger study.[1] They had no previous preschool experience. Eighty-one children were in the intervention group and 83 children were in the comparison group.[2] Individual children were first randomly assigned to a *Direct Instruction* intervention group or a Mediated Learning comparison group; then children in each condition were randomly assigned to classrooms (six preschool classes and two kindergarten classes). All the children had disabilities, including language delay (80%), cognitive delay (50%), fine motor delay (60%), gross motor delay (60%), and socio-emotional delay (60%). The mean age of the children in the sample was 4.75 years, their mean IQ was 76.03, and 32% were female. Sixty-two percent of the children were European-American, 29% were African-American, and 10% were Hispanic, Pacific Islander, Asian, Native American, or another race/ethnicity.
Setting	The study took place at the Experimental Education Unit of the University of Washington's Child Development and Mental Retardation Center. Preschool classes were two hours a day, five days a week for 180 days, and kindergarten classes were five and a half hours a day, five days a week for 180 days.
Intervention	The WWC designated the *Direct Instruction* condition as the intervention for this review. The intervention condition included the *Direct Instruction* programs *DISTAR Arithmetic, DISTAR Reading,* and *DISTAR Language,* which are academically-oriented programs characterized by instruction that is fast-paced, teacher-directed, prescribed, and explicit, with all children receiving instruction on a pre-specified sequence of activities at the same time. Although there is information on the length of each school program, there is no information about the duration or frequency of *Direct Instruction.*
Comparison	The WWC designated the Mediated Learning condition as the comparison condition for this review. The Mediated Learning curriculum included units that were two to three weeks long and focused on topics such as identifying patterns, making comparisons, identifying feelings, and planning ahead. Mediated Learning is a cognitively-oriented, non-academic program that emphasizes processes related to input, elaboration, and output. As opposed to the direct approach used by teachers of *Direct Instruction,* teachers using Mediated Learning scaffold children's learning around cognitive processes such as classification and sequencing. Although there is information on the length of each school program, there is no information about the duration or frequency of Mediated Learning.
Primary outcomes and measurement[3]	The primary outcome domains assessed were children's oral language, print knowledge, cognition, and math. Oral language was assessed with five standardized measures and one non-standardized measure. The standardized measures include: the McCarthy Scales of Children's Abilities (MSCA) Verbal Scale; the Peabody Picture Vocabulary Test-Revised (PPVT-R);[4] the Test of Early Language Development (TELD);[4] the Preschool Language Assessment Inventory (PLAI); and the Basic Language Concepts Test-Errors (BLCT; criterion referenced measure). The non-standardized measure is the Mean Length of Utterance (MLU) derived from language samples. Print knowledge was assessed with one standardized measure, the Test of Early Reading Ability (TERA).[4] Cognition was assessed with three standardized measures, the MSCA Composite Scale (General Cognitive Index), the MSCA Perceptual Scale, and the MSCA Memory Scale. Math was assessed with one standardized measure, the MSCA Quantitative Scale (see Appendices A2.1–2.4 for detailed descriptions of outcome measures).[5] Data analyses were based on the children's scores after the first year of participation in the four-year study period.

(continued)

Appendix A1 Study characteristics: Cole, Dale, Mills, & Jenkins, 1993 (randomized controlled trial with attrition problems) *(continued)*

Characteristic	Description
Teacher training	Each class was staffed by a head teacher and an assistant teacher. Other staff included related services personnel and practicum students. All head teachers had Master's degrees in special education. Of the eight *Direct Instruction* head teachers, three received their degrees from a University of Oregon program that emphasized *Direct Instruction,* four received in-service training from that program, and one was trained on-site. Two of the Mediated Learning teachers were trained in Mediated Learning at Vanderbilt University and two additional teachers received training from Dr. Feuerstein (an Israeli psychologist whose work is the basis for Mediated Learning) and consultation on curriculum and procedures. The five additional teachers in the Mediated Learning program received ongoing professional development from the teachers who were trained at Vanderbilt.

1. This sample of children without previous preschool experience is a sub-sample drawn from a related study (Dale & Cole, 1988) that did not meet WWC evidence screens because effect sizes could not be calculated. In consultation with the study authors, the sub-sample used in Cole et al. (1993) was determined to be an acceptable study with which to determine the effectiveness of *Direct Instruction* in comparison with Mediated Learning. The ECE topic includes studies with preschool and kindergarten children when the majority (60% or more) of children in the sample are in preschool. In this study, there were six preschool classes and two kindergarten classes per year, and the mean age of the children was under five, indicating that the study meets this criterion for inclusion.
2. The study was classified as "meets evidence standards with reservations" due to severe overall attrition. Based on the number of classes and children in the original study, the sample size at assignment was 368 children with disabilities [Cole et al. (1993) stated that the full sample included just 206 children]. However, the analysis sample was 164 children. Based upon the inconsistency between the figures at assignment, the study was downgraded for severe overall attrition.
3. The BLCT, MLU, and PLAI are not included in the overall rating of effectiveness because they were not administered to the full sample of children in the study and cannot be used to determine the overall effectiveness of the intervention. Additionally, they are not included in the later appendices for this study because the WWC was unable to obtain the analysis sample sizes for each of these measures. The WWC did not include the MSCA Composite Scale in the WWC intervention report because the WWC includes the individual measures used to develop the composite. There was one additional outcome measure, the MSCA Motor Scale; however, this measure is not included in this report because it is not relevant to the WWC review.
4. The authors reported raw scores and standardized scores for the PPVT-R, the TELD, and the TERA. The WWC report includes the standardized scores only because they are adjusted for age.
5. For further details about the outcomes included in the early childhood education topic review, please see the Early Childhood Education Protocol.

Appendix A2.1 Outcome measures in the oral language domain

Characteristic	Description
McCarthy Scales of Children's Abilities (MSCA) Verbal subtest	A subtest from a standardized measure that assesses children's receptive and expressive language (as cited in Cole et al., 1993, and Cole et al., 1991).
Peabody Picture Vocabulary Test-Revised (PPVT-R)	A standardized measure of children's receptive vocabulary that requires them to identify pictures that correspond to spoken words (as cited in Cole et al., 1993, and Cole et al., 1991).
Test of Early Language Development (TELD)	A standardized measure of children's semantic and syntactic skills in receptive and expressive language (as cited in Cole et al., 1993, and Cole et al., 1991).
Basic Language Concepts Test (BLCT)	A criterion-referenced measure of children's receptive and expressive language that focuses on common nouns, adjectives, plural, and tense. This test was developed as part of the *Direct Instruction* program and measures the number of errors (as cited in Cole et al., 1993, and Cole et al., 1991) and a lower score reflects a better outcome.
Preschool Language Assessment Inventory (PLAI)	A standardized measure that assesses children's ability to respond to increasingly abstract language (as cited in Cole et al., 1993, and Cole et al., 1991).
Mean Length of Utterance (MLU)	A non-standardized measure of children's expressive language based on 20-minute language samples (as cited in Cole et al., 1993, and Cole et al., 1991).

Appendix A2.2 Outcome measure in the print knowledge domain

Characteristic	Description
Test of Early Reading Ability (TERA)[1]	A standardized measure of young children's early reading skills that captures the following constructs: awareness of print in environmental contexts, vocabulary, listening, comprehension, knowledge of alphabet, and concepts about printed language (as cited in Cole et al., 1993).

1. By name, this measure sounds like it should be captured under the early reading/writing domain; however, the description of the measure identifies constructs that are pertinent to print knowledge such as knowing the alphabet, understanding print conventions, and environmental print.

Appendix A2.3 Outcome measures in the cognition domain

Characteristic	Description
MSCA General Cognitive Index[1]	A standardized measure of children's general intellectual functioning that is based on their scores of the Perceptual, Verbal, and Quantitative measures (as cited in Cole et al., 1993 and Cole et al., 1991).
MSCA Perceptual Scale subtest	A subtest from a standardized measure that assesses children's ability to conceptualize and reason (as cited in Cole et al., 1993, and Cole et al., 1991).
MSCA Memory Scale subtest	A subtest from a standardized measure that assesses children's short-term recall of words, pictures, numbers, and sound sequences (as cited in Cole et al., 1993, and Cole et al., 1991).

1. The WWC did not include the MSCA General Cognitive Index in the overall rating of effectiveness. However, the results for this measure are reported in Appendix A4.2.

Appendix A2.4 Outcome measure in the math domain

Characteristic	Description
MSCA Quantitative subtest	A subtest from a standardized measure that assesses children's mathematical ability (as cited in Cole et al., 1993).

Appendix A3.1 Summary of study findings included in the rating for the oral language domain[1]

Outcome measure	Study sample	Sample size (children)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *Direct Instruction* group[3]	Authors' findings from the study: Mean outcome (standard deviation[2]) — Mediated Learning group[3]	WWC calculations: Mean difference[4] (*Direct Instruction* – Mediated Learning)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Cole et al., 1993 (randomized controlled trial with attrition problems)[8]								
MSCA Verbal	3–7 year olds	164	39.20[9] (10.80)	37.70 (9.80)	1.50	0.14	ns	+6
PPVT-R Scale Score	3–7 year olds	164	81.70 (16.20)	84.60 (14.70)	–2.90	–0.19	ns	–7
TELD Language Quotient	3–7 year olds	164	89.50 (13.70)	88.20 (11.80)	1.30	0.10	ns	+4
Domain average[10] for oral language						0.02	ns	+1

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices. Composite scores and subgroup findings from the same study are not included in these ratings, but are reported in Appendices A4.1 and A4.2.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *Direct Instruction* group mean equals the Mediated Learning group mean plus the mean difference.
4. Positive differences and effect sizes favor the *Direct Instruction* group; negative differences and effect sizes favor the Mediated Learning group. The mean differences were computed by the WWC and took into account the pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effect when the *Direct Instruction* group had lower pretest scores than the Mediated Learning group and underestimate the intervention's effect when the *Direct Instruction* group had higher pretest scores than the Mediated Learning group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the *Direct Instruction* condition versus the percentile rank of the average student in the Mediated Learning condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the *Direct Instruction* group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Cole et al. (1993), no corrections for clustering or multiple comparisons were needed.
9. For all means and standard deviations in Appendix A3.1 the study authors reported data to a single decimal place. For formatting purposes, the WWC added zero in the second decimal place.
10. This row provides the study average, which, in this instance, is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A3.2 Summary of study findings included in the rating for the print knowledge domain[1]

Outcome measure	Study sample	Sample size (children)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *Direct Instruction* group[3]	Authors' findings from the study: Mean outcome (standard deviation[2]) — Mediated Learning group[3]	WWC calculations: Mean difference[4] (*Direct Instruction* – Mediated Learning)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Cole et al., 1993 (randomized controlled trial with attrition problems)[8]								
TERA Reading Quotient	3–7 year olds	164	78.00[9] (13.80)	79.10 (11.10)	–1.10	–0.09	ns	–3
Domain average[10] for print knowledge						–0.09	ns	–3

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices. Composite scores and subgroup findings from the same study are not included in these ratings, but are reported in Appendices A4.1 and A4.2.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *Direct Instruction* group mean equals the Mediated Learning group mean plus the mean difference.
4. Positive differences and effect sizes favor the *Direct Instruction* group; negative differences and effect sizes favor the Mediated Learning group. The mean differences were computed by the WWC and took into account the pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effect when the *Direct Instruction* group had lower pretest scores than the Mediated Learning group and underestimate the intervention's effect when the *Direct Instruction* group had higher pretest scores than the Mediated Learning group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the *Direct Instruction* condition versus the percentile rank of the average student in the Mediated Learning condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the *Direct Instruction* group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Cole et al. (1993), no corrections for clustering or multiple comparisons were needed.
9. For all means and standard deviations in Appendix A3.2 the study authors reported data to a single decimal place. For formatting purposes, the WWC added zero in the second decimal place.
10. This row provides the study average, which, in this instance, is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A3.3 Summary of study findings included in the rating for the cognition domain[1]

Outcome measure	Study sample	Sample size (children)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *Direct Instruction* group[3]	Authors' findings from the study: Mean outcome (standard deviation[2]) — Mediated Learning group[3]	WWC calculations: Mean difference[4] (*Direct Instruction* – Mediated Learning)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Cole et al., 1993 (randomized controlled trial with attrition problems)[8]								
MSCA Perceptual	3–7 year olds	164	38.20[9] (10.80)	38.50 (10.80)	–0.30	–0.03	ns	–1
MSCA Memory	3–7 year olds	164	36.90 (10.70)	37.30 (10.70)	–0.40	–0.04	ns	–1
Domain average[10] for cognition						–0.03	ns	–1

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices. Composite scores and subgroup findings from the same study are not included in these ratings, but are reported in Appendices A4.1 and A4.2.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *Direct Instruction* group mean equals the Mediated Learning group mean plus the mean difference.
4. Positive differences and effect sizes favor the *Direct Instruction* group; negative differences and effect sizes favor the Mediated Learning group. The mean differences were computed by the WWC and took into account the pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effect when the *Direct Instruction* group had lower pretest scores than the Mediated Learning group and underestimate the intervention's effect when the *Direct Instruction* group had higher pretest scores than the Mediated Learning group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the *Direct Instruction* condition versus the percentile rank of the average student in the Mediated Learning condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the *Direct Instruction* group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Cole et al. (1993), no corrections for clustering or multiple comparisons were needed.
9. For all means and standard deviations in Appendix A3.3 the study authors reported data to a single decimal place. For formatting purposes, the WWC added zero in the second decimal place.
10. This row provides the study average, which, in this instance, is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A3.4 Summary of study findings included in the rating for the math domain[1]

Outcome measure	Study sample	Sample size (children)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *Direct Instruction* group[3]	Authors' findings from the study: Mean outcome (standard deviation[2]) — Mediated Learning group[3]	WWC calculations: Mean difference[4] (*Direct Instruction* – Mediated Learning)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Cole et al., 1993 (randomized controlled trial with attrition problems)[8]								
MSCA Quantitative	3–7 year olds	164	38.70[9] (9.70)	39.60 (9.00)	1.80	0.19	ns	+8
Domain average[10] for math						0.19	ns	+8

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices. Composite scores and subgroup findings from the same study are not included in these ratings, but are reported in Appendices A4.1 and A4.2.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *Direct Instruction* group mean equals the Mediated Learning group mean plus the mean difference.
4. Positive differences and effect sizes favor the *Direct Instruction* group; negative differences and effect sizes favor the Mediated Learning group. The mean differences were computed by the WWC and took into account the pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effect when the *Direct Instruction* group had lower pretest scores than the Mediated Learning group and underestimate the intervention's effect when the *Direct Instruction* group had higher pretest scores than the Mediated Learning group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the *Direct Instruction* condition versus the percentile rank of the average student in the Mediated Learning condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the *Direct Instruction* group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Cole et al. (1993), no corrections for clustering or multiple comparisons were needed.
9. For all means and standard deviations in Appendix A3.4 the study authors reported data to a single decimal place. For formatting purposes, the WWC added zero in the second decimal place.
10. This row provides the study average, which, in this instance, is also the domain average. The WWC-computed domain average effect size is a simple average rounded to two decimal places. The domain improvement index is calculated from the average effect size.

Appendix A4.1 Summary of subgroup findings for the oral language domain[1]

Outcome measure	Study sample	Sample size (children)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *Direct Instruction* group[3]	Authors' findings from the study: Mean outcome (standard deviation[2]) — Mediated Learning group[3]	WWC calculations: Mean difference[4] (*Direct Instruction* – Mediated Learning)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Cole et al., 1991 (randomized controlled trial with attrition problems; children with language delays subgroup)[8]								
BLCT	3–7 year olds	58	38.40[9] (15.10)	37.80 (17.40)	–0.60	–0.04	ns	–1
PLAI	3–7 year olds	52	25.10 (11.30)	25.30 (9.00)	–0.20	–0.02	ns	–1
MLU	3–7 year olds	59	3.80 (1.10)	3.70 (1.10)	0.10	0.09	ns	+4
PPVT-R Standard Score	3–7 year olds	103	79.80 (14.40)	84.00 (12.90)	–4.20	–0.30	ns	–12
TELD Quotient	3–7 year olds	92	84.00 (10.10)	84.90 (10.20)	–0.90	–0.09	ns	–4

ns = not statistically significant

1. This appendix presents subgroup findings for measures that fall in the oral language domain for a related study. Cole et al. (1991) is not a pure subgroup analysis of Cole et al. (1993) (both studies are subsamples of a study that did not meet WWC evidence screens); however, Cole et al. (1991) compares the effects of *Direct Instruction* with the effects of Mediated Learning for children with language delays using a related subsample.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *Direct Instruction* group mean equals the Mediated Learning group mean plus the mean difference.
4. Positive differences and effect sizes favor the *Direct Instruction* group and negative differences and effect sizes favor the Mediated Learning group. The mean differences were computed by the WWC and took into account the pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effect when the *Direct Instruction* group had lower pretest scores than the Mediated Learning group and underestimate the intervention's effect when the *Direct Instruction* group had higher pretest scores than the Mediated Learning group. The BLCT is a measure of the number of errors, but the mean difference is calculated so that a positive effect indicates that the *Direct Instruction* group performed better (had fewer errors) than the Mediated Learning group.
5. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the *Direct Instruction* condition versus the percentile rank of the average student in the Mediated Learning condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the *Direct Instruction* group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Cole et al. (1991; 1993), no corrections for clustering were needed.
9. For all means and standard deviations in Appendix A4.1 the study authors reported data to a single decimal place. For formatting purposes, the WWC added zero in the second decimal place.

Appendix A4.2 Summary of subgroup findings for the cognition domain[1]

Outcome measure	Study sample	Sample size (children)	Authors' findings from the study: Mean outcome (standard deviation[2]) — *Direct Instruction* group[3]	Authors' findings from the study: Mean outcome (standard deviation[2]) — Mediated Learning group[3]	WWC calculations: Mean difference[4] (*Direct Instruction* – Mediated Learning)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at $\alpha = 0.05$)	WWC calculations: Improvement index[7]
Cole et al., 1993 (randomized controlled trial with attrition problems)[8]								
MSCA GCI	3–7 year olds	164	77.90[9] (18.20)	76.60 (16.90)	1.30	0.07	ns	+3
Cole et al., 1991 (randomized controlled trial with attrition problems; children with language delays subgroup)[8]								
MSCA GCI	3–7 year olds	105	77.00 (14.00)	78.90 (14.40)	–1.90	–0.13	ns	–5

ns = not statistically significant

1. This appendix presents composite subgroup findings for measures that fall in the cognition domain. Total group scores were used for rating purposes and are presented in Appendix A3.3. As noted in Appendix A2.3, the WWC does not include the MSCA GCI in the overall rating of effectiveness because the WWC includes the individual measures used to develop the composite. Cole et al. (1991) is not a pure subgroup analysis of Cole et al. (1993) (both studies are subsamples of a study that did not meet WWC evidence screens); however, Cole et al. (1991) compares the effects of *Direct Instruction* with the effects of Mediated Learning for children with language delays.
2. The standard deviation across all students in each group shows how dispersed the participants' outcomes are; a smaller standard deviation on a given measure would indicate that participants had more similar outcomes.
3. The *Direct Instruction* group mean equals the Mediated Learning group mean plus the mean difference.
4. Positive differences and effect sizes favor the *Direct Instruction* group; negative differences and effect sizes favor the Mediated Learning group. The mean differences were computed by the WWC and took into account the pretest difference between the study groups. The resulting effect sizes may overestimate the intervention's effect when the *Direct Instruction* group had lower pretest scores than the Mediated Learning group and underestimate the intervention's effect when the *Direct Instruction* group had higher pretest scores than the Mediated Learning group.
5. For an explanation of effect size calculation, see Technical Details of WWC-Conducted Computations.
6. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
7. The improvement index represents the difference between the percentile rank of the average student in the *Direct Instruction* condition versus the percentile rank of the average student in the Mediated Learning condition. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the *Direct Instruction* group.
8. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools (corrections for multiple comparisons were not done for findings not included in the overall intervention rating). For an explanation about the clustering correction, see the WWC Tutorial on Mismatch. See Technical Details of WWC-Conducted Computations for the formulas the WWC used to calculate statistical significance. In the case of Cole et al. (1991; 1993), no corrections for clustering were needed.
9. For all means and standard deviations in Appendix A4.2 the study authors reported data to a single decimal place. For formatting purposes, the WWC added zero in the second decimal place.

Appendix A5.1 *Direct Instruction* rating for the oral language domain

The WWC rates an intervention's effects in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of oral language, the WWC rated *Direct Instruction* as having no discernible effects. It did not meet the criteria for positive effects, potentially positive effects, mixed effects, potentially negative effects, or negative effects, as no studies showed statistically significant or substantively important effects, either positive or negative.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative.*

  Met. The study did not show statistically significant or substantively important effects, either positive or negative.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. Only one study examined effects on oral language.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

  Met. The study did not show statistically significant or substantively important negative effects.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

  Not met. The one study did not show statistically significant or substantively important positive effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

  Not met. The single study showed indeterminate effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

  Not met. The study did not show statistically significant or substantively important effects, either positive or negative.

OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

  Not met. The single study showed indeterminate effects.

(continued)

Appendix A5.1 *Direct Instruction* rating for the oral language domain *(continued)*

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

  Not met. The study did not show statistically significant or substantively important negative effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

  Met. The study did not show statistically significant or substantively important positive effects.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. Only one study examined effects on oral language.

AND

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

  Met. The study did not show statistically significant or substantively important positive effects.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A5.2 *Direct Instruction* rating for the print knowledge domain

The WWC rates an intervention's effects in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of print knowledge, the WWC rated *Direct Instruction* as having no discernible effects. It did not meet the criteria for positive effects, potentially positive effects, mixed effects, potentially negative effects, or negative effects, as no studies showed statistically significant or substantively important effects, either positive or negative.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative.*

  Met. The study did not show statistically significant or substantively important effects, either positive or negative.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. Only one study examined effects on print knowledge.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

  Met. The study did not show statistically significant or substantively important negative effects.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

  Not met. The one study did not show statistically significant or substantively important positive effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

  Not met. The single study showed indeterminate effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

  Not met. The study did not show statistically significant or substantively important effects, either positive or negative.

OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

  Not met. The single study showed indeterminate effects.

(continued)

Appendix A5.2 *Direct Instruction* rating for the print knowledge domain *(continued)*

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

  Not met. The study did not show statistically significant or substantively important negative effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

  Met. The study did not show statistically significant or substantively important positive effects.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. Only one study examined effects on print knowledge.

AND

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

  Met. The study did not show statistically significant or substantively important positive effects.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A5.3 *Direct Instruction* rating for the cognition domain

The WWC rates an intervention's effects in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of cognition, the WWC rated *Direct Instruction* as having no discernible effects. It did not meet the criteria for positive effects, potentially positive effects, mixed effects, potentially negative effects, or negative effects, as no studies showed statistically significant or substantively important effects, either positive or negative.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative.*

  Met. The study did not show statistically significant or substantively important effects, either positive or negative.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. Only one study examined effects on cognition.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

  Met. The study did not show statistically significant or substantively important negative effects.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

  Not met. The one study did not show statistically significant or substantively important positive effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

  Not met. The single study showed indeterminate effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

  Not met. The study did not show statistically significant or substantively important effects, either positive or negative.

OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

  Not met. The single study showed indeterminate effects.

(continued)

Appendix A5.3 *Direct Instruction* rating for the cognition domain *(continued)*

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

  Not met. The study did not show statistically significant or substantively important negative effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

  Met. The study did not show statistically significant or substantively important positive effects.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. Only one study examined effects on cognition.

AND

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

  Met. The study did not show statistically significant or substantively important positive effects.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A5.4 *Direct Instruction* rating for the math domain

The WWC rates an intervention's effects in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative.[1]

For the outcome domain of math, the WWC rated *Direct Instruction* as having no discernible effects. It did not meet the criteria for positive effects, potentially positive effects, mixed effects, potentially negative effects, or negative effects, as no studies showed statistically significant or substantively important effects, either positive or negative.

Rating received

No discernible effects: No affirmative evidence of effects.

- Criterion 1: None of the studies shows a statistically significant or substantively important effect, either *positive* or *negative.*

  Met. The study did not show statistically significant or substantively important effects, either positive or negative.

Other ratings considered

Positive effects: Strong evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *positive* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. Only one study examined effects on math.

AND

- Criterion 2: No studies showing statistically significant or substantively important *negative* effects.

  Met. The study did not show statistically significant or substantively important negative effects.

Potentially positive effects: Evidence of a positive effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect.

  Not met. The one study did not show statistically significant or substantively important positive effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *negative* effect and fewer or the same number of studies showing *indeterminate* effects than showing statistically significant or substantively important *positive* effects.

  Not met. The single study showed indeterminate effects.

Mixed effects: Evidence of inconsistent effects as demonstrated through either of the following criteria.

- Criterion 1: At least one study showing a statistically significant or substantively important *positive* effect, and at least one study showing a statistically significant or substantively important *negative* effect, but no more such studies than the number showing a statistically significant or substantively important *positive* effect.

  Not met. The study did not show statistically significant or substantively important effects, either positive or negative.

OR

- Criterion 2: At least one study showing a statistically significant or substantively important effect, and more studies showing an *indeterminate* effect than showing a statistically significant or substantively important effect.

  Not met. The single study showed indeterminate effects.

(continued)

Appendix A5.4 *Direct Instruction* rating for the math domain *(continued)*

Potentially negative effects: Evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: At least one study showing a statistically significant or substantively important *negative* effect.

  Not met. The study did not show statistically significant or substantively important negative effects.

AND

- Criterion 2: No studies showing a statistically significant or substantively important *positive* effect, or more studies showing statistically significant or substantively important *negative* effects than showing statistically significant or substantively important *positive* effects.

  Met. The study did not show statistically significant or substantively important positive effects.

Negative effects: Strong evidence of a negative effect with no overriding contrary evidence.

- Criterion 1: Two or more studies showing statistically significant *negative* effects, at least one of which met WWC evidence standards for a strong design.

  Not met. Only one study examined effects on math.

AND

- Criterion 2: No studies showing statistically significant or substantively important *positive* effects.

  Met. The study did not show statistically significant or substantively important positive effects.

1. For rating purposes, the WWC considers the statistical significance of individual outcomes and the domain-level effect. The WWC also considers the size of the domain-level effect for ratings of potentially positive or potentially negative effects. See the WWC Intervention Rating Scheme for a complete description.

Appendix A6 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Centers	Children	
Oral language	1	1	164	Small
Print knowledge	1	1	164	Small
Phonological processing	0	0	0	na
Early reading/writing	0	0	0	na
Cognition	1	1	164	Small
Math	1	1	164	Small

na = not applicable/not studied

1. A rating of "moderate to large" requires at least two studies and two schools across studies in one domain and a total sample size across studies of at least 350 students or 14 classrooms. Otherwise, the rating is "small."