What Works Clearinghouse

INSTITUTE OF EDUCATION SCIENCES

Early Childhood Education December 14, 2006

Phonological Awareness Training

Practice description

Phonological Awareness Training is a general practice aimed at enhancing young children's phonological awareness abilities. Phonological awareness refers to the ability to detect or manipulate the sounds in words independent of meaning. Phonological awareness is a precursor to reading. *Phonological Awareness Training* can involve various training activities that focus on teaching children to identify, detect, delete, segment, or blend segments of spoken words (i.e., words, syllables, onsets and rimes, phonemes) or that focus on teaching children to detect, identify, or produce rhyme or alliteration. Three related What Works Clearinghouse (WWC) intervention reports review two curricula for phonological awareness—*DaisyQuest* and *Sound Foundations*—and a similar practice—*Phonological Awareness Training plus Letter Knowledge Training.*

Research

Four studies of *Phonological Awareness Training* met the WWC evidence standards and two studies met the WWC evidence standards with reservations.[1] Together, these six studies included more than 100 preschool children from Washington State and the Pacific Northwest and examined intervention effects on children's phonological processing. Most of the children studied were from economically disadvantaged families, and about one-fourth of the children had developmental delays. This report focuses on immediate posttest findings to determine the effectiveness of the intervention.[2]

Effectiveness

Phonological Awareness Training was found to have positive effects on phonological processing.

	Oral language	*Print knowledge*	*Phonological processing*	*Early reading/ writing*	*Cognition*	*Math*
Rating of effectiveness	N/A	N/A	Positive effects	N/A	N/A	N/A

(continued)

1. To be eligible for the WWC's review, the Early Childhood Education (ECE) interventions had to be implemented in English in center-based settings with children ages 3 to 5 or in preschool. Two additional studies are not included in the overall effectiveness ratings because they compared variations of *Phonological Awareness Training* interventions to each other, which does not allow the effects of *Phonological Awareness Training* to be determined. See the section titled "Findings for comparisons between variations of *Phonological Awareness Training*" and Appendices A4.1–A4.3 for findings from these two studies.
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
Improvement index[3]	N/A	N/A	Average: +27 percentile points Range: –27 to +50 percentile points	N/A	N/A	N/A

Additional practice information

Developer and contact

Phonological Awareness Training does not have a single developer responsible for providing information or materials. The interventions described in this report were developed by the study authors and are not available for distribution through a common developer. However, readers interested in using *Phonological Awareness Training* practices in their classroom can refer to sources available through internet searches for information. A list of examples follows which has not been reviewed or endorsed by the WWC:

- Phonological Awareness: Instructional and Assessment Guidelines: http://www.ldonline.org/article/6254.
- Ideas and Activities for Developing Phonological Awareness Skills: A Teacher Resource Supplement to the Virginia Early Intervention Reading Initiative: http://www.pen.k12.va.us/VDOE/Instruction/Reading/findings.pdf.
- Reading Rockets: Teacher Toolbox—Phonological Awareness: The Phive Phones of Reading: http://www.readingrockets.org/firstyear/fyt.php?SUB=33.
- Reading Rockets: Problems Involving Phonological and Phonemic Awareness: http://www.readingrockets.org/helping/ target/phonologicalphonemic.
- Phonological Awareness Skills and Spelling Skills: http://cla.calpoly.edu/~jrubba/phon/phonaware.html.
- Vaughn Gross Center for Reading and Language Arts, University of Texas at Austin: http://www.texasreading.org/utcrla/materials/primary_phono_awareness.asp.
- Phonological Awareness and Reading Recovery: http://www.readingrecovery.org/sections/reading/phonics.asp.
- Improving Reading Fluency: Phonological Awareness Training: http://www.speechpathology.com/Articles/article_detail.asp?article_id=68.
- Florida Center for Reading Research: http://www.fcrr.org.
- University of Oregon: http://www.reading.uoregon.edu.
- National Reading Panel: http://www.nationalreadingpanel.org.

Scope of use

Published research studies examining *Phonological Awareness Training* practices began appearing in the mid- to late 1970s and early 1980s. Information is not available on the number or demographics of children or centers using these practices.

Teaching

Phonological Awareness Training practices can be used by teachers with individual children, in pairs, or in small group settings. These practices generally are used as a supplement to the regular classroom curriculum, and they have been used with specific subpopulations of students, such as students with learning disabilities. *Phonological Awareness Training* practices vary in their scope and may include such activities as rhyme detection training (e.g., teachers may engage children in a game involving rhyming words and questions about which word in a series of three does not sound like the others), blending training (e.g., teachers may say three sounds and teach children how to

3. These numbers show the average and range of improvement indices for all findings across the studies.

Additional practice information *(continued)*

blend the sounds together to make a word), and segmentation training (e.g., teachers may say a short word such as "cat" and teach children how to separate the word into the three sounds that make up the word) at the phoneme, syllable, or word level.[4]

Cost

Information is not available about the costs of teacher training and implementation of *Phonological Awareness Training* practices.

Research

Eleven studies reviewed by the WWC investigated the effects of *Phonological Awareness Training* practices in center-based settings.[5] Four of the studies (Majsterek, Shorr, & Erion, 2000; O'Connor, Jenkins, Leicester, & Slocum, 1993a, b, c[6]) were randomized controlled trials that met WWC evidence standards. Two studies (Slocum, O'Connor, & Jenkins, 1993a, b[7]) were randomized controlled trials that met WWC evidence standards with reservations because of high overall attrition.[8] Two additional studies met standards (Yeh, 2003) and met standards with reservations (Maslanka & Joseph, 2002) and are included in this report; however, they compare different variations of *Phonological Awareness Training* to each other, which does not allow the effects of *Phonological Awareness Training* to be determined. Therefore, these studies are discussed separately, and the findings are not included in the intervention ratings. The three remaining studies did not meet WWC evidence screens.

Met evidence standards

Majsterek et al. (2000) included 40 three- to five-year-old low-income preschool children attending a Head Start program in Washington State. Forty-five percent of the children were female. Majsterek et al. compared a phonological processing outcome for an intervention group that participated in rhyme detection training with a comparison group that participated in semantic training (i.e., training that focused on word meanings).

O'Connor et al. (1993a) included 22 four- to six-year-old developmentally delayed children attending the Experimental Education Unit at the University of Washington. Eighty percent of the children had significant language delays. O'Connor et al. compared phonological processing outcomes for an intervention group that was taught blending skills with outcomes for children in a no-treatment comparison group that participated in regular preschool activities.

4. Readers who are unfamiliar with the terminology related to *Phonological Awareness Training* and the development of reading may find it helpful to consult the glossary of terms available from the National Institute for Literacy (http://www.nifl.gov/partnershipforreading/glossary/glossary.html) and the definitions of outcome measures in Appendices A2.1–A2.3.

5. Three single-case design studies were identified but are not included in this review because the WWC does not yet have standards for reviewing single-case design studies.

6. O'Connor et al. (1993) is counted as three studies (O'Connor et al., 1993a, b, c) because it examined three different but relevant *Phonological Awareness Training* interventions. The WWC designated blending versus comparison as O'Connor et al. (1993a), segmenting versus comparison as O'Connor et al. (1993b), and rhyming versus comparison as O'Connor et al. (1993c). The same comparison group was used in all three studies.

7. Slocum et al. (1993) is counted as two studies (Slocum et al., 1993a, b) because it examined two different but relevant *Phonological Awareness Training* interventions. The WWC designated the blend-then-segment versus word manipulation-then-segment condition as Slocum et al. (1993a) and the segment-then-blend versus word manipulation-then-blend condition as Slocum et al. (1993b).

8. A teacher strike interrupted the intervention and may have contributed to the high rates of attrition.

Research *(continued)*

O'Connor et al. (1993b) included 24 four- to six-year-old developmentally delayed children attending the Experimental Education Unit at the University of Washington. Eighty percent of the children had significant language delays. O'Connor et al. compared phonological processing outcomes for an intervention group that was taught segmenting skills with outcomes for children in a no-treatment comparison group that participated in regular preschool activities.

O'Connor et al. (1993c) included 23 four- to six-year-old developmentally delayed children attending the Experimental Education Unit at the University of Washington. Eighty percent of the children had significant language delays. O'Connor et al. compared phonological processing outcomes for an intervention group that was taught rhyming skills with outcomes for children in a no-treatment comparison group that participated in regular preschool activities.

Met evidence standards with reservations

Slocum et al. (1993a) included 16 low-achieving preschool children attending four Head Start classrooms in an urban area of the Pacific Northwest. The study is part of a larger two-phase cross-over design study (see Slocum et al., 1993b, below) that included two treatment groups and two comparison groups. For the purposes of this review, the WWC includes only the results from the phase one instructional period, during which the intervention involved blending, because the groups switched conditions in the phase two instructional period, which does not allow the effects of the prior condition to be separated from the effects of the later condition. In this study, phonological processing outcomes of children in the intervention group were compared with those of children in the corresponding comparison group who participated in word manipulation activities during phase one.

Slocum et al. (1993b) included 19 low-achieving preschool children attending four Head Start classrooms in an urban area of the Pacific Northwest. The study is part of a larger two-phase cross-over design study (see Slocum et al., 1993a, above) that included two treatment groups and two comparison groups. For the purposes of this review, the WWC includes only the results from the phase one instructional period, during which the intervention involved segmenting, because the groups switched conditions in the phase two instructional period, which does not allow the effects of the prior condition to be separated from the effects of the later condition. In this study, phonological processing outcomes of children in the intervention group were compared with those of children in the corresponding comparison group who participated in word manipulation activities during phase one.

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.[9]

Phonological processing. Six studies examined outcomes in the domain of phonological processing. Four studies showed statistically significant and positive effects and two studies showed substantively important and positive effects.

Majsterek et al. (2000) reported findings for one measure in this outcome domain, but no statistically significant difference

9. 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 the statistical significance. In the case of *Phonological Awareness Training,* corrections for clustering and multiple comparisons were needed.

Effectiveness *(continued)*

was found (as calculated by the WWC).[10] The findings favored the intervention group, however, and were large enough to categorize the effect as substantively important and positive, according to WWC criteria.

O'Connor et al. (1993a) compared the blending and no-treatment comparison groups on nine measures in this outcome domain.[11] They reported statistically significant differences favoring the blending group for the three blending measures and one nonstatistically significant effect for a rhyming measure. They did not report statistical significance for the other five measures. The WWC confirmed the statistical significance reported by the authors. In this study, the effect was statistically significant and positive, according to WWC criteria.

O'Connor et al. (1993b) compared the segmenting and no-treatment comparison groups on the same nine measures in this outcome domain.[11] They reported statistically significant differences favoring the segmenting group for the three segmentation measures and one blending measure. They reported one nonstatistically significant finding for a rhyming measure. The authors did not report statistically significant findings for the other four measures, but some measures showed large effect sizes favoring the intervention group. The WWC confirmed the statistical significance reported by the authors for two of the segmentation measures and the blending measure, but not for the third segmentation measure. In this study, the effect was statistically significant and positive, according to WWC criteria.

O'Connor et al. (1993c) compared the rhyming and no-treatment comparison groups on the same nine measures in this outcome domain.[11] They reported statistically significant differences favoring the rhyming group for two of the three rhyming measures and no statistically significant differences for the third rhyming measure. The authors did not report statistical significance for the other six measures. The WWC was unable to confirm the statistical significance reported by the authors; however, the findings were large enough to categorize the effect as substantively important and positive, according to WWC criteria.

Data from Slocum et al. (1993a) were used to compare the blend-then-segment intervention group with the word manipulation-then-segment comparison group on two measures in this outcome domain. The authors did not report statistical significance information for either of these measures; however, the WWC determined that there was a statistically significant difference favoring the intervention group on the blending measure. In this study, the effect was statistically significant and positive, according to WWC criteria.

Data from Slocum et al. (1993b) were used to compare the segment-then-blend intervention group with the word manipulation-then-blend comparison group on two measures in this outcome domain. The authors did not report statistical significance information for either of these measures; however, the WWC determined that there was a statistically significant difference favoring the intervention group on the segmenting measure. In this study, the effect was statistically significant and positive, 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,[9] the size of the difference between participants in the intervention condition and the comparison condition, and the consistency in findings across studies (see the WWC Intervention Rating Scheme).

10. The WWC cannot use the author-reported significance levels because they are from an analysis of gain scores, which was not appropriate given that the pretest and posttest were two different tests in this study. The authors provided covariate adjusted posttest means.

11. O'Connor et al. (1993a, b, c) also administered a phonological mastery test to the intervention group children to assess how well they learned tasks in the intervention they received, but these were not administered to the comparison group children and were not considered in the impact analyses. Therefore, the WWC does not include results from this measure in this report.

The WWC found *Phonological Awareness Training* to have positive effects for phonological processing

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 entirely based on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analysis. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results. The average improvement index for phonological processing is +27 percentile points across the six studies, with a range of –27 to +50 percentile points across findings.

Findings for comparisons between variations of *Phonological Awareness Training*

The two studies described below do not contribute to the overall rating of effectiveness because the comparison involves two different variations of *Phonological Awareness Training,* which does not allow the effects of *Phonological Awareness Training* to be determined. However, the WWC believes that the findings from these comparisons provide useful information to practitioners who may be interested in comparing the effects of one variation of a practice to another variation of the same practice. The WWC reports the individual study findings here and in Appendices A4.1–A4.3.

Yeh (2003) included 44 four- to five-year-old low-income children attending four Head Start classrooms in two Head Start centers in the Boston, Massachusetts area. Forty-one percent of the children were Hispanic, 41% were African-American, 7% were Asian, and 11% were Caucasian. Yeh compared print knowledge, phonological processing, and early reading/writing outcomes for a group that participated in phonemic segmentation training with a group that participated in rhyme/alliteration training.

Maslanka and Joseph (2002) included 19 three- to five-year-old children attending a private preschool in a central Ohio suburban community. Most of the children studied were Caucasian. About half of the children were female. Maslanka and Joseph compared phonological processing outcomes for a group that used sound boxes with outcomes for a group that used sound sorts.

Print knowledge. Yeh (2003) analyzed findings for one measure in this outcome domain. The findings favored the phonemic segmentation group over the rhyme/alliteration group. Although the difference was not statistically significant (as calculated by the WWC), it was large enough to categorize the difference between the two groups as substantively important, according to WWC criteria. The improvement index for print knowledge is +33 percentile points for the one print knowledge outcome in this study.

Phonological processing. Yeh (2003) analyzed findings for four measures in this outcome domain.[12] The findings favored the phonemic segmentation group over the rhyme/alliteration group for all measures. Although the differences were not statistically significant (as calculated by the WWC), they were large enough to categorize the difference between the two groups as substantively important, according to WWC criteria. The average improvement index for phonological processing is +20 percentile points, with a range of +9 to +33 percentile points across findings.

12. Yeh (2003) also reported a combined phoneme awareness score. The WWC does not include the measure in this report, however, because it is a composite of the individual tests and does not provide additional information beyond what is included for the individual tests.

The WWC found *Phonological Awareness Training* to have positive effects for phonological processing *(continued)*

Maslanka and Joseph (2002) compared the sound box and sound sort groups on four measures in this outcome domain.[13] The findings favored the concrete representation of the sounds in words (sound box) condition over the categorization of words based on sounds alone (sound sort) condition. The authors did not report statistically significant differences between groups for any of these measures, but the findings were large enough to categorize the difference between the two groups as substantively important, according to WWC criteria. The average improvement index for phonological processing is +11 percentile points, with a range of –10 to +29 percentile points across findings.

Early reading/writing. Yeh (2003) analyzed findings for one measure in this outcome domain. The findings favored the phonemic segmentation group over the rhyme/alliteration group. Although the difference was not statistically significant (as calculated by the WWC), it was large enough to categorize the difference between the two groups as substantively important, according to WWC criteria. The improvement index for early reading/writing is +26 percentile points for the one early reading/writing outcome in this study.

Summary

The WWC reviewed 11 studies on *Phonological Awareness Training.* Four of these studies met WWC standards and two studies met WWC standards with reservations. Two additional studies that either met standards or met standards with reservations are described in this report but are not included in the overall rating of effectiveness. The remaining studies did not meet WWC evidence screens. Based on the six studies included in the overall rating of effectiveness, the WWC found positive effects for phonological processing.[14] Based on the study that compared *Phonological Awareness Training* that focused on phonemes with *Phonological Awareness Training* that focused solely on rhyme and alliteration, the WWC found potentially positive effects on print knowledge, phonological processing, and early reading/writing. Based on the study that compared *Phonological Awareness Training* that focused on the concrete representation of the sounds in words with *Phonological Awareness Training* that included categorization of words based on sounds alone, the WWC found potentially positive effects on phonological processing. The evidence presented in this report may change as new research emerges.

References

Met WWC evidence standards

Majsterek, D. J., Shorr, D. N., & Erion, V. L. (2000). Promoting early literacy through rhyme detection activities during Head Start circle-time. *Child Study Journal, 30*(3), 143–151.

O'Connor, R. E., Jenkins, J. R., Leicester, N., & Slocum, T. A. (1993a). Teaching phonological awareness to young children with learning disabilities. *Exceptional Children, 59*(6), 532–546. (Study: blending intervention versus no-treatment comparison group.)

O'Connor, R. E., Jenkins, J. R., Leicester, N., & Slocum, T. A. (1993b). Teaching phonological awareness to young children with learning disabilities. *Exceptional Children, 59*(6), 532–546. (Study: segmenting intervention versus no-treatment comparison group.)

O'Connor, R. E., Jenkins, J. R., Leicester, N., & Slocum, T. A. (1993c). Teaching phonological awareness to young children with learning disabilities. *Exceptional Children, 59*(6), 532–546. (Study: rhyming intervention versus no-treatment comparison group.)

13. Maslanka and Joseph (2003) also reported findings for the overall score of the Test of Phonological Awareness, Kindergarten Version and the overall score of the Test of Phonological Awareness, Early Education Version. The WWC does not include these measures in the review, however, because there was not enough information available to compute effect sizes.

14. In O'Connor et al. (1993a, b, c) and Slocum et al. (1993a, b), the effects on measures that were more aligned with the nature of the *Phonological Awareness Training* (i.e., blending, segmenting) were larger than the effects on measures less aligned with the nature of the *Phonological Awareness Training.*

References *(continued)*

Yeh, S. S. (2003). An evaluation of two approaches for teaching phonemic awareness to children in Head Start. *Early Childhood Research Quarterly, 18*(4), 513–529.

Met WWC evidence standards with reservations

Maslanka, P., & Joseph, L. M. (2002). A comparison of two phonological awareness techniques between samples of preschool children. *Reading Psychology, 23*(4), 271–288.

Slocum, T. A., O'Connor, R. E., & Jenkins, J. R. (1993a). Transfer among phonological manipulation skills. *Journal of Educational Psychology, 85*(4), 618–630. (Study: blend-then-segment intervention versus word manipulation-then-segment comparison group.)

Slocum, T. A., O'Connor, R. E., & Jenkins, J. R. (1993b). Transfer among phonological manipulation skills. *Journal of Educational Psychology, 85*(4), 618–630. (Study: segment-then-blend intervention versus word manipulation-then-blend comparison group.)

Additional source:

Slocum, T. A. (1992). The learning and transfer of two phonemic manipulation skills. *Dissertation Abstracts International, 52*(08), 2889A. (UMI No. 9203306).

Did not meet WWC evidence screens

Hatcher, P. J., Hulme, C., & Snowling, M. J. (2004). Explicit phoneme training combined with phonic reading instruction helps young children at risk of reading failure. *Journal of Child Psychology and Psychiatry, 45*(2), 338–358.[15]

Laing, S. P., & Espeland, W. (2005). Low intensity phonological awareness training in a preschool classroom for children with communication impairments. *Journal of Communication Disorders, 38,* 65–82.[16]

Layton, L., Deeny, K., Upton, G., & Tall, G. (1998). A pre-school training programme for children with poor phonological awareness: Effects on reading and spelling. *Journal of Research in Reading, 21*(1), 36–52.[16]

For more information about specific studies and WWC calculations, please see the WWC *Phonological Awareness Training* Technical Appendices.

15. Confound: there was only one group in each study condition, so the analysis could not separate the effects of the intervention from the effects of group membership.

16. Lack of evidence for baseline equivalence: the study, which used a quasi-experimental design, did not establish that the comparison group was equivalent to the intervention group at baseline.

Appendix

Appendix A1.1 Study characteristics: Majsterek, Shorr, & Erion, 2000 (randomized controlled trial)

Characteristic	Description
Study citation	Majsterek, D. J., Shorr, D. N., & Erion, V. L. (2000). Promoting early literacy through rhyme detection activities during Head Start circle-time. *Child Study Journal, 30*(3), 143–151.
Participants	The study included 40 low-income three- to five-year-old preschool children. The mean age of the intervention group was 55.2 months; the mean age of the comparison group was 54.3 months. Forty-five percent of the sample was female, and all children included in the analyses spoke English as their primary language. Four groups of children were randomly assigned to the intervention and comparison conditions.
Setting	The study took place at a Head Start program in Washington State.
Intervention	The intervention group participated in *Phonological Awareness Training* that focused on the concepts of rhyming and rhyme detection. The children were introduced to these concepts using pictures from *DaisyQuest* software[1] and objects from a rhyme box. Children generated rhyming words for the objects in the box. Each teaching session began with a review of the previous session's activity. The intervention group participated in nine 10-minute sessions during their regular circle time over a four-week period.
Comparison	The comparison group participated in training with a semantic emphasis (i.e., a focus on word meaning), with activities focusing on synonyms, comparative-superlative, position in space, and reasoning. The comparison group participated in nine 10-minute sessions during their regular circle time over the same four-week period.
Primary outcomes and measurement	The primary outcome domain was children's phonological processing. This domain was assessed with one researcher-developed measure called the rhyme detection test. (See Appendix A2.2 for more detailed descriptions of outcome measures.)
Teacher training	Implementation of both the intervention and comparison conditions was conducted by the first author of the study. The WWC found no reason to believe that the person implementing the intervention and comparison conditions was not equally trained and motivated to implement each condition.

1. A separate WWC intervention report details the effects of the *DaisyQuest* software on children's phonological processing skills.

Appendix A1.2 Study characteristics: O'Connor, Jenkins, Leicester, & Slocum, 1993a (randomized controlled trial)

Characteristic	Description
Study citation	O'Connor, R. E., Jenkins, J. R., Leicester, N., & Slocum, T. A. (1993a). Teaching phonological awareness to young children with disabilities. *Exceptional Children, 59*(6), 532–546. (Study: blending intervention versus no-treatment comparison group.)
Participants	*O'Connor et al. (1993a, b, c).* The study began with 55 four- to six-year-old developmentally delayed preschool children. Five children were removed from the study due to high pretest scores or autism, and three children left the preschool, leaving a sample of 47 children. Of these children, 80% had significant language delays and some had additional disabilities (e.g., physical disabilities, mental retardation, or behavior disorders). Children were blocked by mean age and cognitive ability and then randomly assigned to one of four groups: blending, segmenting, rhyming, or a no-treatment comparison. *O'Connor et al. (1993a).* The same comparison group was used in all of the O'Connor et al. studies included in this report. Results for the 22 children who had been randomly assigned to the blending or the comparison conditions are included in this report.
Setting	The study took place at the preschool located in the Experimental Education Unit of the University of Washington.
Intervention	The blending intervention group participated in blending training that took place in two phases. Phase one of the training lasted three weeks, and phase two lasted four weeks. During phase one, the children focused on learning one specific blending task (e.g., blend two to three phonemes when presented as continuous sounds). During phase two, the children reviewed the task learned in phase one and learned about other blending tasks (e.g., blend words with separated sounds and blend onset-rime). The children were taught in groups of three to five and met for 10 minutes four times a week. Results for phase two are not included in this report because the effects of the second condition cannot be separated from the effects of the first condition.
Comparison	Comparison group children participated in routine preschool activities, such as listening to stories read by their teachers or "circle time" oral language activities. Additionally, the researcher met twice with each child in the comparison group during phase two training to practice isolated sounds used in training.
Primary outcomes and measurement	The primary outcome domain was children's phonological processing. Nine nonstandardized subtests measured phonological processing skills. There were three blending subtests, three segmenting subtests, and three rhyming subtests. The study also administered a phonological mastery test to the intervention group children to assess how well they learned tasks in the intervention they received, but it is not included in this review because the test was not administered to the comparison group children and was not considered in the impact analyses. (See Appendix A2.2 for more detailed descriptions of outcome measures.)
Teacher training	Three graduate students with teaching experience provided instruction. The instructors met with the researcher each Monday to practice teaching formats for the week. Instructors were observed during their sessions and received additional training as needed. Instructors alternated teaching the conditions described in O'Connor et al. (1993a, b, c).

Appendix A1.3 Study characteristics: O'Connor, Jenkins, Leicester, & Slocum, 1993b (randomized controlled trial)

Characteristic	Description
Study citation	O'Connor, R. E., Jenkins, J. R., Leicester, N., & Slocum, T. A. (1993b). Teaching phonological awareness to young children with disabilities. *Exceptional Children, 59*(6), 532–546. (Study: segmenting intervention versus no-treatment comparison group.)
Participants	*O'Connor et al. (1993a, b, c).* The study began with 55 four- to six-year-old developmentally delayed preschool children. Five children were removed from the study due to high pretest scores or autism, and three children left the preschool, leaving a sample of 47 children. Of these children, 80% had significant language delays and some had additional disabilities (e.g., physical disabilities, mental retardation, or behavior disorders). Children were blocked by mean age and cognitive ability and randomly assigned to one of four groups: blending, segmenting, rhyming, or a no-treatment comparison. *O'Connor et al. (1993b).* The same comparison group was used in all of the O'Connor et al. studies included in this report. Results for the 24 children who had been randomly assigned to the segmenting or the comparison conditions are included in this report.
Setting	The study took place at the preschool located in the Experimental Education Unit of the University of Washington.
Intervention	The segmenting intervention group participated in segmenting training that took place in two phases. Phase one of the training lasted three weeks, and phase two lasted four weeks. During phase one, the children were instructed to separate sounds by stretching each sound in a two-to-three phoneme word as they said the word slowly. During phase two, the children were taught three skills: separate words into onset-rime, say each sound in a word, and indicate which sound is the first. The children were taught in groups of three to five and met for 10 minutes four times a week. Results for phase two are not included in this report because the effects of the second condition cannot be separated from the effects of the first condition.
Comparison	Comparison group children participated in routine preschool activities, such as listening to stories read by their teachers or "circle time" oral language activities. Additionally, the researcher met twice with each child in the comparison group during phase two training to practice isolated sounds used in training.
Primary outcomes and measurement	The primary outcome domain was children's phonological processing. Nine nonstandardized subtests measured auditory phonological skills. There were three blending subtests, three segmenting subtests, and three rhyming subtests. The study also administered a phonological mastery test to the intervention group children to assess how well they learned tasks in the intervention they received, but it is not included in this review because the test was not administered to the comparison group children and was not considered in the impact analyses. (See Appendix A2.2 for more detailed descriptions of outcome measures.)
Teacher training	Three graduate students with teaching experience provided instruction. The instructors met with the researcher each Monday to practice teaching formats for the week. Instructors were observed during their sessions and received additional training as needed. Instructors alternated teaching the conditions described in O'Connor et al. (1993a, b, c).

Appendix A1.4 Study characteristics: O'Connor, Jenkins, Leicester, & Slocum, 1993c (randomized controlled trial)

Characteristic	Description
Study citation	O'Connor, R. E., Jenkins, J. R., Leicester, N., & Slocum, T. A. (1993c). Teaching phonological awareness to young children with disabilities. *Exceptional Children, 59*(6), 532–546. (Study: rhyming intervention versus no-treatment comparison group.)
Participants	*O'Connor et al. (1993a, b, c).* The study began with 55 four- to six-year-old developmentally delayed preschool children. Five children were removed from the study due to high pretest scores or autism, and three children left the preschool, leaving a sample of 47 children. Of these children, 80% had significant language delays and some had additional disabilities (e.g., physical disabilities, mental retardation, or behavior disorders). Children were blocked by mean age and cognitive ability and randomly assigned to one of four groups: blending, segmenting, rhyming, or a no-treatment comparison. *O'Connor et al. (1993c).* The same comparison group was used in all of the O'Connor et al. studies included in this report. Results for the 23 children who had been randomly assigned to the rhyming or the comparison conditions are included in this report.
Setting	The study took place at the preschool located in the Experimental Education Unit of the University of Washington.
Intervention	The rhyming intervention group participated in rhyming training that took place in two phases. Phase one of the training lasted three weeks, and phase two lasted four weeks. During phase one, the children were given examples of rhyme, allowed to rhyme in a group, and were asked to make a rhyme. During phase two, children were asked to continue to make rhymes and additionally were asked to identify if pairs of words rhyme and to select a word that does not rhyme from a trio of words. The children were taught in groups of three to five and met for 10 minutes four times a week. Results for phase two are not included in this report because the effects of the second condition cannot be separated from the effects of the first condition.
Comparison	Comparison group children participated in routine preschool activities, such as listening to stories read by their teachers or "circle time" oral language activities. Additionally, the researcher met twice with each child in the comparison group during phase two training to practice isolated sounds used in training.
Primary outcomes and measurement	The primary outcome domain was children's phonological processing. Nine nonstandardized subtests measured auditory phonological skills. There were three blending subtests, three segmenting subtests, and three rhyming subtests. The study also administered a phonological mastery test to the intervention group children to assess how well they learned tasks in the intervention they received, but it is not included in this review because the test was not administered to the comparison group children and was not considered in the impact analyses. (See Appendix A2.2 for more detailed descriptions of outcome measures.)
Teacher training	Three graduate students with teaching experience provided instruction. The instructors met with the researcher each Monday to practice teaching formats for the week. Instructors were observed during their sessions and received additional training as needed. Instructors alternated teaching the conditions described in O'Connor et al. (1993a, b, c).

Appendix A1.5 Study characteristics: Yeh, 2003 (randomized controlled trial)

Characteristic	Description
Study citation	Yeh, S. S. (2003). An evaluation of two approaches for teaching phonemic awareness to children in Head Start. *Early Childhood Research Quarterly, 18*(4), 513–529. This study is not included in the overall effectiveness rating because it compared variations of *Phonological Awareness Training* to each other, which does not allow the effects of *Phonological Awareness Training* to be determined.
Participants	The study included 44 low-income four- to five-year-old children. Forty-one percent were Hispanic, 41% were African-American, 7% were Asian, and 11% were Caucasian. Four classrooms were matched on student achievement levels and randomly assigned to the phonemic segmentation and the rhyming/alliteration groups.
Setting	The study took place in four classrooms from two Head Start programs in the Boston, Massachusetts area.
Phonemic segmentation group	The children in this group participated in segmentation training concerning the segmentation, blending, and substitution of phonemes. The activities for this group were developed based on the Phono-Graphix program. To support children's learning, teachers modeled the expected behaviors (e.g., by exaggerating pronunciation), encouraged participation, and reinforced correct responses until eventually children were able to match sounds and graphemes and sound out words on their own. Instruction was conducted in small groups of three to five children for 20–25 minutes a week over a nine-week period.
Rhyming/alliteration group	The children in this group participated in rhyming and alliteration training that required them to rhyme and give words that have the same first consonant. They were also taught through bookmaking activities, during which they glued pictures into blank books and told stories to accompany the pictures. The activities for this group were developed from a commercially available phonemic awareness curriculum. Instruction was conducted in small groups of three to five children for 20–25 minutes a week over a nine-week period.
Primary outcomes and measurement	The primary outcome domains assessed were children's print knowledge, phonological processing, and early reading/writing. Print knowledge was assessed with a nonstandardized measure of letter-sound matching. Phonological awareness was assessed by four nonstandardized measures: phoneme blending, phoneme segmentation, phoneme deletion, and phoneme substitution. A combined phoneme awareness variable was also created based on the scores from the individual measures; however, this measure is not included in this review because the WWC includes the four individual measures of phonological processing. Children's early reading/writing was measured with a nonstandardized test of oral reading. (See Appendices A2.1–A2.3 for more detailed descriptions of outcome measures.)
Teacher training	Teachers received in-class modeling and coaching on the appropriate phonological instructional approach over a period of three weeks. During this three-week period, a group of children was taught by a consultant-trainer, then the teacher taught another group of children as the consultant provided coaching.

Appendix A1.6 Study characteristics: Maslanka & Joseph, 2002 (quasi-experimental design)

Characteristic	Description
Study citation	Maslanka, P., & Joseph, L. M. (2002). A comparison of two phonological awareness techniques between samples of preschool children. *Reading Psychology, 23*(4), 271–288. This study is not included in the overall effectiveness rating because it compared variations of *Phonological Awareness Training* to each other, which does not allow the effects of *Phonological Awareness Training* to be determined.
Participants	The study began with 20 three- to five-year-old children from middle-income families. One child left the preschool during the intervention, leaving a sample of 19 children. Fifty-three percent of the children were female, 95% were Caucasian, and 5% were Hispanic. Children were randomly selected to participate in the sound box or sound sort groups.
Setting	The study took place at a private preschool in a suburban community in Ohio.
Sound box group	Children in the sound box group were provided with cardboard with sound boxes drawn on the front and back and poker chips. A sound box is a drawn rectangle that has been divided into sections, with each section representing a sound in the word. In this study, the children used two-section and three-section sound boxes, corresponding with two-sound and three-sound words. As the children say the phonemes in a word, they move an object or write a letter in the box. In this study, poker chips were used for this purpose. After giving the children a sound box, the children participated in a choral responding exercise followed by a beginning and ending sounds exercise. The last activity involved having a student lead the group by saying the word in the picture, saying each sound of the word, and having the other children move their poker chips in the box as the lead student said each phoneme. Children were instructed in small groups for about 15 minutes a day for 26 days.
Sound sort group	Children in the sound sort group participated in sorting lessons using black and white illustrated picture cards from another curriculum. Each sorting lesson involved children sorting pictures into categories by beginning sounds, consonant sounds, middle sounds, or ending sounds. At the beginning of each lesson, the experimenter placed two pictures on the table and gave each child two category pictures. Then the experimenter presented two pictures at a time and said the word describing the picture, told the child which sound to focus on (beginning, consonant, middle, or end), and modeled the sorting procedure for about seven cards. The modeling procedure involved the experimenter turning a card over with the picture facing the child, saying the word and the target sound, and putting the card below the category picture card that had the same sound. The children did the same sorting procedure with seven different cards. Throughout the activity children were given corrective guidance and feedback. Children were instructed in small groups for about 15 minutes a day for 26 days.
Primary outcomes and measurement	The primary outcome domain was children's phonological processing, which was assessed with three standardized measures: the Phonological Awareness Test (PAT), the Test of Phonological Awareness, Kindergarten Version (TOPA-KV), and the Test of Phonological Awareness, Early Education Version (TOPA-EV). The researchers included 10 subtests from the PAT as well as total scores for each subtest. The WWC does not include the TOPA-KV or the TOPA-EV in this review because complete data were not reported in the article and effect sizes could not be calculated. (See Appendix A2.2 for more detailed descriptions of outcome measures.)
Teacher training	Both conditions were implemented by the lead toddler teacher at the private preschool where the interventions took place. The teacher was a graduate student in School Psychology and received training by a school psychology professor. The WWC found no reason to believe that the person implementing the intervention and comparison condition was not equally trained and motivated to implement each condition.

Appendix A1.7 Study characteristics: Slocum, O'Connor, & Jenkins 1993a (randomized controlled trial with attrition problems)

Characteristic	Description
Study citation	Slocum, T. A., O'Connor, R. E., & Jenkins, J. R. (1993a). Transfer among phonological manipulation skills. *Journal of Educational Psychology, 85*(4), 618–630. (Study: blend-then-segment intervention group versus word manipulation-then-segment group.)
Participants	*Slocum* et al. *(1993a, b).* The study began with 48 low-income preschool children randomly assigned to two intervention and two comparison conditions. Thirteen children left the preschools over the course of the study, leaving a sample of 35 children in the four groups.[1] Children in special education were excluded from the study sample. The mean age of the child participants was 5.2 years. Fifty-seven percent were female, 77% were Black, 14% were Asian, and 9% were Caucasian. *Slocum* et al. *(1993a).* Results for the 16 children who had been randomly assigned to the blend-then-segment intervention group and the word manipulation-then-segment comparison group are included in this report.
Setting	The study took place in four Head Start preschool classrooms in an urban area of the Pacific Northwest.
Intervention	The study included two phases. The children in the blend-then-segment intervention group were taught blending during phase one and segmenting during phase two. Children were taught how to blend using auditory blending instruction. Children were taught how to segment using onset-rime segmentation instruction. Instruction was one-on-one, lasted about 10 minutes per session, and continued until the child achieved mastery, which was defined as responding correctly to four of five items in a set of items for two days in a row. The WWC only includes the results from the phase one instructional period in this report because there was no appropriate comparison condition to estimate effects of training at the completion of phase two training.
Comparison	Children in the word manipulation-then-segment comparison group were taught word manipulation in phase one and segmenting in phase two. Comparison group children were randomly matched with intervention group children, and they participated in word manipulation activities until their intervention group counterpart achieved mastery in his or her first instructional phase. Instruction was one-on-one and lasted about 10 minutes per session.
Primary outcomes and measurement	The primary outcome domain was children's phonological processing, which was assessed with two nonstandardized measures of onset-rime blending and onset-rime segmenting. (See Appendix A2.2 for more detailed descriptions of outcome measures.)
Teacher training	The intervention was implemented by six graduate students in education. They were trained in the instructional procedures until they could deliver the instruction to adults with 100% accuracy. After training, they practiced with nonstudy children until they reached 100% accuracy.

1. The study was downgraded by the WWC due to overall attrition, which was 27%. A teacher strike interrupted the intervention and may have contributed to the high rates of attrition. There were 9 children in the intervention group and 17 children in the combined comparison group.

Appendix A1.8 Study characteristics: Slocum, O'Connor, & Jenkins 1993b (randomized controlled trial with attrition problems)

Characteristic	Description
Study citation	Slocum, T. A., O'Connor, R. E., & Jenkins, J. R. (1993b). Transfer among phonological manipulation skills. *Journal of Educational Psychology, 85*(4), 618–630. (Study: segment-then-blend intervention group versus word manipulation-then-blend group.)
Participants	*Slocum et al. (1993a, b).* The study began with 48 low-income preschool children randomly assigned to two intervention and two comparison conditions. Thirteen children left the preschools over the course of the study, leaving a sample of 35 children in the four groups.[1] Children in special education were excluded from the study sample. The mean age of the child participants was 5.2 years. Fifty-seven percent were female, 77% were Black, 14% were Asian, and 9% were Caucasian. *Slocum et al. (1993b).* Results for the 19 children who had been randomly assigned to the segment-then-blend intervention group and the word manipulation-then-blend comparison group are included in this report.
Setting	The study took place in four Head Start preschool classrooms in an urban area of the Pacific Northwest.
Intervention	The study included two phases. Children in the segment-then-blend intervention group were taught segmenting during phase one and blending during phase two. Children were taught how to blend using auditory blending instruction. Instruction was one-on-one, lasted about 10 minutes per session, and continued until the child achieved mastery, which was defined as responding correctly to four of five items in a set of items for two days in a row. The WWC only includes the results from the phase one instructional period in this report because there was no appropriate comparison condition to estimate effects of training at the completion of phase two training.
Comparison	Children in the word manipulation-then-blend comparison group were taught word manipulation in phase one and blending in phase two. Comparison group children were randomly matched with intervention group children, and they participated in word manipulation activities until their intervention group counterpart achieved mastery in his or her first instructional phase. Instruction was one-on-one and lasted about 10 minutes per session.
Primary outcomes and measurement	The primary outcome domain was children's phonological processing, which was assessed with two nonstandardized measures of onset-rime blending and onset-rime segmenting. (See Appendix A2.2 for more detailed descriptions of outcome measures.)
Teacher training	The intervention was implemented by six graduate students in education. They were trained in the instructional procedures until they could deliver the instruction to adults with 100% accuracy. After training, they practiced with nonstudy children until they reached 100% accuracy.

1. The study was downgraded by the WWC due to overall attrition, which was 27%. A teacher strike interrupted the intervention and may have contributed to the high rates of attrition. There were 9 children in the intervention group and 17 children in the combined comparison group.

Appendix A2.1 Outcome measure in the print knowledge domain

Outcome measure	Description
Letter-sound matching	A measure in which children are presented with each of the letters of the alphabet except "q" and asked what sound or sounds the letter makes to assess children's phonemic awareness (as cited in Yeh, 2003).

Appendix A2.2 Outcome measures in the phonological processing domain

Outcome measure	Description
Rhyme detection test	A researcher-developed measure that assesses children's ability to identify rhyming words. The child is presented with four pictures in a row, the teacher points to and identifies each picture, and then the teacher asks the child which of the final three words pictured rhymes with the initial one (as cited in Majsterek et al., 2000).
Blending: continuous sounds	A researcher-developed measure that requires children to blend stretched words (e.g., "SSSSaaaammm") to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Blending: onset-rime	A researcher-developed measure that requires children to blend onset-rimes (e.g., "S-am") to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Blending: separated sounds	A researcher-developed measure that requires children to blend separated sounds (e.g., "S-a-m") to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Segmenting: all sounds	A researcher-developed measure that requires children to separate all of the sounds in two- and three-phoneme words (e.g., say all of the sounds in "mob") to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Segmenting: onset-rime	A researcher-developed measure that requires children to separate words into onset-rime after being trained on a number of examples (e.g., say "m-ob") to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Segmenting: first sound	A researcher-developed measure that requires children to identify the first sound in words (e.g., what is the first sound in "mob"?) to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Rhyming: production	A researcher-developed measure that requires children to provide a rhyming word for a word that is presented to them (e.g., what is a word that rhymes with "land"?) to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Rhyming: oddity	A researcher-developed measure that requires children to identify the nonrhyming word among a set of three words (e.g., "cat," "hat," and "bell") to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Rhyming: recognition	A researcher-developed measure that requires children to identify whether or not a pair of words rhyme (e.g., "dime" and "time") to assess children's auditory phonological skills (as cited in O'Connor et al., 1993a, b, c).
Phoneme blending	A measure that requires children to blend separate sounds (e.g., /p/ /i/ /g/) into a word (e.g., "pig") to assess children's phonemic awareness (as cited in Yeh, 2003).
Phoneme segmentation	A measure that requires children to identify each of the individual sounds in words (e.g., "dog": "What is the first sound? Second sound? Last sound?") to assess children's phonemic awareness (as cited in Yeh, 2003).
Phoneme deletion	A measure that requires children to say words and then delete certain sounds (e.g., "say 'sip' without the 's'") to assess children's phonemic awareness (as cited in Yeh, 2003).

(continued)

Appendix A2.2 Outcome measures in the phonological processing domain *(continued)*

Outcome measure	Description
Phoneme substitution	A measure that asks children what sound is needed to change one word into another word (e.g., "cat" into "mat") to assess children's phonemic awareness (as cited in Yeh, 2003).
Rhyming: discrimination	A subtest of a standardized test—the Phonological Awareness Test (PAT)—that requires children to identify whether or not a pair of words rhymes to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Rhyming: production	A subtest of the PAT that requires children to provide a rhyming word for a word that is presented to them to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Rhyming: total	A composite measure of the rhyming-discrimination and rhyming-production subtests of the PAT (as cited in Maslanka & Joseph, 2002).
Segmentation: sentences	A subtest of the PAT that requires children to clap once for each word in a sentence that is read to them to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Segmentation: syllables	A subtest of the PAT that requires children to clap once for each syllable or word part to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Segmentation: phonemes	A subtest of the PAT in which children are presented with a word and required to say each sound in the word to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Segmentation: total	A composite measure of the segmentation-sentences, segmentation-syllables, and segmentation-phonemes subtests of the PAT (as cited in Maslanka & Joseph, 2002).
Isolation: initial	A subtest of the PAT in which children are presented with a word and required to say its initial sound to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Isolation: medial	A subtest of the PAT in which children are presented with a word and required to say its middle sound to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Isolation: final	A subtest of the PAT in which children are presented with a word and required to say its final sound to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Isolation: total	A composite measure of the isolation-initial, isolation-medial, and isolation-final subtests of the PAT (as cited in Maslanka & Joseph, 2002).
Blending: syllables	A subtest of the PAT in which children are presented with separate word parts and required to say the word to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Blending: phonemes	A subtest of the PAT in which children are presented with separate sounds and required to blend them into a word to assess children's phonological skills (as cited in Maslanka & Joseph, 2002).
Blending: total	A composite measure of the blending-syllables and blending-phonemes subtests of the PAT (as cited in Maslanka & Joseph, 2002).
Onset-rime segmentation	A researcher-developed measure that requires children to separate the initial phoneme from the rime (e.g., separate /k/ from "at" in the word "cat") in three-phoneme words to assess children's phonological processing skills (as cited in Slocum et al., 1993a, b).
Onset-rime blending	A researcher-developed measure that requires children to blend all of the sounds in three-phoneme words (e.g., children hear "/k/….at" and say "cat") to assess children's phonological processing skills (as cited in Slocum et al., 1993a, b).

Appendix A2.3 Outcome measure in the early reading/writing domain

Outcome measure	Description
Oral reading	A researcher-developed measure that requires children to read simple sentences with mostly consonant-vowel-consonant words (e.g., "Fat cat sat on mat.") to assess children's decoding ability (as cited in Yeh, 2003).

Appendix A3 Summary of study findings included in the rating for the phonological processing domain[1]

Outcome measure	Study sample	Sample size (classrooms or groups/children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Phonological Awareness Training* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Phonological Awareness Training* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at α = 0.05)	WWC calculations: Improvement index[7]
Majsterek et al., 2000 (randomized controlled trial)[8]								
Rhyme detection test	3–5 year olds	4/40	5.40 (3.28)	4.31 (2.91)	1.09	0.34	ns	+13
Average[9] for phonological processing (Majsterek et al., 2000)						0.34	ns	+13
O'Connor et al., 1993a (randomized controlled trial)[10]								
Blending: continuous sounds	4–6 year olds	22	8.30 (2.80)	2.40 (3.70)	5.90	1.73	Statistically significant	+46
Blending: onset-rime	4–6 year olds	22	5.20 (3.40)	0.80 (2.10)	4.40	1.50	Statistically significant	+43
Blending: separated sounds	4–6 year olds	22	5.20 (3.20)	0.70 (0.60)	4.50	1.88	Statistically significant	+47
Segmenting: all sounds	4–6 year olds	22	0.00[11] (0.00)	0.00[11] (0.00)	0.00	0.00	ns	0
Segmenting: onset-rime	4–6 year olds	22	0.00[11] (0.00)	0.00[11] (0.00)	0.00	0.00	ns	0
Segmenting: first sound	4–6 year olds	22	0.00[11] (0.00)	0.00[11] (0.00)	0.00	0.00	ns	0
Rhyming: production	4–6 year olds	22	2.80 (4.20)	1.80 (4.00)	1.00	0.24	ns	+9
Rhyming: oddity	4–6 year olds	22	1.30 (2.50)	1.50 (3.00)	–0.20	–0.07	ns	–3
Rhyming: recognition	4–6 year olds	22	5.20 (2.20)	5.20 (2.30)	0.00	0.00	ns	0
Average[9] for phonological processing (O'Connor et al., 1993a)						0.59	ns	+22

(continued)

Appendix A3 Summary of study findings included in the rating for the phonological processing domain *(continued)*

Outcome measure	Study sample	Sample size (classrooms or groups/children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Phonological Awareness Training* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Phonological Awareness Training* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at α = 0.05)	WWC calculations: Improvement index[7]
O'Connor et al., 1993b (randomized controlled trial)[12]								
Blending: continuous phonemes	4–6 year olds	24	6.90 (3.80)	2.40 (3.70)	4.50	1.16	Statistically significant	+38
Blending: onset-rime	4–6 year olds	24	0.80 (1.50)	0.80 (2.10)	0.00	0.00	ns	0
Blending: separated sounds	4–6 year olds	24	1.60 (1.30)	0.70 (0.60)	0.90	0.83	ns	+30
Segmenting: all sounds	4–6 year olds	24	3.70 (3.60)	0.00[13] (0.00)	3.70	1.34	Statistically significant	+41
Segmenting: onset-rime	4–6 year olds	24	2.40 (2.80)	0.00[13] (0.00)	2.40	1.12	Statistically significant	+37
Segmenting: first sound	4–6 year olds	24	2.10 (3.70)	0.00[13] (0.00)	2.10	0.74	ns	+27
Rhyming: production	4–6 year olds	24	2.80 (3.80)	1.80 (4.00)	1.00	0.25	ns	+10
Rhyming: oddity	4–6 year olds	24	1.70 (2.20)	1.50 (3.00)	0.20	0.07	ns	+3
Rhyming: recognition	4–6 year olds	24	5.50 (2.20)	5.20 (2.30)	0.30	0.13	ns	+5
Average[9] for phonological processing (O'Connor et al., 1993b)						0.63	ns	+24
O'Connor et al., 1993c (randomized controlled trial)[14]								
Blending: continuous phonemes	4–6 year olds	23	4.70 (3.30)	2.40 (3.70)	2.30	0.63	ns	+24
Blending: onset-rime	4–6 year olds	23	2.00 (2.90)	0.80 (2.10)	1.20	0.45	ns	+18
Blending: separated sounds	4–6 year olds	23	2.30 (2.20)	0.70 (0.60)	1.60	0.94	ns	+33

(continued)

Appendix A3 Summary of study findings included in the rating for the phonological processing domain *(continued)*

Outcome measure	Study sample	Sample size (classrooms or groups/children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Phonological Awareness Training* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Phonological Awareness Training* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at α = 0.05)	WWC calculations: Improvement index[7]
O'Connor et al., 1993c (randomized controlled trial) *(continued)*								
Segmenting: all sounds	4–6 year olds	23	0.20 (0.40)	0.00[15] (0.00)	0.20	0.67	ns	+25
Segmenting: onset-rime	4–6 year olds	23	0.10 (0.30)	0.00[15] (0.00)	0.10	0.44	ns	+17
Segmenting: first sound	4–6 year olds	23	0.40 (1.40)	0.00[15] (0.00)	0.40	0.38	ns	+15
Rhyming: production	4–6 year olds	23	6.40 (3.80)	1.80 (4.00)	4.60	1.14	ns	+37
Rhyming: oddity	4–6 year olds	23	4.30 (3.40)	1.50 (3.00)	2.80	0.84	ns	+30
Rhyming: recognition	4–6 year olds	23	7.10 (1.90)	5.20 (2.30)	1.90	0.87	ns	+31
Average[9] for phonological processing (O'Connor et al., 1993c)						0.71	ns	+26
Slocum et al., 1993a (randomized controlled trial with attrition problems)[16]								
Onset-rime blending	Preschool children	16	4.87 (2.28)	0.86 (0.90)	4.01	2.08	Statistically significant	+48
Onset-rime segmenting	Preschool children	16	0.00 (0.00)	1.14 (2.27)	–1.14	–0.73	ns	–27
Average[9] for phonological processing (Slocum et al., 1993a)						0.68	ns	+25
Slocum et al., 1993b (randomized controlled trial with attrition problems)[17]								
Onset-rime blending	Preschool children	19	0.53 (1.80)	0.70 (1.34)	–0.17	–0.10	ns	–4
Onset-rime segmenting	Preschool children	19	8.78 (1.69)	1.00 (3.16)	7.78	2.89	Statistically significant	+50

(continued)

Appendix A3 Summary of study findings included in the rating for the phonological processing domain *(continued)*

Outcome measure	Study sample	Sample size (classrooms or groups/children)	Author's findings from the study: Mean outcome (standard deviation[2]) — *Phonological Awareness Training* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Comparison group[3]	WWC calculations: Mean difference[4] (*Phonological Awareness Training* – comparison)	WWC calculations: Effect size[5]	WWC calculations: Statistical significance[6] (at α = 0.05)	WWC calculations: Improvement index[7]
Average[9] for phonological processing (Slocum et al., 1993b)						1.39	Statistically significant	+42
Domain average[9] for phonological processing across all studies						0.72	na	+27

ns = not statistically significant
na = not applicable

1. This appendix reports findings considered for the effectiveness rating and the improvement indices. Findings for Yeh (2003) and Maslanka and Joseph (2002) are reported in Appendix A4.2 because they compare two *Phonological Awareness Training* interventions to each other, which does not allow the effects of *Phonological Awareness Training* to be determined. Subtest findings from the same studies are not included in these ratings, but are reported in Appendix A5.
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. For Slocum et al. (1993a, b), the standard deviations were calculated from graphs in the original article.
3. For Majsterek et al. (2000), the posttest means are covariate-adjusted means provided by the study author. For Slocum et al. (1993a, b), the means were calculated from graphs in the original article, and the intervention group mean equals the comparison group mean plus the mean difference.
4. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group. For Slocum et al. (1993a, b), 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 effects when the intervention group had lower pretest scores than the comparison group and underestimate the intervention's effect when the intervention group had higher pretest scores than the comparison 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 intervention condition and that of the average student in the comparison condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results.
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 Majsterek et al. (2000), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.
9. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.
10. In the case of O'Connor et al. (1993a), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
11. O'Connor et al. (1993a) reported means and standard deviations of zero for both the intervention and comparison groups.
12. In the case of O'Connor et al. (1993b), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
13. O'Connor et al. (1993b) reported means and standard deviations of zero for the comparison group.
14. In the case of O'Connor et al. (1993c), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
15. O'Connor et al. (1993c) reported means and standard deviations of zero for the comparison group.
16. In the case of Slocum et al. (1993a), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
17. In the case of Slocum et al. (1993b), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.

Appendix A4.1 Summary of findings for comparisons between variations of *Phonological Awareness Training* for the print knowledge domain[1]

Outcome measure	Study sample	Sample size (classrooms or groups/children)	Author's findings from the study: Mean outcome (standard deviation[2]) — First *Phonological Awareness Training* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Second *Phonological Awareness Training* group[4]	WWC calculations: Mean difference[5] (first group – second group)	WWC calculations: Effect size[6]	WWC calculations: Statistical significance[7] (at α = 0.05)	WWC calculations: Improvement index[8]
Yeh, 2003 (randomized controlled trial)[9]								
Letter-sound matching	4–5 year olds	4/44	8.17 (7.41)	2.00 (4.64)	6.17	0.97	ns	+33
Domain average[10] for phonological processing						0.97	ns	+33

ns = not statistically significant

1. This appendix presents a summary of study findings for measures that fall in the print knowledge domain for one study that is not included in the overall effectiveness ratings.
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 first *Phonological Awareness Training* group mean equals the second *Phonological Awareness Training* group mean plus the mean difference. The first *Phonological Awareness Training* group is the phonemic segmentation condition.
4. The second *Phonological Awareness Training* group is the rhyming/alliteration condition.
5. Positive differences and effect sizes favor the first *Phonological Awareness Training* group; negative differences and effect sizes favor the second *Phonological Awareness Training* 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 effects when the first *Phonological Awareness Training* group had lower pretest scores than the second *Phonological Awareness Training* group and underestimate the effects when the first *Phonological Awareness Training* group had higher pretest scores than the second *Phonological Awareness Training* group.
6. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
8. The improvement index represents the difference between the percentile rank of the average student in the first *Phonological Awareness Training* condition and that of the average student in the second *Phonological Awareness Training* condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results to the first *Phonological Awareness Training* condition.
9. 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 Yeh (2003), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.
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 A4.2 Summary of findings for comparisons between variations of *Phonological Awareness Training* for the phonological processing domain[1]

Outcome measure	Study sample	Sample size (classrooms or groups/children)	Author's findings from the study: Mean outcome (standard deviation[2]) — First *Phonological Awareness Training* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Second *Phonological Awareness Training* group[4]	WWC calculations: Mean difference[5] (first group – second group)	WWC calculations: Effect size[6]	WWC calculations: Statistical significance[7] (at α = 0.05)	WWC calculations: Improvement index[8]
Yeh, 2003 (randomized controlled trial)[9]								
Phoneme blending	4–5 year olds	4/44	3.98 (3.10)	3.24 (3.71)	0.74	0.21	ns	+9
Phoneme segmentation	4–5 year olds	4/44	10.23 (11.93)	3.67 (8.95)	6.56	0.61	ns	+23
Phoneme deletion	4–5 year olds	4/44	1.09 (2.79)	0.43 (0.87)	0.66	0.31	ns	+12
Phoneme substitution	4–5 year olds	4/44	4.00 (5.67)	0.00 (0.00)	4.00	0.96	ns	+33
Average[10] for phonological processing (Yeh, 2003)						0.52	ns	+20
Maslanka & Joseph, 2002 (quasi-experimental design)[11]								
Rhyming: total	3–5 year olds	19	13.90 (5.60)	15.40 (6.17)	–1.50	–0.24	ns	–10
Segmentation: total	3–5 year olds	19	14.10 (4.34)	12.90 (3.84)	1.20	0.28	ns	+11
Isolation: total	3–5 year olds	19	9.40 (3.14)	6.50 (3.72)	2.90	0.80	ns	+29
Blending: total	3–5 year olds	19	8.20 (3.93)	7.30 (1.80)	0.90	0.29	ns	+11
Average[10] for phonological processing (Maslanka & Joseph, 2002)						0.28	ns	+11
Domain average[10] for phonological processing across all studies						0.40	na	+16

ns = not statistically significant
na = not applicable

1. This appendix presents a summary of study findings for measures that fall in the phonological processing domain for two studies that are not included in the overall effectiveness ratings.
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.

(continued)

Appendix A4.2 Summary of findings for comparisons between variations of *Phonological Awareness Training* for the phonological processing domain *(continued)*

3. For Yeh (2003), the first *Phonological Awareness Training* group mean equals the second *Phonological Awareness Training* group mean plus the mean difference. For Yeh (2003), the first *Phonological Awareness Training* group is the phonemic segmentation condition. For Maslanka and Joseph (2002), the first *Phonological Awareness Training* group is the sound box condition.
4. For Yeh (2003), the second *Phonological Awareness Training* group is the rhyming/alliteration condition. For Maslanka and Joseph (2002), the second *Phonological Awareness Training* group is the sound sort condition.
5. Positive differences and effect sizes favor the first *Phonological Awareness Training* group; negative differences and effect sizes favor the second *Phonological Awareness Training* group. For Yeh (2003), 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 effects when the first *Phonological Awareness Training* group had lower pretest scores than the second *Phonological Awareness Training* group and underestimate the effects when the first *Phonological Awareness Training* group had higher pretest scores than the second *Phonological Awareness Training* group.
6. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
8. The improvement index represents the difference between the percentile rank of the average student in the first *Phonological Awareness Training* condition and that of the average student in the second *Phonological Awareness Training* condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results to the first *Phonological Awareness Training* condition.
9. 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 Yeh (2003), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.
10. The WWC-computed average effect sizes for each study and for the domain across studies are simple averages rounded to two decimal places. The average improvement indices are calculated from the average effect sizes.
11. In the case of Maslanka and Joseph (2002), no corrections for clustering or multiple comparisons were needed.

Appendix A4.3 Summary of findings for comparisons between variations of *Phonological Awareness Training* for the early reading/writing domain[1]

Outcome measure	Study sample	Sample size (classrooms/ children)	Author's findings from the study: Mean outcome (standard deviation[2]) — First *Phonological Awareness Training* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Second *Phonological Awareness Training* group[4]	WWC calculations: Mean difference[5] (first group – second group)	WWC calculations: Effect size[6]	WWC calculations: Statistical significance[7] (at α = 0.05)	WWC calculations: Improvement index[8]
Yeh, 2003 (randomized controlled trial)[9]								
Oral reading	4–5 year olds	4/44	1.71 (2.77)	0.24 (0.77)	1.47	0.70	ns	+26
Domain average[10] for early reading/writing						0.70	ns	+26

ns = not statistically significant

1. This appendix presents a summary of study findings for measures that fall in the early reading/writing domain for one study that is not included in the overall effectiveness ratings.
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 first *Phonological Awareness Training* group mean equals the second *Phonological Awareness Training* group mean plus the mean difference. The first *Phonological Awareness Training* group is the phonemic segmentation condition.
4. The second *Phonological Awareness Training* group is the rhyming/alliteration condition.
5. Positive differences and effect sizes favor the first *Phonological Awareness Training* group; negative differences and effect sizes favor the second *Phonological Awareness Training* 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 effects when the first *Phonological Awareness Training* group had lower pretest scores than the second *Phonological Awareness Training* group and underestimate the effects when the first *Phonological Awareness Training* group had higher pretest scores than the second *Phonological Awareness Training* group.
6. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
8. The improvement index represents the difference between the percentile rank of the average student in the first *Phonological Awareness Training* condition and that of the average student in the second *Phonological Awareness Training* condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results to the first *Phonological Awareness Training* condition.
9. 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 Yeh (2003), a correction for clustering was needed, so the significance levels may differ from those reported in the original study.
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 A5 Summary of subtest findings for comparisons between variations of *Phonological Awareness Training* for the phonological processing domain[1]

Outcome measure	Study sample	Sample size (children)	Author's findings from the study: Mean outcome (standard deviation[2]) — First *Phonological Awareness Training* group[3]	Author's findings from the study: Mean outcome (standard deviation[2]) — Second *Phonological Awareness Training* group[4]	WWC calculations: Mean difference[5] (first group – second group)	WWC calculations: Effect size[6]	WWC calculations: Statistical significance[7] (at α = 0.05)	WWC calculations: Improvement index[8]
Maslanka & Joseph, 2002 (quasi-experimental design)[9]								
Rhyming: discrimination	3–5 year olds	19	8.60 (1.58)	8.40 (1.90)	0.20	0.11	ns	+4
Rhyming: production	3–5 year olds	19	5.30 (4.24)	7.00 (4.62)	–1.70	–0.37	ns	–14
Segmentation: sentences	3–5 year olds	19	7.80 (1.56)	8.00 (2.83)	–0.20	–0.08	ns	–3
Segmentation: syllables	3–5 year olds	19	5.40 (2.60)	4.70 (1.42)	0.70	0.32	ns	+13
Segmentation: phonemes	3–5 year olds	19	0.90 (1.83)	0.20 (0.42)	0.70	0.52	ns	+20
Isolation: initial	3–5 year olds	19	5.30 (3.97)	4.10 (4.33)	1.20	0.28	ns	+11
Isolation: medial	3–5 year olds	19	0.40 (1.33)	0.00 (0.00)	0.40	0.42	ns	+16
Isolation: final	3–5 year olds	19	3.70 (3.77)	2.40 (2.40)	1.30	0.40	ns	+16
Blending: syllables	3–5 year olds	19	6.30 (2.83)	6.30 (1.20)	0.00	0.00	ns	0
Blending: phonemes	3–5 year olds	19	1.90 (1.69)	1.00 (0.60)	0.90	0.69	ns	+26

ns = not statistically significant

1. This appendix presents subtest findings for measures that fall in the phonological processing domain. Total test scores were used for rating purposes and are presented in Appendix A3.
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 first *Phonological Awareness Training* group is the sound box condition.
4. The second *Phonological Awareness Training* group is the sound sort condition.

(continued)

Appendix A5 Summary of subtest findings for comparisons between variations of *Phonological Awareness Training* for the phonological processing domain *(continued)*

5. Positive differences and effect sizes favor the first *Phonological Awareness Training* group; negative differences and effect sizes favor the second *Phonological Awareness Training* group.
6. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
7. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
8. The improvement index represents the difference between the percentile rank of the average student in the first *Phonological Awareness Training* condition and that of the average student in the second *Phonological Awareness Training* condition. The improvement index can take on values between –50 and +50, with positive numbers denoting favorable results to the first *Phonological Awareness Training* condition.
9. 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 Maslanka and Joseph (2002), no correction for clustering was needed.

Appendix A6 *Phonological Awareness Training* rating for the phonological processing domain

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.[1]

For the outcome domain of phonological processing, the WWC rated *Phonological Awareness Training* as having positive effects. The remaining ratings (potentially positive effects, mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered because *Phonological Awareness Training* was assigned the highest applicable rating.

Rating received
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.
Met. Four of the six studies that reported phonological processing outcomes found statistically significant and positive effects. Two of these four studies met WWC evidence standards for a strong design.
• Criterion 2: No studies showing statistically significant or substantively important *negative* effects.
Met. No studies showed statistically significant or substantively important negative 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.