What Works Clearinghouse

ies INSTITUTE OF EDUCATION SCIENCES

Beginning Reading December 2008[1]

Lindamood Phonemic Sequencing (LiPS)®

Program Description[2]

The *Lindamood Phonemic Sequencing (LiPS)*® program (formerly called the *Auditory Discrimination in Depth*® *[ADD]* program) is designed to teach students skills to decode 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. After students are able to produce, label, and organize the sounds, subsequent activities in sequencing, reading, and spelling use the oral aspects of sounds to identify and order them within words. The program also offers direct instruction in letter patterns, sight words, and context clues in reading. The *LiPS*® program is individualized to meet students' needs and is often used with students who have learning disabilities or reading difficulties. The version of the program tested here involved computer-supported activities.

Research

One study of *LiPS*® meets What Works Clearinghouse (WWC) evidence standards. The study included 150 first-grade students in five elementary schools. The WWC considers the extent of evidence for *LiPS*® to be small for alphabetics and comprehension. No studies that meet WWC standards with or without reservations addressed fluency or general reading achievement.[3]

Effectiveness

Based on one study, *LiPS*® was found to have potentially positive effects on alphabetics and no discernible effects on comprehension. Findings on fluency and general reading achievement were not reported in the study.

	Alphabetics	*Fluency*	*Comprehension*	*General reading achievement*
Rating of effectiveness	Potentially positive	na	No discernible effect	na
Improvement index[4]	Average: +17 percentile points Range: –1 to +35 percentile points	na	Average: +6 percentile points Range: 0 to +20 percentile points	na

na = not applicable

1. This report has been updated to include reviews of 12 studies that have been released since 2005. Of the additional studies, all 12 were not within the scope of the protocol. A complete list and disposition of all studies reviewed are provided in the references.
2. The descriptive information for this program was obtained from a publicly available source: the program's website (http://www.lindamoodbell.com/programs/lips.html, downloaded October 2008). The WWC requests developers to review the program description sections for accuracy from their perspective. Further verification of the accuracy of the descriptive information for this program is beyond the scope of this review.
3. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.
4. These numbers show the average and range of student-level improvement indices for all findings across the study.

Additional program information

Developer and contact

Developed by Patricia Lindamood and Phyllis Lindamood, the *LiPS*® program is published by Pro-Ed and is available through a number of professional distributors and publishers. Address: 416 Higuera Street, San Luis Obispo, CA 93401. Web: http://www.lindamoodbell.com. Telephone: (800) 233-1819.

Scope of use

Auditory Discrimination in Depth® was developed in the late 1960s and early 1970s. It was revised and renamed *LiPS*® in 1998. The program is frequently provided in centers or clinics, including program-endorsed Lindamood-Bell Learning Centers. The program is available for purchase by the public. According to the program authors, the program is used widely for remedial purposes but exact numbers were not available.

Teaching

The program is designed for readers in kindergarten through third grade or for struggling readers. Teachers work with students in whole classes or in 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 four to six months for one hour a day or four to six weeks for four 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 *Lindamood Phonemic Sequencing Teacher's Manual*.

Cost

A kit of materials designed for one-on-one or small group instruction can be purchased for $298. The classroom kit costs $498. Kits include a trainer's manual and all student materials (tiles, blocks, colored felts, and picture cards). Some of these materials are also sold separately. Information is not available on the cost of training for instructors or on how much it costs for students to receive instruction at a licensed center.

Research

Thirty-seven studies reviewed by the WWC investigated the effects of *LiPS*®. One study (Torgesen, Wagner, Rashotte, & Herron, 2003) is a randomized controlled trial that meets WWC evidence standards. No studies are randomized controlled trials or quasi-experimental designs that meet WWC evidence standards with reservations. The remaining 36 studies do not meet either WWC evidence standards or eligibility screens.

Torgesen et al. (2003) included 150 low-achieving first-grade students in five elementary schools. At two schools, students were randomly assigned to either *LiPS*® (formerly *Auditory Discrimination in Depth*®) or *Read, Write and Type*™ *(RWT)*, a reading software program. At three additional schools, students were randomly assigned to either *LiPS*®, *RWT*, or a regular instruction control group.[5] The beginning reading review presents data relevant to comparisons of *LiPS*® with *RWT* and of *LiPS*® with a regular instruction control group.[6]

Extent of evidence

The WWC categorizes the extent of evidence in each domain as small or medium to large (see the What Works Clearinghouse Extent of Evidence Categorization Scheme). The extent of evidence takes into account the number of studies and the total sample size across the studies that meet WWC evidence standards with or without reservations.[7]

5. Description of the assignment procedure was based on personal communication with the first study author on September 7, 2006.
6. The WWC review of beginning reading includes all comparison groups that meet evidence standards because all schools provide some type of reading instruction, and there is no typical comparison condition.
7. The Extent of Evidence Categorization was developed to tell readers how much evidence was used to determine the intervention rating, focusing on the number and size of studies. Additional factors associated with a related concept—external validity, such as the students' demographics and the types of settings in which studies took place—are not taken into account for the categorization. Information about how the extent of evidence rating was determined for *LiPS*® is in Appendix A5.

Research *(continued)*

The WWC considers the extent of evidence for *LiPS*® to be small for alphabetics and comprehension. No studies that meet WWC evidence standards with or without reservations examined the effectiveness of *LiPS*® in the fluency or general reading achievement domains.

Effectiveness

Findings

The WWC review of interventions for Beginning Reading addresses student outcomes in four domains: alphabetics, fluency, comprehension, and general reading achievement. The study included in this report covers two domains: alphabetics and comprehension. The findings below present the authors' estimates and WWC-calculated estimates of the size and the statistical significance of the effects of *LiPS*® on students.[8]

Alphabetics. The Torgesen et al. (2003) study findings for alphabetics are based on the performance of *LiPS*® students and comparison students on three measures of phonological awareness and two measures of phonics.

- When the *LiPS*® group was compared with the *Read, Write and Type*™ group, the study authors found that there were no statistically significant differences between the groups on the three phonological awareness measures (Phoneme Blending, Phoneme Elision, and Phoneme Segmenting subtests of the Comprehensive Test of Phonological Processes) and the two phonics measures (Word Identification and Word Attack subtests of the Woodcock Reading Mastery Test).
- When the *LiPS*® group was compared with the regular classroom instruction group, the authors reported, and the WWC confirmed, statistically significant positive effects for *LiPS*® on two of the phonological awareness measures (Phoneme Elision and Phoneme Segmenting), but the authors did not find statistically significant effects on the third phonological awareness measure—Phoneme Blending. The authors reported, and the WWC confirmed, statistically significant positive effects of *LiPS*® on the two phonics measures (Word Identification and Word Attack).

Comprehension. The Torgesen et al. (2003) study findings for comprehension are based on the performance of *LiPS*® students and comparison students on the Passage Comprehension subtest of the Woodcock Reading Mastery Test and a Vocabulary subtest of the Stanford Binet Intelligence test (reported as a verbal IQ measure).

- When the *LiPS*® group was compared with the *Read, Write and Type*™ group, the authors found that there was no statistically significant difference between the groups on the comprehension measures.
- When the *LiPS*® group was compared with the regular classroom instruction group, the authors found that *LiPS*® had no statistically significant effect on the Vocabulary subtest. The authors found a statistically significant positive effect on the Passage Comprehension subtest. In WWC computations, the effect was not statistically significant.

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, 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).

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, see the WWC Tutorial on Mismatch. For the formulas the WWC used to calculate the statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Torgeson et al. (2003), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.

The WWC found *LiPS*® to have potentially positive effects for alphabetics and no discernible effects for comprehension

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 alphabetics is +17 percentile points in the one study across two comparison groups, with a range of –1 to +35 percentile points across findings.

The average improvement index for comprehension is +6 percentile points in the one study across two comparison groups, with a range of 0 to +20 percentile points across findings.

Summary

The WWC reviewed 37 studies on *LiPS*®. One of these studies meets WWC evidence standards; no studies meet WWC evidence standards with reservations; the remaining 36 studies do not meet either WWC evidence standards or eligibility screens. Based on the one study, the WWC found potentially positive effects in alphabetics and no discernible effects in comprehension. The conclusions presented in this report may change as new research emerges.

References

Meets WWC evidence standards

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). Retrieved from Florida Center for Reading Research website: http://www.fcrr.org/TechnicalReports/RWTfullrept.pdf.

Studies that fall outside the Beginning Reading protocol or do not meet WWC evidence standards

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 does not meet evidence standards because the measures of effect cannot be attributed solely to the intervention—the intervention was combined with another intervention.

Adair, J., Nadeau, S., Conway, T., Gonzalez-Rothi, L., Heilman, P., Green, I., et al. (2000). Alterations in the functional anatomy of reading induced by rehabilitation of an alexic patient. *Neuropsychiatry, Neuropsychology and Behavioral Neurology, 13*(4), 303–311. This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

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. This study is ineligible for review because it does not use a sample within the age or grade range specified in the protocol.

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. This study is ineligible for review because it does not use a sample within the age or grade range specified in the protocol.

Howard, M. P. (1986). Effects of pre-reading training in auditory conceptualization on subsequent reading achievement. *Dissertation Abstracts International, 47*(03), 847A. (UMI No. 8612677)

References *(continued)*

(Study: Arco, Indiana, and Santa Maria, California) The study does not meet evidence standards because the measures of effect cannot be attributed solely to the intervention—there was only one unit of analysis in one or both conditions.

Howard, M. P. (1986). Effects of pre-reading training in auditory conceptualization on subsequent reading achievement. *Dissertation Abstracts International, 47*(03), 847A. (UMI No. 8612677) (Study: Arco, Indiana, kindergarten) The study does not meet evidence standards because the measures of effect cannot be attributed solely to the intervention—there was only one unit of analysis in one or both conditions.

Howard, M. P. (1986). Effects of pre-reading training in auditory conceptualization on subsequent reading achievement. *Dissertation Abstracts International, 47*(03), 847A. (UMI No. 8612677) (Study: Arco, Indiana, first-grade longitudinal) The study does not meet evidence standards because the measures of effect cannot be attributed solely to the intervention—there was only one unit of analysis in one or both conditions.

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. This study is ineligible for review because it does not use a sample within the age or grade range specified in the protocol.

Leask, A. (2007). The effect of phonological awareness intervention on non-word spelling ability in school-aged children: An analysis of qualitative change. *Advances in Speech-Language Pathology, 9*(1), 1–16. The study is ineligible for review because it does not use a comparison group.

Lindamood-Bell Learning Processes. (2003). Lindamood-Bell Learning Processes: Beginning reading submissions. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: Intervention in kindergarten through 2nd grade) This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

Lindamood-Bell Learning Processes. (2003). Lindamood-Bell Learning Processes: Beginning reading submissions. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: Kindergarten results from school project in Oregon) The study is ineligible for review because it does not use a comparison group.

Lindamood-Bell Learning Processes. (2003). Lindamood-Bell Learning Processes: Beginning reading submissions. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: Kindergarten through 3rd grade results from learning centers across the United States) The study is ineligible for review because it does not use a comparison group.

Lindamood-Bell Learning Processes. (2003). Lindamood-Bell Learning Processes: Beginning reading submissions. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: Kindergarten through 3rd grade results from school project in Colorado) The study is ineligible for review because it does not use a comparison group.

Lindamood-Bell Learning Processes. (2004). Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: K–3 Lindamood-Bell focus students 2002 summary) This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

Lindamood-Bell Learning Processes. (2004). Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: Kindergarten students in Oregon 2001–02) This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

References *(continued)*

Lindamood-Bell Learning Processes. (2004). Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: Pueblo, Colorado, 2001–02 summary) This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

Lindamood-Bell Learning Processes. (2004). Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: Longitudinal Florida study summary) This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

Lindamood-Bell Learning Processes. (2004). Lindamood-Bell Learning Processes: Interventions for beginning reading evidence report–Report 1, Book I of II. (Available from the Lindamood-Bell Learning Processes, 416 Higuera Street, San Luis Obispo, CA 93401) (Study: Second grade students in Idaho) This study is ineligible for review because it does not provide enough information about its design to assess whether it meets standards.

Lindamood-Bell Learning Processes. 2004 clinical statistics. Retrieved from http://www.lindamoodbell.com/downloads/pdf/research/2004%20Clinical%20Stats.pdf The study is ineligible for review because it does not use a comparison group.

Lindamood-Bell Learning Processes. 2005 clinical statistics. Retrieved from http://www.lindamoodbell.com/downloads/pdf/research/clinical%20stats%202005.pdf The study is ineligible for review because it does not use a comparison group.

Lindamood-Bell Learning Processes. 2006 learning centers' results. Retrieved from http://www.lindamoodbell.com/downloads/pdf/research/2006%20Center%20results.pdf The study is ineligible for review because it does not use a comparison group.

Matson, A. E. (2005). Central auditory processing: A current literature review and summary of interviews with researchers on controversial issues related to auditory processing disorders. St. Louis, MO: Washington University School of Medicine, Program in Audiology and Communication Sciences. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention.

McBride, N. (2005). The effectiveness of second shot and/or Lindamood-Bell on reading achievement of elementary students. University of Nevada, Reno: Dept. of Counseling and Educational Psychology. (UMI No. 3209120) The study is ineligible for review because it does not disaggregate findings for the age or grade range specified in the protocol.

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 evidence standards because the measures of effect cannot be attributed solely to the intervention—there was only one unit of analysis in one or both conditions.

Nelson, R. L., & Damico, J. S. (2006). Qualitative research in literacy acquisition: A framework for investigating reading in children with language impairment. *Clinical Linguistics & Phonetics, 20*(7–8), 631–639. The study is ineligible for review because it does not use a comparison group.

Olson, R. K., & Wise, B. (2006). Computer-based remediation for reading and related phonological disabilities. In Michael C. McKenna, Linda D. Labbo, Ronald D. Kiefer, and David Rein-king, (Eds.), *International handbook of literacy and technology* (pp. 57–74). Mahwah, NJ: Erlbaum. The study is ineligible for review because it does not examine the effectiveness of an intervention.

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. This study is ineligible for review because it does not use a sample within the age or grade range specified in the protocol.

References *(continued)*

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. This study is ineligible for review because it does not disaggregate findings for the age or grade range specified in the protocol.

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 evidence standards because the measures of effect cannot be attributed solely to the intervention—the intervention was combined with another intervention.

Simos, P., Fletcher, J., Bergman, E., Breier, J., Foorman, B., Castillo, E., et al. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58,* 1203–1212. The study is ineligible for review because it does not include an outcome within a domain specified in 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. University of Sheffield, UK: Department for Education Skills Publications. The study is ineligible for review because it is not a primary analysis of the effectiveness of an intervention.

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. This study is ineligible for review because it does not disaggregate findings for the age or grade range specified in the protocol.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., et al. (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 does not meet evidence standards because the measures of effect cannot be attributed solely to the intervention—the intervention was combined with another intervention.

Truch, S. (1994). Stimulating basic reading processes using *Auditory Discrimination in Depth. Annals of Dyslexia, 44*, 60–80. This study is ineligible for review because it does not use a sample within the age or grade range specified in 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 does not meet evidence standards because the measures of effect cannot be attributed solely to the intervention—the intervention was combined with another intervention.

Appendix

Appendix A1 Study Characteristics: Torgeson, Wagner, Rashotte, & Herron, 2003 (randomized clinical trial)

Characteristic	Description
Study citation	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). Retrieved from Florida Center for Reading Research website: http://www.fcrr.org/TechnicalReports/RWTfullrept.pdf.
Participants	The study included 150 first-grade students in five elementary schools. All students scored in the lowest 35% on a letter-sound knowledge measure. At two of the schools, 16 students were randomly assigned to *LiPS*® and 16 to *Read, Write and Type!*™ *(RWT)*. At three schools, 38 students were randomly assigned to *LiPS*®, 38 to *RWT*, and 42 to a control group (J. K. Torgesen, personal communication, September 7, 2006). Two students left the *LiPS*® and *RWT* groups, and one student left the control group. The final sample for the analysis comparing *LiPS*® to *RWT* included 52 *LiPS*® students and 53 *RWT* students across five schools. The final sample for the analysis comparing *LiPS*® to control students included 36 *LiPS*® students and 41 control students across three schools. Approximately 34% of the sample were minority children (primarily African-American). Approximately 35% of the sample received free/reduced-price lunch, but the socioeconomic status of the students varied.
Setting	Five elementary schools (locations unknown).
Intervention	Students assigned to the *LiPS*® program were divided into groups of three children and received four 50-minute sessions a week from October through May. A trained teacher devoted half of each session to direct instruction. The remainder of the time the students worked individually on the computer practicing the same skills with the teacher in a support role.
Comparison	*RWT* students had the same format and time of instruction as did the *LiPS*® students, but they had different activities. *RWT* teachers began their sessions with warm-up activities, and students then worked on computers, with teachers lending support (particularly when students had difficulties). The computer component emphasized phonological awareness, letter sound correspondence, and phonemic decoding as students expressed themselves in written language. The control group had classroom instruction and support typically available to them (J. K. Torgesen, personal communication, September 7, 2006). Two of the three schools with regular instruction comparison groups used Open Court's *Collections for Young Scholars* as the whole-class reading curriculum.
Primary outcomes and measurement	The authors assessed students at the end of the study period using a battery of tests. All students in the sample were given the Phoneme Blending, Phoneme Elision, and Phoneme Segmenting subtests of the Comprehensive Test of Phonological Processes and the Word Attack, Word Identification, and Passage Comprehension subtests of the Woodcock Reading Mastery Test. Students in the study were also given the Vocabulary subtest of the Stanford Binet Intelligence Scale, which the authors used as a proxy for verbal IQ. Other outcomes were reported in the study but were not included in this review either because they were outside the scope of the beginning reading review (developmental spelling and probability of reading disability) or because sufficient information on the measure name, description, or validity and reliability was not reported (word efficiency and nonword efficiency). For a more detailed description of these outcome measures, see Appendices A2.1 and A2.2.
Staff/teacher training	No information was provided on teacher training.

Appendix A2.1 Outcome measures for the alphabetics domain

Outcome measure	Description
Phonological awareness	
Comprehensive Test of Phonological Processes (CTOPP): Phoneme Blending subtest	The Phoneme Blending subtest measures the student's ability to blend separately presented sounds together to form words. This is a standardized test (as cited in Torgesen et al., 2003).
CTOPP: Phoneme Elision subtest	The Phoneme Elision subtest measures the student's ability to manipulate sounds in words. This is a standardized test (as cited in Torgesen et al., 2003).
CTOPP: Phoneme Segmenting subtest	The Phoneme Segmenting subtest measures the student's ability to isolate and pronounce the sounds in words. This is a standardized test (as cited in Torgesen et al., 2003).
Phonics	
Woodcock Reading Mastery Test: Word Identification subtest	The Word Identification subtest is a measure of word reading vocabulary in which the student reads list of words of increasing difficulty. This is a standardized test (as cited in Torgesen et al., 2003).
Woodcock Reading Mastery Test: Word Attack subtest	The Word Attack subtest is a measure of phonemic reading ability in which the student reads nonwords. This is a standardized test (as cited in Torgesen et al., 2003).

Appendix A2.2 Outcome measures for the comprehension domain

Outcome measure	Description
Comprehension	
Woodcock Reading Mastery Test: Passage Comprehension subtest	The Passage Comprehension subtest measures the student's ability to comprehend the meaning of short passages. This is a standardized test (as cited in Torgesen et al., 2003).
Vocabulary	
Stanford Binet Intelligence Scale: Vocabulary subtest	The measure is based on the Vocabulary subtest of the Stanford Binet Intelligence Scale. The Vocabulary subtest measures the student's ability to provide names of pictures and definitions of words. This is a standardized test (as cited in Torgesen et al., 2003).

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

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *LiPS*® group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*LiPS*® - comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Torgeson, Wagner, Rashotte, & Herron, 2003 (randomized controlled trial)[7]								
Phonological Awareness								
Comparison #1: *LiPS*® vs. *Read, Write, and Type*								
CTOPP: Phoneme Blending subtest	Grade 1	104	18.80 (5.30)	18.90 (4.90)	–0.1	–0.02	ns	–1
CTOPP: Phoneme Elision subtest	Grade 1	104	14.30 (4.50)	13.50 (4.50)	0.8	0.18	ns	+7
CTOPP: Phoneme Segmentation subtest	Grade 1	104	16.20 (6.60)	15.30 (5.30)	0.8	0.15	ns	+6
Comparison #2: *LiPS*® vs. regular instruction/support								
CTOPP: Phoneme Blending subtest	Grade 1	77	20.60 (4.50)	18.20 (5.40)	2.4	0.48	ns	+18
CTOPP: Phoneme Elision subtest	Grade 1	77	15.30 (4.20)	12.50 (4.60)	2.8	0.63	Statistically significant	+23
CTOPP: Phoneme Segmentation subtest	Grade 1	77	15.60 (3.70)	11.70 (4.50)	3.9	0.93	Statistically significant	+32
Phonics								
Comparison #1: *LiPS*® vs. *Read, Write, and Type*								
Woodcock Reading Mastery Test: Word Attack subtest	Grade 1	104	109.70 (14.00)	106.30 (13.60)	3.4	0.24	ns	+10
Woodcock Reading Mastery Test: Word Identification subtest	Grade 1	104	107.10 (14.30)	105.10 (13.40)	2.0	0.14	ns	+6

(continued)

Appendix A3.1 Summary of study findings included in the rating for the alphabetics domain[1] *(continued)*

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *LiPS*® group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*LiPS*® - comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Torgeson, Wagner, Rashotte, & Herron, 2003 (randomized controlled trial)[7]								
Comparison #2: *LiPS*® vs. regular instruction/support								
Woodcock Reading Mastery Test: Word Attack subtest	Grade 1	77	113.70 (12.20)	99.50 (14.50)	14.2	1.04	Statistically significant	+35
Woodcock Reading Mastery Test: Word Identification subtest	Grade 1	77	110.60 (12.20)	100.10 (15.60)	10.5	0.74	Statistically significant	+27
Average for alphabetics, Comparison #1 (Torgeson et al., 2003)[8]						**0.14**	**ns**	**+6**
Average for alphabetics, Comparison #2 (Torgeson et al., 2003)[8]						**0.76**	**Statistically significant**	**+28**
Domain average for alphabetics across comparisons (Torgeson et al., 2003)[8]						**0.45**	**Statistically significant**	**+17**

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the alphabetics domain.
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. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. 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.
7. 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. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Torgeson et al. (2003), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
8. 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.

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

Outcome measure	Study sample	Sample size (students)	Authors' findings from the study: Mean outcome (standard deviation)[2] — *LiPS®* group	Authors' findings from the study: Mean outcome (standard deviation)[2] — Comparison group	WWC calculations: Mean difference[3] (*LiPS®*-comparison)	WWC calculations: Effect size[4]	WWC calculations: Statistical significance[5] (at $\alpha = 0.05$)	WWC calculations: Improvement index[6]
Torgeson, Wagner, Rashotte, & Herron, 2003 (randomized controlled trial)[7]								
Reading comprehension								
Comparison #1: *LiPS®* vs. *Read, Write, and Type*								
Woodcock Reading Mastery Test: Passage Comprehension subtest	Grade 1	104	99.90 (12.50)	99.30 (10.50)	0.6	0.05	ns	+2
Comparison #2: *LiPS®* vs. regular instruction/support								
Woodcock Reading Mastery Test: Passage Comprehension subtest	Grade 1	77	102.20 (10.00)	95.40 (14.40)	6.8	0.54	ns	+20
Vocabulary								
Comparison #1: *LiPS®* vs. *Read, Write, Type*								
Stanford Binet Intelligence Scale: Vocabulary subtest	Grade 1	104	95.50	95.50	0.0	0.00	ns	0
Comparison #2: *LiPS®* vs. regular instruction/support								
Stanford Binet Intelligence Scale: Vocabulary subtest	Grade 1	77	96.10 (12.50)	95.90 (11.30)	0.2	0.02	ns	+1
Average for comprehension, Comparison #1 (Torgeson et al., 2003)[8]						**0.03**	**ns**	**+1**
Average for comprehension, Comparison #2 (Torgeson et al., 2003)[8]						**0.28**	**ns**	**+11**
Domain average for comprehension across comparisons (Torgeson et al., 2003)[8]						**0.15**	**ns**	**+6**

ns = not statistically significant

1. This appendix reports findings considered for the effectiveness rating and the average improvement indices for the comprehension domain.
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. Positive differences and effect sizes favor the intervention group; negative differences and effect sizes favor the comparison group.
4. For an explanation of the effect size calculation, see Technical Details of WWC-Conducted Computations.
5. Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups.
6. 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.
7. 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. For the formulas the WWC used to calculate statistical significance, see Technical Details of WWC-Conducted Computations. In the case of Torgeson et al. (2003), a correction for multiple comparisons was needed, so the significance levels may differ from those reported in the original study.
8. 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.

Appendix A4.1 *LiPS*® rating for the alphabetics domain

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

For the outcome domain of alphabetics, the WWC rated *LiPS*® as having potentially positive effects. It did not meet the criteria for positive effects because only one study met WWC evidence standards. The remaining ratings (mixed effects, no discernible effects, potentially negative effects, and negative effects) were not considered, as *LiPS*® was assigned the highest applicable rating.

Rating received

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

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

  Met. One comparison within one study showed statistically significant positive effects.

AND

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

  Met. No studies showed statistically significant or substantively important negative effects or indeterminate effects.

Other ratings considered

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

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

  Not Met. Only one study met WWC evidence standards for a strong design.

AND

- 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. For a complete description, see the WWC Intervention Rating Scheme.

Appendix A4.2 *LiPS*® rating for the comprehension domain

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

For the outcome domain of comprehension, the WWC rated *LiPS*® as having no discernible effects. The remaining ratings (potentially negative effects and negative effects) were not considered, as *LiPS*® was assigned the highest applicable rating.

Rating received

No discernible effects: No affirmative evidence of effects.

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

  Met. No study showed a statistically significant or substantively important effect, either positive or negative.

Other ratings considered

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

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

  Not Met. Only one study met the WWC evidence standards for a strong design.

AND

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

  Met. No study showed statistically significant or substantively important negative effects.

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

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

  Not met. No study showed a statistically significant or substantively important positive effect.

AND

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

  Not met. No study showed a statistically significant or substantively important negative effect, but one study showed indeterminate effects.

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

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

  Not met. No study showed a statistically significant or substantively important effect, either positive or negative.

OR

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

  Not met. No study showed a statistically significant or substantively important effect, while one study showed indeterminate 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. For a complete description, see the WWC Intervention Rating Scheme.

Appendix A5 Extent of evidence by domain

Outcome domain	Number of studies	Sample size		Extent of evidence[1]
		Schools	Students	
Alphabetics	1	5	146	Small
Fluency	0	0	0	na
Comprehension	1	5	146	Small
General reading achievement	0	0	0	na

na = not applicable/not studied

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