WWC Intervention Report

A summary of findings from a systematic review of the evidence

ies INSTITUTE OF EDUCATION SCIENCES

Beginning Reading Updated November 2015

Lindamood Phoneme Sequencing® (LiPS®)

Report Contents

This intervention report presents findings from a systematic review of *LiPS*® conducted using the WWC Procedures and Standards Handbook, version 3.0, and the Beginning Reading review protocol, version 3.0.

Program Description[1]

The *Lindamood Phoneme Sequencing*® (*LiPS*®) program (formerly known as *Auditory Discrimination in Depth*® *[ADD*®*]*) is designed to improve reading and spelling skills by teaching students the skills needed to decode and encode words and to identify individual sounds and blends in words. Initial activities engage students in discovering the lip, tongue, and mouth actions needed to produce specific sounds. Once students are able to produce, label, and organize the sounds with their mouths, they work on identifying and ordering them within words during sequencing, reading, and spelling activities. The program also offers direct instruction in letter patterns, sight words, and context clues in reading. *LiPS*® is designed for emergent readers in grades K–3 and can be used with students of all ages who have learning disabilities or reading difficulties. The program is individualized to meet students' needs. This report focuses on the effectiveness of *LiPS*® on beginning readers in regular education classes. Studies reviewed in this report include research on a version of the *LiPS*® program that utilizes computer-supported activities.

Research[2]

The What Works Clearinghouse (WWC) identified two studies of *LiPS*® that both fall within the scope of the Beginning Reading topic area and meet WWC group design standards. Two studies meet WWC group design standards without reservations, and no studies meet WWC group design standards with reservations. Together, these studies included 97 beginning readers in first grade in five elementary schools.

The WWC considers the extent of evidence for *LiPS*® on the reading achievement of elementary students to be small for two outcome domains—comprehension and alphabetics. There were no studies that meet WWC design standards in the two other domains, so this intervention report does not report on the effectiveness of *LiPS*® for those domains. (See the Effectiveness Summary on p. 5 for more details of effectiveness by domain.)

Effectiveness

LiPS® was found to have potentially positive effects on comprehension and mixed effects on alphabetics for beginning readers.

Table 1. Summary of findings[3]

Outcome domain	Rating of effectiveness	Improvement index *(percentile points)*		Number of studies	Number of students	Extent of evidence
		Average	Range			
Comprehension	Potentially positive effects	+20	na	1	74	Small
Alphabetics	Mixed effects	+6	–13 to +24	2	97	Small

na = not applicable

Program Information

Background

Developed by Patricia Lindamood and Phyllis Lindamood, the *LiPS*® program is published by PRO-ED, Inc. and is available through a number of professional distributors and publishers. *ADD*® was developed in 1969 and revised in 1975. It was revised again and renamed the *Lindamood Phoneme Sequencing*® *(LiPS®)* program in 1998. Address: 416 Higuera Street, San Luis Obispo, CA 93401. Website: http://www.lindamoodbell.com. Telephone: (800) 233-1819.

Program details

During *LiPS*® instruction, teachers work with students in whole-class, small-group, and one-on-one settings to help them become aware of the mouth actions that produce speech sounds. Instructors help students verify sounds within words and teach them to self-correct in reading, spelling, and speech. The program developer recommends that instruction last 4–6 months for 1 hour a day, or 4–6 weeks for 4 hours a day. Computer-supported activities are available for the program. Lindamood Bell offers *LiPS*® workshops to train teachers, but teachers can also learn to administer the program from the *LiPS*® *Teacher's Manual*.

Cost

A kit of materials designed for classroom use costs $444.95.[4] Kits include the *LiPS*® *Teacher's Manual, Teacher's Guide*, and companion DVD, along with all student materials (stories, magnets, magnetic board, mirror, and picture cards). Some of these materials are also sold separately. A 3-day professional development workshop on *LiPS*® is available for a cost of $879. During this workshop, all steps of the intervention are introduced, modeled, and practiced.

Research Summary

The WWC identified eight eligible studies that investigated the effects of *LiPS*® on literacy for beginning readers. An additional 40 studies were identified but do not meet WWC eligibility criteria for review in this topic area. Citations for all 48 studies are in the References section, which begins on p. 7.

Table 2. Scope of reviewed research

Grade	1
Delivery method	Small group
Program type	Supplement

The WWC reviewed eight of those studies against group design standards. Two studies (Gunn, 1996; Torgesen, Wagner, Rashotte, Herron, & Lindamood, 2010) are randomized controlled trials that meet WWC group design standards without reservations. Those two studies are summarized in this report. Six studies do not meet WWC group design standards.

Summary of studies meeting WWC group design standards without reservations

Gunn (1996) conducted a cluster randomized controlled trial that examined the effects of *LiPS*® on first-grade students attending two elementary schools in one Pacific Northwest school district during the 1995–96 school year. Within each school, groups of students were randomly assigned either to the intervention condition (*LiPS*®) or the comparison condition (a basal reading program). The WWC based its effectiveness rating on findings from the 23 first-grade students who participated in the study; 11 students in the *LiPS*® group and 12 in the basal reading group. Eleven students were assigned to receive a modified version of the *LiPS*® program, which is not eligible for review. Outcomes for these students were not used to determine the WWC effectiveness rating and are not presented in this report. The study reported student outcomes after approximately 8 weeks of program implementation.

Torgesen et al. (2010) conducted a randomized controlled trial that examined the effects of *LiPS*® on first-grade students attending three elementary schools in 2 consecutive school years. Within each school, students identified as at risk for difficulty learning to read were randomly assigned either to the *LiPS*® group, the *Read, Write and Type (RWT*) group, or the comparison group (*Open Court's Collections for Young Scholars*, a basal reading program). Approximately 75% of students in the intervention groups received the intervention as supplemental reading instruction, in addition to standard reading instruction. The WWC based its effectiveness rating on outcomes measured at the end of the instructional period, which lasted from October through May, from 74 first-grade students who participated in the study; 35 students in the *LiPS*® group and 39 students in the comparison group. The study also reported supplemental findings for an additional 34 students assigned to the *RWT* group, as well as 1-year follow-up findings for the *LiPS*® vs. comparison contrast.

Summary of studies meeting WWC group design standards with reservations

No studies of *LiPS*® met WWC group design standards with reservations.

Effectiveness Summary

The WWC review of *LiPS*® for the Beginning Reading topic area includes outcomes for students in four domains: comprehension, alphabetics, reading fluency, and general reading achievement. The two studies of *LiPS*® that meet WWC group design standards reported findings in two of the four domains: (a) comprehension and (b) alphabetics. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *LiPS*® on beginning readers. Additional comparisons are presented as supplemental findings. The supplemental findings do not factor into the intervention's rating of effectiveness. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 23.

Summary of effectiveness for the comprehension domain

One study that meets WWC group design standards without reservations reported findings in the comprehension domain.

Torgesen et al. (2010) found a statistically significant positive effect of *LiPS*® on the Woodcock Reading Mastery Test–Revised (WRMT-R) Passage Comprehension subtest when compared to the standard reading instruction alone. The WWC confirmed the finding. The WWC characterizes this study finding as a statistically significant positive effect.

Thus, for the comprehension domain, one study showed statistically significant and substantively important positive effects. This results in a rating of potentially positive effects, with a small extent of evidence.

Table 3. Rating of effectiveness and extent of evidence for the comprehension domain

Rating of effectiveness	Criteria met
Potentially positive effects *Evidence of a positive effect with no overriding contrary evidence.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *comprehension* domain was positive, statistically significant, and substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 74 students in three schools reported evidence of effectiveness in the *comprehension* domain.

Summary of effectiveness for the alphabetics domain

Two studies that meet WWC group design standards without reservations reported findings in the alphabetics domain.

Gunn (1996) found effects of *LiPS*® on word accuracy/fluency and phonemic decoding that were not statistically significant but were negative and substantively important, as measured by the WRMT-R Word Identification and Word Attack subtests. The WWC confirmed these findings. The WWC characterizes these study findings as a substantively important negative effect.

Torgesen et al. (2010) found statistically significant and substantively important positive effects of *LiPS*® on word accuracy/fluency, phonemic decoding, reading comprehension, phonological awareness, and spelling when compared to the basal reading program alone. These outcomes were drawn from a number of assessments, including the WRMT-R Word Identification and Word Attack subtests, the Comprehensive Test of Phonological Processes (CTOPP) Phoneme Elision and Segmenting Words subtests, a developmental spelling analysis (Tangel & Blachman, 1992), and the Test of Word Reading Efficiency (TOWRE) Phonemic Decoding and Word Efficiency subtests. The WWC confirmed the statistical significance of these findings after adjusting for multiple comparisons. The WWC characterizes these study findings as a statistically significant positive effect.

Thus, for the alphabetics domain, one study showed a statistically significant positive effect and one study showed a substantively important negative effect. This results in a rating of mixed effects, with a small extent of evidence.

Table 4. Rating of effectiveness and extent of evidence for the alphabetics domain

Rating of effectiveness	Criteria met
Mixed effects *Evidence of inconsistent effects.*	In the two studies that reported findings, the estimated impact of the intervention on outcomes in the *alphabetics* domain was positive and statistically significant in one study, and negative and substantively important in another study.
Extent of evidence	**Criteria met**
Small	Two studies that included 97 students in five schools reported evidence of effectiveness in the *alphabetics* domain.

References

Studies that meet WWC group design standards without reservations

Gunn, B. K. (1996). *An investigation of three approaches to teaching phonological awareness to first-grade students and the effects on word recognition* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9706736)

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia, 60(*1), 40–56.

Additional source:

Torgesen, J., Wagner, R., Rashotte, C., & Herron, J. (2003). *Summary of outcomes from first grade study with Read, Write, and Type and Auditory Discrimination in Depth instruction and software with at-risk children* (FCRR Tech. Rep. No. 2). Tallahassee, FL: Florida Center for Reading Research. Retrieved from: http://www.fcrr.org

Studies that meet WWC group design standards with reservations

None.

Studies that do not meet WWC group design standards

Howard, M. P. (1986a). *Effects of pre-reading training in auditory conceptualization on subsequent reading achievement* (Study: Arco, Indiana, first-grade longitudinal) (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8612677) The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

Howard, M. P. (1986b). *Effects of pre-reading training in auditory conceptualization on subsequent reading achievement* (Study: Arco, Indiana, kindergarten) (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8612677) The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

Howard, M. P. (1986c). *Effects of pre-reading training in auditory conceptualization on subsequent reading achievement* (Study: Arco, Indiana, and Santa Maria, California) (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 8612677) The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

McGuinness, C., McGuinness, D., & Donohue, J. (1995). Phonological training and the alphabet principle: Evidence for reciprocal causality. *Reading Research Quarterly*, *30*(4), 830–852. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

Sadoski, M., & Willson, V. L. (2006). Effects of a theoretically based large-scale reading intervention in a multicultural urban school district. *American Educational Research Journal*, *43*(1), 137–154. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

Van der Laan, A. R. (2006). *The effects of the Lindamood Phoneme Sequencing Program on reading fluency and comprehension of at-risk first-graders* (Unpublished doctoral dissertation). Eastern Michigan University, Ypsilanti. The study does not meet WWC group design standards because the measures of effectiveness cannot be attributed solely to the intervention.

Studies that are ineligible for review using the Beginning Reading Evidence Review Protocol

Aaron, P. G., Joshi, R. M., Gooden, R., & Bentum, K. E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading. *Journal of Learning Disabilities, 41*(1), 67–84. The study is ineligible for review because it is out of scope of the protocol.

Adair, J., Nadeau, S., Conway, T., Gonzalez-Rothi, L., Heilman, P., Green, I., & Heilman, K. (2000). Alterations in the functional anatomy of reading induced by rehabilitation of an alexic patient. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 13*(4), 303–311. The study is ineligible for review because it does not use a sample aligned with the protocol.

Alexander, A., Anderson, H., Heilman, P., Voeller, K., & Torgesen, J. (1991). Phonological awareness training and the remediation of analytic decoding deficits in a group of severe dyslexics. *Annals of Dyslexia, 41*, 193–206. The study is ineligible for review because it does not use a sample aligned with the protocol.

Burke, C., Lisbeth, H., & Evangelou, T. (2005). *A project of hope: Lindamood-Bell center in a school project final evaluation report*. San Diego, CA: San Diego Association of Governments. The study is ineligible for review because it is out of scope of the protocol.

Castiglioni-Spalten, M., & Ehri, L. C. (2003). Phonemic awareness instruction: Contribution of articulatory segmentation to novice beginners' reading and spelling. *Scientific Studies of Reading, 7*(1), 25–52. The study is ineligible for review because it is out of scope of the protocol.

Colón, E. (2005). *Utility of the Lindamood Phoneme Sequencing Program (LiPS) for classroom-based reading instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3204371) The study is ineligible for review because it does not use an eligible design.

Conway, T., Heilman, P., Gonzalez-Rothi, L., Alexander, A., Adair, J., Crosson, B., & Heilman, K. (1998). Treatment of a case of phonological alexia with agraphia using the Auditory Discrimination in Depth (ADD) program. *Journal of the International Neuropsychological Society, 4*, 608–620. The study is ineligible for review because it does not use a sample aligned with the protocol.

Kennedy, K., & Backman, J. (1993). Effectiveness of the Lindamood Auditory Discrimination in Depth program with students with learning disabilities. *Learning Disabilities Research and Practice*, *8*(4), 253–259. The study is ineligible for review because it does not use a sample aligned with the protocol.

Kutrumbos, B. (1993). *The effect of phonemic training on unskilled readers: A school-based study* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9333368) The study is ineligible for review because it is out of scope of the protocol.

Lindamood-Bell Learning Processes. (n.d.). *2004 clinical statistics*. San Luis Obispo, CA: Author. Retrieved from http://www.lindamoodbell.com The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (n.d.). *2005 clinical statistics*. San Luis Obispo, CA: Author. Retrieved from http://www.lindamoodbell.com The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (n.d.). *2006 learning centers' results*. San Luis Obispo, CA: Author. Retrieved from http://www.lindamoodbell.com The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2003a). *Lindamood-Bell Learning Processes: Beginning reading submissions (Study: Intervention in kindergarten through 2nd grade)*. San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2003b). *Lindamood-Bell Learning Processes: Beginning reading submissions (Study: Kindergarten results from school project in Oregon)*. San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2003c). *Lindamood-Bell Learning Processes: Beginning reading submissions (Study: Kindergarten through 3rd grade results from learning centers across the United States)*. San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2003d). *Lindamood-Bell Learning Processes: Beginning reading submissions (Study: Kindergarten through 3rd grade results from school project in Colorado*). San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2004a). *Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II (Study: K–3 Lindamood-Bell focus students 2002 summary*). San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2004b). *Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1*, *Book I of II (Study: Kindergarten students in Oregon 2001–02)*. San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2004c). *Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II (Study: Longitudinal Florida study summary)*. San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2004d). *Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II (Study: Pueblo, Colorado, 2001–02 summary)*. San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

Lindamood-Bell Learning Processes. (2004e). *Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II (Study: Second grade students in Idaho)*. San Luis Obispo, CA: Author. The study is ineligible for review because it does not use an eligible design.

McBride, N. (2005). *The effectiveness of Second Shot and/or Lindamood-Bell on reading achievement of elementary students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3209120) The study is ineligible for review because it does not use a sample aligned with the protocol.

McCollom, P. (2000). *Effects of phonemic awareness training and the influence of phonemic skills on early reading success with pre-school and kindergarten children* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3030634) The study is ineligible for review because it is out of scope of the protocol.

McIntyre, L., Protz, S., & McQuarrie, L. (2008). Exploring the potential of LiPS instruction for beginning readers. *Developmental Disabilities Bulletin, 36*(1-2), 18–48. The study is ineligible for review because it does not use an eligible design.

Olson, R. K., & Wise, B. (2006). Computer-based remediation for reading and related phonological disabilities. In M. C. McKenna, L. D. Labbo, R. D. Kiefer, & D. Reinking (Eds.), *International handbook of literacy and technology* (pp. 57–74). Mahwah, NJ: Erlbaum. The study is ineligible for review because it does not use an eligible design.

Olson, R. K., Wise, B. W., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the post-training development of word recognition. *Scientific Studies of Reading*, *1*(3), 235–253. The study is ineligible for review because it does not use a sample aligned with the protocol.

Owen, K. E. (2004). *Effects of Lindamood-Bell on third and fourth grade reading achievement in Pueblo School District No. 60* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3134416) The study is ineligible for review because it is out of scope of the protocol.

Pokorni, J. L., Worthington, C. K., & Jamison, P. J. (2004). Phonological awareness intervention: Comparison of Fast ForWord, Earobics, and LiPS. *The Journal of Educational Research*, *97*(3), 147–157. The study is ineligible for review because it does not use a sample aligned with the protocol.

Shaw, D. M., & Disney, L. (2012). Expanding access, knowledge, and participation for learning disabled young adults with low literacy. *Journal of Research and Practice for Adult Literacy, Secondary, and Basic Education*, *1*(3), 148–160. The study is ineligible for review because it does not use a sample aligned with the protocol.

Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M., … Papanicolaou, A. C. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58,* 1203–1212. The study is ineligible for review because it is out of scope of the protocol.

Swanson, T. (1996). *Metaphonological instruction for poor readers in upper grades: An investigation of treatment effectiveness* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9632544) The study is ineligible for review because it does not use a sample aligned with the protocol.

Torgerson, C., Brooks, G., & Hall, J. (2006). *A systematic review of the research literature on the use of phonics in the teaching of reading and spelling* (Research Report No. RR711). South Yorkshire, UK: University of Sheffield, Department for Education Skills Publications. The study is ineligible for review because it does not use an eligible design.

Torgesen, J. K., Alexander, P. A., Wagner, R. K., Rashotte, C. A., Voeller, K. K. S., Conway, T., & Rose, E. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34*(1), 33–58. The study is ineligible for review because it does not use a sample aligned with the protocol.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*(4), 579–593. The study is ineligible for review because it is out of scope of the protocol.

Truch, S. (1994). Stimulating basic reading processes using Auditory Discrimination in Depth. *Annals of Dyslexia, 44*, 60–80. The study is ineligible for review because it does not use an eligible design.

Truch, S. (2004). *Remedial outcomes with different reading programs*. San Diego, CA: International Dyslexia Association. The study is ineligible for review because it does not use an eligible design.

Vanderberg, L., Pierce, M., & Disney, L. (2011). Reading intervention outcomes for adults with disabilities in a vocational rehabilitation setting: Results of a 3-year research and demonstration grant. *Rehabilitation Counseling Bulletin, 54*(4), 210–222. The study is ineligible for review because it does not use a sample aligned with the protocol.

Wise, B. W., Ring, J., and Olson, R. K. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology, 72*(4), 271–304. The study is ineligible for review because it is out of scope of the protocol.

Wise, B. W., Ring, J., & Olson, R. K. (2000). Individual differences in gains from computer-assisted remedial reading. *Journal of Experimental Child Psychology, 77*(3), 197–235. The study is ineligible for review because it is out of scope of the protocol.

Zyburt, G. (2010). *Comprehensive school reform with a focus on literacy* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3414047) The study is ineligible for review because it does not use an eligible design.

Appendix A.1: Research details for Gunn, 1996

Gunn, B. K. (1996). *An investigation of three approaches to teaching phonological awareness to first-grade students and the effects on word recognition* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 9706736)

Table A1. Summary of findings **Meets WWC group design standards without reservations**

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Alphabetics	23 students	–13	No

Setting

The study included two elementary schools in one Pacific Northwest school district.

Study sample

Thirty-five first-grade students from five classrooms in two schools who were identified as having low phonological awareness skills based on performance on the Test of Phonological Awareness (TOPA) participated in the study. The students were organized into six groups, and these groups were randomly assigned to one of three conditions: *Complete Auditory Discrimination in Depth* (*CADD*, which is now known as *LiPS*®, included 11 students), the basal reading program (*BASAL*, which included 13 students), or *Modified Auditory Discrimination in Depth* (*MADD*, which included 11 students). The *MADD* condition is not eligible for review, as it is a modified version of the *LiPS*® program. Therefore, the 11 students in the MADD condition are excluded from this review.

The *CADD* group's analytic sample included five male and six female students, with a mean age of 6.35 years. The *BASAL* (comparison) group's analytic sample included three male and nine female students, with a mean age of 6.47 years. All students who participated in the study scored at or below the 15th percentile on the TOPA, which was used as a screening test and administered prior to randomization.

Intervention group

The *LiPS*® program is designed to teach students the skills they need to decode and encode words and to identify individual sounds and blends in words. For this study, instruction focused on five components of the *CADD* program: (a) setting the climate for learning, (b) identifying and classifying speech sounds, (c) tracking speech sounds, (d) spelling, and (e) reading.

As a supplement to regular classroom reading instruction, *CADD* instruction was delivered to small groups of five to six students. Instruction was provided for 30 minutes every day for 8 weeks (40 sessions total). During this period, all consonant pairs specified in *CADD* guidelines were introduced, but there was insufficient time to introduce consonant groups and some vowel sounds.

Comparison group

Students in the comparison (*BASAL*) group received supplemental reading instruction that incorporated phonological awareness activities from the basal reader adopted for use in the regular classroom. The classroom reading basal at School A was the *Scribner Reading Series*, *Join the Circle* (Scribner, 1997); the classroom reading basal at School B was the *Open Court Series, Blue Pillow Sky* (Open Court, 1989). The supplemental instruction was provided to students in a small room near the first-grade classroom for approximately 30 minutes each day for 40 days (8 weeks).

Outcomes and measurement

Two outcomes in the alphabetics domain met review requirements. These outcomes were the WRMT-R Word Identification subtest and the Word Attack subtest. For a more detailed description of these outcome measures, see Appendix B.

Supplemental findings are presented in Appendix D on WRMT-R results for a 4-week delayed posttest and an 8-week delayed posttest. The supplemental findings do not factor into the intervention's rating of effectiveness.

Outcomes in two domains (alphabetics and reading fluency) did not meet review requirements. The Lindamood Auditory Conceptualization Test (LACT) in the alphabetics domain was over-aligned with the intervention. The Curriculum-Based Oral Reading Fluency Test in the reading fluency domain was developed by the study authors. No reliability information was provided for this outcome, so it did not meet review requirements because the reliability of the test could not be established.

Support for implementation

The two teachers providing instruction for the *CADD*, *MADD*, and *BASAL* programs each received 18 hours of pre-service training on the published formats and guidelines for each program. The two teachers also met with the investigator for 2 hours each week for 8 weeks for additional pre-service training. Daily lesson plans developed by the investigator were also provided to the two teachers.

Appendix A.2: Research details for Torgesen et al., 2010

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Herron, J., & Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of Dyslexia, 60*(1), 40–56.

Table A2. Summary of findings **Meets WWC group design standards without reservations**

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Comprehension	74 students	+20	Yes
Alphabetics	74 students	+24	Yes

Setting

The study included students from three elementary schools.

Study sample

First-grade students were identified as potentially at risk of having difficulty reading using a two-stage process. First, a pool of potential candidates was identified based on low scores (bottom 35%) on a test of letter-sound knowledge. Second, study authors computed a probability of reading difficulty for each student, using logistic regression and based on a combined score from three tests that measured phoneme elision, serial naming of numbers, and vocabulary. Students with the highest probabilities of reading difficulty were eligible for inclusion in the study.

In total, 112 students potentially at risk of reading difficulty were recruited to participate in the study over 2 consecutive school years. Across these 2 years, 36 students were randomly assigned to the *LiPS*® intervention group, 36 students were randomly assigned to another

intervention (*RWT*), and 40 students were randomly assigned to the comparison group. The final study sample, after attrition, included 35 students in the *LiPS*® group, 34 students in the *RWT* group, and 39 students in the comparison group. The *RWT* condition does not factor into the intervention's rating of effectiveness, as the comparison group's use of a basal reader provided a more appropriate counterfactual to test the effectiveness of *LiPS*®; however, *LiPS*® vs. *RWT* contrasts are presented as supplemental findings in Appendix D. These supplemental findings in the comprehension, alphabetics, and reading fluency domains contrast an oral language approach used in the *LiPS*® intervention with an approach focused more heavily on spelling and writing in *RWT*. The supplemental findings do not factor into the intervention's rating of effectiveness.

About 56% of the total sample were male, 33% were minority (mostly African American), and about 35% received free or reduced-price lunch. The average age at the beginning of instruction was 6.5 years.

Intervention group

The *LiPS*® program is designed to teach students the skills they need to decode and encode words and to identify individual sounds and blends in words. For this study, as a supplement to regular classroom reading instruction, students were instructed in groups of three, and received four 50-minute sessions per week throughout the school year (i.e., from October through May). On average, students received 84.5 hours of *LiPS*® instruction.

Comparison group

Students in the comparison group did not receive any supplemental reading instruction. In two of the schools, the standard reading instruction was *Open Court's Collections for Young Scholars*. The third school did not have a standard reading curriculum, but instead allowed teachers to choose their materials for reading instruction.

Outcomes and measurement

Assessments were administered immediately following the delivery of the interventions in May of a given school year. Outcomes in the alphabetics domain were measured using the WRMT-R Word Attack and Word Identification subtests; the TOWRE Word Efficiency and Phonemic Decoding Efficiency subtests; the CTOPP Blending Words, Segmenting Words, Phoneme Elision, and Rapid Letter Naming subtests; and a developmental spelling analysis (Tangel & Blachman, 1992). Outcomes in the comprehension domain were measured using the WRMT-R Passage Comprehension subtest. The CTOPP Rapid Digit Naming subtest was excluded from this review, since it was out of scope of the Beginning Reading Protocol.

Outcomes were also measured 1 year following the delivery of the intervention. Reading fluency was measured using the Gray Oral Reading Test–Third Edition (GORT-3) Text Reading Rate subtest. Alphabetics was measured using the WRMT-R, CTOPP, and the Wide Range Achievement Test–Revised (WRAT-R) Spelling subtest. Comprehension was measured using the WRMT-R Passage Comprehension subtest and the GORT-3 Comprehension subtest. These 1-year follow-up assessments are presented as supplemental findings in Appendix D. The supplemental findings do not factor into the intervention's rating of effectiveness.

Support for implementation

Teachers received 18 hours of pre-service training in *LiPS*® at the beginning of each year. Biweekly 3-hour staff meetings were held with teachers to discuss instructional or behavioral issues in their classrooms. Supervisors with special expertise in the *LiPS*® program attended roughly half of these staff meetings.

Appendix B: Outcome measures for each domain

Comprehension	
Gray Oral Reading Test–Third Edition (GORT-3) Comprehension subtest	This assessment required students to answer five multiple-choice comprehension questions after reading each paragraph. Paragraphs increased in difficulty as the assessment proceeded (as cited in Torgesen et al., 2010). The assessment was administered 1 year following the end of the intervention.
Woodcock Reading Mastery Test–Revised (WRMT-R) Passage Comprehension subtest	This assessment required students to silently read a paragraph and supply a key missing word (as cited in Torgesen et al., 2010). The assessment was administered immediately following the delivery of the intervention (in May of a given school year) and again 1 year following the end of the intervention.
Alphabetics	
Letter identification	
Comprehensive Test of Phonological Processes (CTOPP) Rapid Letter Naming subtest	This assessment required students to name lowercase letters as quickly as possible (as cited in Torgesen et al., 2010). The assessment was administered immediately following the delivery of the intervention (in May of a given school year) and again 1 year following the end of the intervention.
Phonics	
Developmental Spelling Analysis	This assessment measured accuracy of phonemic representations in spelling (Tangel & Blachman, 1992) and was administered immediately following the delivery of the intervention (in May of a given school year). Inter-rater reliability for this assessment was measured at .98 (as cited in Torgesen et al., 2010).
Test of Word Reading Efficiency (TOWRE) Phonemic Decoding Efficiency subtest	This assessment required students to read as many nonwords as possible in 45 seconds (as cited in Torgesen et al., 2010). The assessment was administered immediately following the delivery of the intervention (in May of a given school year) and again 1 year following the end of the intervention.
TOWRE Word Efficiency subtest	This assessment required students to read as many words as accurately as possible within 45 seconds (as cited in Torgesen et al., 2010). The assessment was administered immediately following the delivery of the intervention (in May of a given school year).
Wide Range Achievement Test–Revised (WRAT-R) Spelling subtest	This assessment (Jastak & Jastak, 1978) measured a student's spelling skills (as cited in Torgesen et al., 2010). The assessment was administered 1 year following the end of the intervention.
WRMT-R Word Attack subtest	This assessment required participants to read either nonsense words or words that occur infrequently in everyday usage to measure students' ability to apply phonic and structural analysis skills. Internal consistency was measured as greater than .80, and inter-rater reliability was .95 (as cited in Gunn [1996] and Torgesen et al. [2010]). The assessment was administered immediately following the delivery of the intervention (in the final 2 days of the intervention in the Gunn [1996] study and in May of a given school year in the Torgesen et al. [2010] study). This assessment was also administered 1 year following the end of the intervention in the Torgesen et al. (2010) study.
WRMT-R Word Identification subtest	This assessment required participants to identify words that appear in isolation. Internal consistency was measured as greater than .80, and inter-rater reliability was .95 (as cited in Gunn [1996] and Torgesen et al. [2010]). The assessment was administered immediately following the delivery of the intervention (in the final 2 days of the intervention in the Gunn [1996] study and in May of a given school year in the Torgesen et al. [2010] study). This assessment was also administered 1 year following the end of the intervention in the Torgesen et al. (2010) study.
Phonological awareness	
CTOPP Blending Words subtest	This 29-item assessment required students to put a series of sounds together to make a word (as cited in Torgesen et al., 2010). The assessment was administered immediately following the delivery of the intervention (in May of a given school year) and again 1 year following the end of the intervention.
CTOPP Phoneme Elision subtest	This 25-item assessment required students to indicate the remainder of a word after part is omitted (as cited in Torgesen et al., 2010). The assessment was administered immediately following the delivery of the intervention (in May of a given school year) and again 1 year following the end of the intervention.
CTOPP Segmenting Words subtest	This 26-item assessment required participants to read aloud a word and then repeat it one sound at a time (as cited in Torgesen et al., 2010). The assessment was administered immediately following the delivery of the intervention (in May of a given school year) and again 1 year following the end of the intervention.

Reading fluency	
GORT-3 Text Reading Accuracy subtest	This assessment required students to read a series of paragraphs that gradually increased in difficulty level. Reading accuracy was assessed by recording the number of errors that occurred at each level of difficulty (as cited in Torgesen et al., 2010). The assessment was administered 1 year following the end of the intervention. Results from this assessment are presented as supplemental findings in Appendix D.
GORT-3 Text Reading Rate subtest	This assessment required students to read a series of paragraphs that gradually increased in difficulty level. Reading rate was measured by the average amount of time it took to read each paragraph (as cited in Torgesen et al., 2010). The assessment was administered 1 year following the end of the intervention. Results from this assessment are presented as supplemental findings in Appendix D.

Appendix C.1: Findings included in the rating for the comprehension domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Torgesen et al., 2010[a]								
Woodcock Reading Mastery Test–Revised (WRMT-R)	Grade 1	3 schools/ 74 students	102.20 (10.00)	95.40 (14.40)	6.80	0.54	+20	.02
Domain average for comprehension (Torgesen et al., 2010)						**0.54**	**+20**	**Statistically significant**
Domain average for comprehension across all studies						**0.54**	**+20**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. na = not applicable.

[a] For Torgesenet al. (2010), the WWC did not need to make corrections for clustering, multiple comparisons, or to adjust for baseline differences. Study authors combined the *LiPS*® and *RWT* samples in their analysis and presented significance levels and effect sizes for these two intervention groups relative to the comparison group. Effect sizes and significance levels reported in this table, which are designed to reflect the effects of *LiPS*® relative to the comparison group, were calculated by the WWC. WRMT-R outcomes are reported at post-test only, as a pretest of reading comprehension was not administered. This study is characterized as having a statistically significant positive effect because the estimated effect is positive and statistically significant. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

Appendix C.2: Findings included in the rating for the alphabetics domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Gunn, 1996[a]								
Woodcock Reading Mastery Test–Revised (WRMT-R) Word Attack subtest	Grade 1	2 schools/ 23 students	2.82 (2.99)	4.25 (4.58)	–1.43	–0.35	–14	.57
WRMT-R Word Identification subtest	Grade 1	2 schools/ 23 students	6.36 (6.48)	8.75 (8.37)	–2.39	–0.31	–12	.62
Domain average for alphabetics (Gunn, 1996)						**–0.33**	**–13**	**Not statistically significant**
Torgesen et al., 2010[b]								
Comprehensive Test of Phonological Processes (CTOPP) Blending Words subtest	Grade 1	3 schools/ 74 students	20.60 (4.50)	18.20 (5.40)	2.40	0.48	+18	.05
CTOPP Phoneme Elision subtest	Grade 1	3 schools/ 74 students	15.50 (4.20)	12.50 (4.60)	3.00	0.67	+25	.01
CTOPP Rapid Letter Naming subtest	Grade 1	3 schools/ 74 students	1.20 (0.30)	1.20 (0.30)	0.0	0	0	> .99
CTOPP Segmenting Words subtest	Grade 1	3 schools/ 74 students	15.60 (3.70)	11.70 (4.50)	3.90	0.93	+32	< .01
Developmental Spelling Analysis	Grade 1	3 schools/ 74 students	25.10 (2.70)	23.40 (3.20)	1.70	0.57	+21	.03

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Test of Word Reading Efficiency (TOWRE) Phonemic Decoding Efficiency subtest	Grade 1	3 schools/ 74 students	16.80 (7.60)	10.60 (7.70)	6.20	0.80	+29	< .01
TOWRE Word Efficiency subtest	Grade 1	3 schools/ 74 students	26.90 (11.10)	21.00 (11.40)	5.90	0.52	+20	.03
WRMT-R Word Attack subtest	Grade 1	3 schools/ 74 students	113.70 (12.10)	99.50 (15.00)	14.20	1.03	+35	< .01
WRMT-R Word Identification subtest	Grade 1	3 schools/ 74 students	110.60 (12.20)	100.60 (15.60)	10.00	0.70	+26	< .01
Domain average for alphabetics (Torgesen et al., 2010)						**0.63**	**+24**	**Statistically significant**
Domain average for alphabetics across all studies						**0.15**	**+6**	**na**

Table Notes: For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of each study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. na = not applicable.

[a] For Gunn (1996), effect sizes are calculated based on posttest-only means, since a corresponding pretest was not conducted. The author combined the *CADD* and *MADD* samples in significance tests. Effect sizes and significance levels reported in this table, which are designed to reflect the effects of *CADD* relative to the comparison group, were calculated by the WWC. For Gunn (1996), no corrections for multiple comparisons were needed. A clustering correction was needed but did not affect whether any of the contrasts were found to be statistically significant. This study is characterized as having a substantively important negative effect because the mean effect within the domain is negative and not statistically significant but is substantively important. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

[b] For Torgesen et al. (2010) the WWC calculated the program group mean for the CTOPP Phoneme Elision subtest using a difference-in-differences approach by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information. All other outcomes are reported using posttest-only means, since a corresponding pretest was not conducted. The study authors combined the *LiPS*® and *RWT* samples in their analysis and presented significance levels and effect sizes for these two intervention groups relative to the comparison group. Effect sizes and significance levels reported in this table, which are designed to reflect the effects of *LiPS*® relative to the comparison group, were calculated by the WWC. For Torgesen et al. (2010), a correction for multiple comparisons was needed and resulted in a WWC-computed critical *p*-value of .044 for the CTOPP Blending Words subtest; therefore, the WWC does not find the result to be statistically significant. All other outcomes that were originally found by the authors to be statistically significant remained so. This study is characterized as having a statistically significant positive effect because at least half of the effects are positive and statistically significant, and no effects are negative and statistically significant. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

Appendix D.1: Description of supplemental findings for the comprehension domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Torgesen et al., 2010[a]								
LiPS® vs. Read, Write, and Type (RWT): Immediate posttest								
Woodcock Reading Mastery Test–Revised (WRMT-R) Passage Comprehension subtest	Grade 1	3 schools/ 69 students	102.20 (10.00)	100.20 (9.60)	2.00	0.20	+8	.41
LiPS® vs. Comparison: 1-year follow-up								
Gray Oral Reading Test–Third Edition (GORT-3) Comprehension subtest	Grade 1	3 schools/ 74 students	99.20 (14.50)	95.60 (13.80)	3.60	0.25	+10	.28
WRMT-R Passage Comprehension subtest	Grade 1	3 schools/ 74 students	98.90 (8.50)	93.70 (12.60)	5.20	0.47	+18	.05
LiPS® vs. RWT: 1-year follow-up								
GORT-3 Comprehension subtest	Grade 1	3 schools/ 69 students	99.20 (14.50)	96.40 (11.80)	2.80	0.21	+8	.39
WRMT-R Passage Comprehension subtest	Grade 1	3 schools/ 69 students	98.90 (8.50)	96.70 (7.60)	2.20	0.27	+11	.27

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding.

[a] In Torgesen et al. (2010), study authors combined the *LiPS®* and *RWT* samples in their analysis and presented significance levels and effect sizes for these two intervention groups relative to the comparison group. Effect sizes and significance levels reported in this table, which are designed to reflect the effects of *LiPS®* relative to the *RWT* group or comparison group, were calculated by the WWC. Effect sizes for the WRMT-R and GORT-3 outcomes are calculated based on posttest-only means, since a corresponding pretest was not conducted. For Torgesen et al. (2010), a correction for multiple comparisons was needed for the 1-year follow-up of *LiPS®* vs. comparison contrast and resulted in a WWC-computed critical *p*-value of .025 for the WRMT-R Passage Comprehension subtest (vs. comparison), 1-year follow-up; therefore, the WWC does not find the result to be statistically significant.

Appendix D.2: Description of supplemental findings for the alphabetics domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Gunn, 1996[a]								
4-week delayed posttest								
Woodcock Reading Mastery Test–Revised (WRMT-R) Word Attack subtest	Grade 1	2 schools/ 23 students	4.91 (4.74)	4.75 (4.25)	0.16	0.03	+1	.96
WRMT-R Word Identification subtest	Grade 1	2 schools/ 23 students	9.82 (7.99)	10.75 (8.35)	–0.93	–0.11	–4	.86

Outcome measure	Study sample	Sample size	Mean (standard deviation) Intervention group	Mean (standard deviation) Comparison group	WWC calculations Mean difference	WWC calculations Effect size	WWC calculations Improvement index	*p*-value
8-week delayed posttest								
WRMT-R Word Attack subtest	Grade 1	2 schools/ 23 students	4.64 (2.94)	13.40 (9.81)	–8.76	–1.14	–37	.08
WRMT-R Word Identification subtest	Grade 1	2 schools/ 23 students	10.70 (7.36)	12.70 (15.20)	–2.00	–0.16	–6	.80
Torgesen et al., 2010[b]								
LiPS® vs. Read, Write, and Type (RWT): Immediate posttest								
Comprehensive Test of Phonological Processes (CTOPP) Blending Words subtest	Grade 1	3 schools/ 69 students	20.60 (4.50)	22.00 (4.00)	–1.40	–0.37	–14	.13
CTOPP Phoneme Elision subtest	Grade 1	3 schools/ 69 students	15.10 (4.20)	13.80 (4.20)	1.30	0.47	+18	.06
CTOPP Rapid Letter Naming subtest	Grade 1	3 schools/ 69 students	1.20 (0.30)	1.30 (0.30)	–0.10	–0.33	–13	.18
CTOPP Segmenting Words subtest	Grade 1	3 schools/ 69 students	16.20 (3.70)	14.60 (4.60)	1.60	0.38	+15	.12
Developmental Spelling subtest	Grade 1	3 schools/ 69 students	25.10 (2.70)	25.00 (2.60)	0.10	0.04	+1	.88
Test of Word Reading Efficiency (TOWRE) Phonemic Decoding Efficiency subtest	Grade 1	3 schools/ 69 students	16.80 (7.60)	12.60 (7.00)	4.20	0.57	+21	.02
TOWRE Word Efficiency subtest	Grade 1	3 schools/ 69 students	26.70 (11.10)	23.50 (9.30)	3.20	0.31	+12	.20
WRMT-R Word Attack subtest	Grade 1	3 schools/ 69 students	116.50 (12.10)	108.30 (12.20)	8.20	0.67	+25	< .01
WRMT-R Word Identification subtest	Grade 1	3 schools/ 69 students	109.00 (12.20)	107.00 (12.40)	2.00	0.16	+6	.51
LiPS® vs. Comparison: 1-year follow-up								
CTOPP Blending Words subtest	Grade 1	3 schools/ 74 students	22.70 (4.20)	21.60 (5.40)	1.10	0.22	+9	.34
CTOPP Phoneme Elision subtest	Grade 1	3 schools/ 74 students	17.60 (4.80)	15.70 (4.40)	1.90	0.41	+16	.08
CTOPP Rapid Letter Naming subtest	Grade 1	3 schools/ 74 students	1.70 (0.30)	1.50 (0.30)	0.20	0.66	+25	.01
CTOPP Rapid Segmenting Words subtest	Grade 1	3 schools/ 74 students	16.10 (3.90)	14.20 (3.50)	1.90	0.51	+19	.03
TOWRE Phonemic Decoding Efficiency subtest	Grade 1	3 schools/ 74 students	26.10 (9.00)	20.20 (12.80)	5.90	0.52	+20	.03
TOWRE Word Efficiency subtest	Grade 1	3 schools/ 74 students	44.30 (12.00)	38.60 (14.40)	5.70	0.42	+16	.07
WRAT-R Spelling subtest	Grade 1	3 schools/ 74 students	37.60 (4.40)	34.90 (4.60)	2.70	0.59	+22	.01

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
WRMT-R Word Attack subtest	Grade 1	3 schools/ 74 students	112.50 (15.70)	99.60 (20.40)	12.90	0.70	+26	< .01
WRMT-R Word Identification subtest	Grade 1	3 schools/ 74 students	106.80 (12.80)	99.80 (14.80)	7.00	0.50	+19	.04
LiPS® vs. Comparison: 1-year follow-up								
CTOPP Blending Words subtest	Grade 1	3 schools/ 69 students	22.50 (4.20)	22.00 (4.00)	0.50	0.12	+5	.62
CTOPP Phoneme Elision subtest	Grade 1	3 schools/ 69 students	17.90 (4.80)	16.50 (4.40)	1.40	0.30	+12	.22
CTOPP Rapid Letter Naming subtest	Grade 1	3 schools/ 69 students	1.70 (0.30)	1.70 (0.30)	0.00	0.00	0	> .99
CTOPP Segmenting Words subtest	Grade 1	3 schools/ 69 students	16.70 (3.90)	14.60 (4.60)	2.10	0.49	+19	.05
TOWRE Phonemic Decoding Efficiency subtest	Grade 1	3 schools/ 69 students	26.10 (9.00)	22.60 (8.40)	3.50	0.40	+15	.10
TOWRE Word Efficiency subtest	Grade 1	3 schools/ 69 students	44.10 (12.00)	42.70 (10.90)	1.40	0.12	+5	.62
WRAT-R Spelling subtest	Grade 1	3 schools/ 69 students	37.60 (4.40)	36.20 (3.20)	1.40	0.36	+14	.14
WRMT-R Word Attack subtest	Grade 1	3 schools/ 69 students	115.30 (15.70)	104.40 (11.90)	10.90	0.77	+28	< .01
WRMT-R Word Identification subtest	Grade 1	3 schools/ 69 students	105.20 (12.80)	103.80 (11.00)	1.40	0.12	+5	.63

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding.

[a] Effect sizes from Gunn (1996) are calculated based on posttest-only means, since a corresponding pretest was not conducted. The author combined the *CADD* and *MADD* samples in significance tests. Effect sizes and significance levels reported in this table, which are designed to reflect the effects of *CADD* relative to the comparison group, were calculated by the WWC. For Gunn (1996), no corrections for multiple comparisons were needed. A correction for clustering was needed and resulted in a WWC-computed *p*-value of .08 for the WRMT-R Word Attack 8-week posttest; therefore, the WWC does not find the result to be statistically significant.

[b] In Torgesen et al. (2010), study authors combined the *LiPS®* and *RWT* samples in their analysis and presented significance levels and effect sizes for these two intervention groups relative to the comparison group. Effect sizes and significance levels reported in this table, which are designed to reflect the effects of *LiPS®* relative to the *RWT* group or comparison group, were calculated by the WWC. For the *LiPS®* vs. *RWT* contrast at posttest, the WWC used a difference-in-differences approach to calculate the program group mean for all outcomes except for the CTOPP Rapid Letter Naming subtest and the Developmental Spelling Analysis by adding the impact of the program (i.e., difference in mean gains between the intervention and comparison groups) to the unadjusted comparison group posttest means. For the *LiPS®* vs. comparison contrast at 1-year follow-up, the WWC calculated the program group mean for only the CTOPP Phoneme Elision subtest using a difference-in-differences approach. For the *LiPS®* vs. *RWT* contrast at 1-year follow-up, the WWC used a difference-in-differences approach to calculate the program group mean for all outcomes except for the CTOPP Rapid Letter Naming subtest and the WRAT-R Spelling subtest. Please see the WWC Procedures and Standards Handbook (version 3.0) for more information. All other outcomes are reported using posttest-only means, since a corresponding pretest was not conducted. A correction for multiple comparisons was needed. After correcting for multiple comparisons, the following outcomes were found not to be statistically significant: WRMT-R Word Attack subtest (vs. *RWT*, posttest; WWC-computed critical *p*-value of .006), TOWRE Phonemic Decoding Efficiency subtest (vs. *RWT*, posttest; WWC-computed critical *p*-value of .011), CTOPP: Segmenting Words subtest (vs. comparison, 1-year follow-up; WWC-computed critical *p*-value of .028), TOWRE Phonemic Decoding Efficiency subtest (vs. comparison, 1-year follow-up; WWC-computed critical *p*-value of .022), WRMT Word Identification subtest (vs. comparison, 1-year follow-up; WWC-computed critical *p*-value of .033), and CTOPP Segmenting Words subtest (vs. *RWT*, 1-year follow-up; WWC-computed critical *p*-value of .011). All other findings reported by the author to be statistically significant remained so.

Appendix D.3: Description of supplemental findings for the reading fluency domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Torgesen et al., 2010[a]								
LiPS® vs. Comparison: 1-year follow-up								
Gray Oral Reading Test–Third Edition (GORT-3) Text Reading Accuracy subtest	Grade 1	3 schools/ 74 students	97.40 (12.80)	92.40 (14.20)	5.00	0.36	+14	.12
GORT-3 Text Reading Rate subtest	Grade 1	3 schools/ 74 students	97.20 (10.70)	92.20 (14.70)	5.00	0.38	+15	.11
LiPS® vs. Read, Write, and Type (RWT): 1-year follow-up								
GORT-3 Text Reading Accuracy subtest	Grade 1	3 schools/ 74 students	97.40 (12.80)	96.80 (11.30)	0.60	0.05	+2	.84
GORT-3 Text Reading Rate subtest	Grade 1	3 schools/ 69 students	97.20 (10.70)	94.70 (9.50)	2.50	0.24	+10	.31

Table Notes: The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding.

[a] In Torgesen et al. (2010), study authors combined the *LiPS®* and *RWT* samples in their analysis and presented significance levels and effect sizes for these two intervention groups relative to the comparison group. Effect sizes and significance levels reported in this table, which are designed to reflect the effects of *LiPS®* relative to the *RWT* group or comparison group, were calculated by the WWC. Effect sizes for the GORT-3 outcomes are calculated based on posttest-only means, since a corresponding pretest was not conducted. No corrections for clustering or multiple comparisons were needed.

Endnotes

[1] The descriptive information for this program was obtained from a publicly available source: the program's website (http://www.lindamoodbell.com, downloaded October 2014). The WWC requests developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in October 2014, and the WWC incorporated feedback from the developer. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.

[2] The literature search reflects documents publicly available by August 2015. The previous intervention report was released in December 2008. This report has been updated to include reviews of 16 studies that have been released since 2008. Of the additional studies, 13 were not within the scope of the review protocol for the Beginning Reading topic area, and one was within the scope of the review protocol but did not meet WWC group design standards. Two studies met WWC group design standards without reservations, and findings from these studies are highlighted in this report. A complete list and disposition of all studies reviewed are provided in the references. The studies in this report were reviewed using the WWC Procedures and Standards Handbook (version 3.0), along with those described in the Beginning Reading review protocol (version 3.0). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[3] For criteria used in the determination of the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 23. These improvement index numbers show the average and range of student-level improvement indices for all findings across the studies.

[4] Cost information for *LiPS*® materials was verified through Lindamood-Bell's publisher, Gander Publishing: http://ganderpublishing.com/product/lips-kit.asp. Cost information for the *LiPS*® training is available on the Lindamood-Bell Workshops Registration Form, posted on the developer's website: http://lindamoodbell.com/wp-content/uploads/2014/10/US-Registration.pdf. All cost information is current as of April 14, 2015.

[5] Supplemental findings are presented for three domains in Appendix D. For the comprehension domain, findings are presented for the following contrasts from Torgesen et al. (2010): 1-year follow-up on the *LiPS*® and comparison conditions, as well as the immediate posttest and 1-year follow-up results for the *LiPS*® and *RWT* conditions. For the alphabetics domain, supplemental findings are presented for Torgesen et al. (2010) on the following contrasts: immediate posttest results for the *LiPS*® and *RWT* conditions, and 1-year follow-up on the *LiPS*®, *RWT*, and comparison conditions. Supplemental findings from 4-week and 8-week delayed posttests are also presented for Gunn (1996). For the reading fluency domain, supplemental findings are presented on the 1-year follow-up on the *LiPS*®, *RWT*, and comparison conditions.

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2015, November). *Beginning Reading intervention report: Lindamood Phoneme Sequencing® (LiPS®)*. Retrieved from http://whatworks.ed.gov

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC group design standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC group design standards with reservations	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
Positive effects	Two or more studies show statistically significant positive effects, at least one of which met WWC group design standards for a strong design, AND No studies show statistically significant or substantively important negative effects.
Potentially positive effects	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
Mixed effects	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows 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, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
Potentially negative effects	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
Negative effects	Two or more studies show statistically significant negative effects, at least one of which met WWC group design standards for a strong design, AND No studies show statistically significant or substantively important positive effects.
No discernible effects	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
Medium to large	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
Small	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

Glossary of Terms

Attrition Attrition occurs when an outcome variable is not available for all participants initially assigned to the intervention and comparison groups. The WWC considers the total attrition rate and the difference in attrition rates across groups within a study.

Clustering adjustment If intervention assignment is made at a cluster level and the analysis is conducted at the student level, the WWC will adjust the statistical significance to account for this mismatch, if necessary.

Confounding factor A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

Design The design of a study is the method by which intervention and comparison groups were assigned.

Domain A domain is a group of closely related outcomes.

Effect size The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

Eligibility A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

Equivalence A demonstration that the analysis sample groups are similar on observed characteristics defined in the review area protocol.

Extent of evidence An indication of how much evidence supports the findings. The criteria for the extent of evidence levels are given in the WWC Rating Criteria on p. 23.

Improvement index Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual starts at the 50th percentile, the measure ranges from –50 to +50.

Intervention An educational program, product, practice, or policy aimed at improving student outcomes.

Intervention report A summary of the findings of the highest-quality research on a given program, product, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against design standards, and summarizes the findings of those that meet WWC design standards.

Multiple comparison adjustment When a study includes multiple outcomes or comparison groups, the WWC will adjust the statistical significance to account for the multiple comparisons, if necessary.

Quasi-experimental design (QED) A quasi-experimental design (QED) is a research design in which study participants are assigned to intervention and comparison groups through a process that is not random.

Randomized controlled trial (RCT) A randomized controlled trial (RCT) is an experiment in which eligible study participants are randomly assigned to intervention and comparison groups.

Rating of effectiveness The WWC rates the effects of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 23.

Single-case design A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.

Glossary of Terms

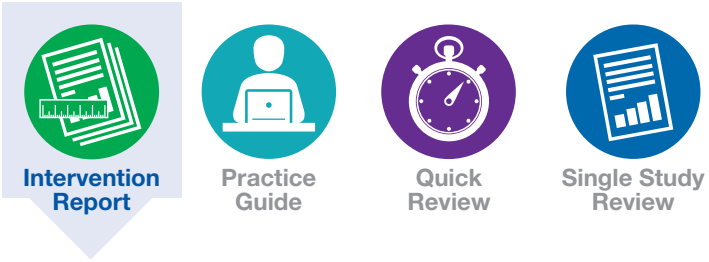
Standard deviation The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.

Statistical significance Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ($p < .05$).

Substantively important A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.

Systematic review A review of existing literature on a topic that is identified and reviewed using explicit methods. A WWC systematic review has five steps: 1) developing a review protocol; 2) searching the literature; 3) reviewing studies, including screening studies for eligibility, reviewing the methodological quality of each study, and reporting on high quality studies and their findings; 4) combining findings within and across studies; and, 5) summarizing the review.

Please see the WWC Procedures and Standards Handbook (version 3.0) for additional details.



An **intervention report** summarizes the findings of high-quality research on a given program, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against evidence standards, and summarizes the findings of those that meet standards.

This intervention report was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.