What Works Clearinghouse™

ies INSTITUTE OF EDUCATION SCIENCES

Early Childhood Education Updated June 2013

Let's Begin with the Letter People®

Report Contents

Program Description[1]

Let's Begin with the Letter People® is an early childhood literacy curriculum that uses 26 thematic units—each of which covers a letter of the alphabet—to develop children's language and early literacy skills. A major focus of the program is phonological awareness, including rhyming, word play, alliteration, and segmentation. Children are encouraged to learn as individuals, in small groups, and in a whole-class environment. Teacher resource books and a set of classroom books and other program materials are available as a program kit.

Research[2]

The What Works Clearinghouse (WWC) identified three studies of *Let's Begin with the Letter People*® that fall within the scope of the Early Childhood Education topic area and meet WWC evidence standards.[3] Three studies meet WWC evidence standards with reservations, and together, they included 885 preschool children ages three to five in 24 preschools in the greater Houston, Texas area and southeastern New York state.

The WWC considers the extent of evidence for *Let's Begin with the Letter People*® on the school readiness of preschool children to be medium to large for two outcome domains—oral language and print knowledge—and small for one domain—phonological processing. There were no studies that meet standards in three other domains, so we do not report on the effectiveness of *Let's Begin with the Letter People*® for those domains in this intervention report. (See the Effectiveness Summary on p. 6 for a full list of all domains.)

Effectiveness

Let's Begin with the Letter People® was found to have no discernible effects on oral language and phonological processing and mixed effects on print knowledge for preschool children.

Table 1. Summary of findings[4]

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of children	Extent of evidence
		Average	Range			
Oral language	No discernible effects	–1	–6 to +3	3	879	Medium to large
Print knowledge	Mixed effects	+6	+1 to +14	3	881	Medium to large
Phonological processing	No discernible effects	–6	na	1	184	Small

na = not applicable

Program Information

Background

Let's Begin with the Letter People® was developed and is distributed by Abrams Learning Trends. Address: 16310 Bratton Lane, Suite 250, Austin, TX 78728-2403. Email: customerservice@abramslearningtrends.com. Web: www.abramslearningtrends.com. Telephone: (800) 227-9120.

Program details

Let's Begin with the Letter People® is an early childhood literacy curriculum that provides five thematically organized Teacher Resource books: *Ready, Set, Go!*; *At Home and At School*; *At the Park and At the Doctor's Office*; *In the Community and At the Zoo*; and *In the Garden and At the Museum*. Each of the Teacher Resource books contains six or seven units and offers varied teaching strategies and suggested activities. Each unit contains five daily menus of activities, including suggestions for interest centers (thematically arranged areas of the preschool classroom where teachers can interact with children individually and in small groups) and "meeting circle" (whole-class activities while children are sitting in a circle), as well as differentiation activities and Family Activity Pages. The Teacher Resource books are designed to provide teachers with a number of choices for teaching knowledge and skills in language and literacy, as well as in science, math, art, music, social development, health and safety, and motor skills. Through the interest centers, children are able to explore, investigate, construct, and apply knowledge. Skills are integrated into the classroom's daily events and are taught using a number of materials such as Letter People® hand puppets, children's literature, Big Books, Little Books and story tapes, songs and rhymes, virtual books, Ready to Read Rebus Books, Me Bag™ (for sharing special items), shapes and symbols, and Letter People® Stickables™. Teachers introduce concepts during meeting circle time that are then explored in the interest centers and other group activities. For instance, the Letter People® puppets (for example, Mr. N) are used to introduce letters, sounds, stories, colors, shapes, and characteristics. *Blueprint for Learning*, the program guide for *Let's Begin with the Letter People*®, provides an overview of the program and its components and includes information teachers can use for setting up their classrooms, as well as various instructional strategies. Teachers are trained during professional development activities and with other resources such as the Teacher Resource books.

Cost

Let's Begin with the Letter People® products can be purchased separately or in various combinations. The introductory set is available for $1,930 and includes the Teacher Resource File ($725), Letter People® Puppets Complete Set ($599), Meeting and Greeting Cards ($129), Let's Sing with the Letter People® CD ($30), Big and Little Books ($399), a read-along CD ($30), and a set of Letter People® Virtual Books ($199). Packages that include additional components at extra cost are also available. Information about the cost of professional development is not available. Additional pricing information for separate products from *Let's Begin with the Letter People*® is available on the website: http://www.abramslearningtrends.com.

Research Summary

The WWC identified nine studies that investigated the effects of *Let's Begin with the Letter People*® on the school readiness of preschool children.

Table 2. Scope of reviewed research

Grade	PK
Delivery method	Individual/Small group/ Whole class
Program type	Curriculum/Supplement

The WWC reviewed four of those studies against group design evidence standards. Three studies (Assel, Landry, Swank, & Gunnewig, 2007; Fischel, Bracken, Fuchs-Eisenberg, Spira, Katz, & Shaller, 2007; and Preschool Curriculum Evaluation Research [PCER] Consortium, 2008, Chapter 6) are randomized controlled trials that meet WWC evidence standards with reservations. Those three studies are summarized in this report. One study does not meet WWC evidence standards. The remaining five studies do not meet WWC eligibility screens for review in this topic area. Citations for all nine studies are in the References section, which begins on p. 9.

Summary of studies meeting WWC evidence standards without reservations

No studies of *Let's Begin with the Letter People*® met WWC evidence standards without reservations.

Summary of studies meeting WWC evidence standards with reservations

Assel et al. (2007) was a randomized controlled trial with high attrition at the subcluster level. As part of the PCER Consortium (2008, Chapter 6) study, schools in Houston, Texas with Title I and non-Title I (universal) preschool classrooms and Head Start centers were randomly assigned (within each program type) to the *Let's Begin with the Letter People*® intervention group, the *Doors to Discovery*™ intervention group, or a comparison group.[5] In the second stage of random assignment, classrooms assigned to either *Let's Begin with the Letter People*® or *Doors to Discovery*™ were randomly assigned either to receive mentoring or not to receive mentoring as part of the implementation of the intervention, creating a total of four intervention groups and one comparison group. For Assel et al. (2007), data were collected in the first year of the study for 550 children (182 *Let's Begin with the Letter People*®, 184 *Doors to Discovery*™, and 184 comparison) in 79 classrooms (25 *Let's Begin with the Letter People*®, 27 *Doors to Discovery*™, and 27 comparison).[6] Pretest data were collected prior to the implementation of the curriculum, and posttest data were collected at the end of the school year. The authors examined effects on oral language, print knowledge, and phonological processing. This review focuses on the comparison between the *Let's Begin with the Letter People*® intervention group and the comparison group. Findings for the oral language and print knowledge outcomes meet WWC evidence standards with reservations because the analytic sample of intervention and comparison group children are demonstrated as equivalent at baseline. Findings for the phonological processing domain do not meet WWC evidence standards due to a lack of baseline equivalence of the intervention and comparison children in the analytic sample at the end of the preschool year. The study also discusses differences in child outcomes between the mentoring and non-mentoring groups, but since the estimated differences are not presented in the paper, they are not included in supplemental analyses in this report.

Fischel et al. (2007) examined the effectiveness of the *Let's Begin with the Letter People*® and the *Waterford Early Reading*™ *Level One* curricula using a randomized controlled trial design. The study was conducted in six Head Start centers in southeastern New York state in the 2001–02, 2002–03, and 2003–04 school years (four centers in Year 1, one additional center in Year 2, and one additional center in Year 3). In each of the two intervention conditions, one of the experimental curricula was used in conjunction with the *High/Scope*® program that had been used by the Head Start programs for more than 10 years. Comparison classrooms used only the *High/Scope*® program. Each Head Start center contained classrooms assigned to at least two different study conditions. In Year 1 of the study, three classrooms were assigned to *Let's Begin with the Letter People*®, three to *Waterford Early Reading*™

Level One, and three to the comparison group. In Years 2 and 3, classrooms had unequal probabilities of assignment to study conditions. Continuing comparison group classrooms could only be assigned to one of the intervention groups: *Let's Begin with the Letter People*® or *Waterford Early Reading*™ *Level One*. New classrooms could be assigned to either intervention or comparison groups. Continuing intervention classrooms remained in the same intervention group. In the second year of the study, five classrooms implemented *Let's Begin with the Letter People*®, five implemented *Waterford Early Reading*™ *Level One*, and five were in the comparison group. In Year 3, four classrooms implemented *Let's Begin with the Letter People*®, four implemented *Waterford Early Reading*™ *Level One*, and three participated in the comparison group. The authors investigated effects on oral language and print knowledge using combined data from all 3 study years. The study sample across 3 years included 35 classrooms (12 *Let's Begin with the Letter People*® classrooms, 12 *Waterford Early Reading*™ *Level One* classrooms, and 11 comparison classrooms).[7]

Although the study used a randomized controlled trial design to assign teachers to intervention or comparison groups, the study included some teachers randomly assigned with unequal probabilities of assignment in Year 2 and Year 3 not accounted for in the analysis. Because of the unequal probabilities of assignment not accounted for in the analysis, under WWC standards, the study must demonstrate baseline equivalence between the intervention and comparison group samples of children used in the analysis of outcomes. Baseline equivalence was established from the data provided by the study authors, and the analyses controlled for pretest measures. A total of 507 children participated in the study across all three conditions. The analysis sample included 336 children (186 *Let's Begin with the Letter People*® and 150 comparison).[8] Pretest data were collected in the fall, and posttest data were collected in the spring, of the preschool year.

The PCER Consortium (2008, Chapter 6) assessed the effectiveness of *Let's Begin with the Letter People*® as part of the second year of the PCER initiative (2003–04 school year). This study included 44 of the original 79 full-day Head Start and public preschool classrooms in Houston, Texas in the year following the Assel et al. (2007) study. In the pilot year (the 2002–03 school year), schools in Houston, Texas with Title I and non-Title I (universal) preschool classrooms and Head Start centers were randomly assigned (within each program type) to either the *Let's Begin with the Letter People*® intervention group, the *Doors to Discovery*™ intervention group, or a comparison group. These schools included a total of 79 classrooms (25 *Let's Begin with the Letter People*®, 27 *Doors to Discovery*™, and 27 comparison). Teachers in participating classrooms from the 2002–03 school year were provided with a description of the national PCER study. From the group of teachers who agreed to participate in the PCER study during the 2003–04 school year, a subset of 45 were randomly selected to participate. After one teacher (and her classroom) dropped out, the sample included 15 *Let's Begin with the Letter People*® classrooms, 14 *Doors to Discovery*™ classrooms, and 15 comparison classrooms.

Although the study used a randomized controlled trial design to assign schools to intervention or comparison conditions in the pilot year, the PCER Consortium (2008, Chapter 6) study analyzed data from the second year of implementation, when children who had been in the classrooms at random assignment had moved on to kindergarten and a new class of children had replaced them. Thus, the study has high attrition at the child level and, under WWC standards, must demonstrate baseline equivalence between the intervention and comparison group sample of children used in the analyses of outcomes.

The authors investigated effects on oral language, print knowledge, phonological processing, and math. The WWC based its effectiveness ratings on findings from comparisons of 94 children who received *Let's Begin with the Letter People*® and 89 comparison children who received a variety of curricula, including teacher-developed, nonspecific curricula. Children who received the *Doors to Discovery*™ curriculum were not included in the comparisons. The study demonstrated the baseline equivalence of the intervention and comparison children in the analytic sample at the end of the preschool year for the outcome measures in the oral language, print knowledge, and phonological

processing domains. The study did not demonstrate baseline equivalence for the analytic samples for the outcomes in the math domain and, therefore, does not meet WWC evidence standards for this outcome. The authors reported on the effects of *Let's Begin with the Letter People*® in the spring of the preschool year and again at the end of kindergarten. The kindergarten findings are not reported here because information about the baseline equivalence of the outcome measures for the kindergarten sample was not provided in the report. The authors also reported findings on the Social Skills Rating Scale; however, these findings are not reported here because the current Early Childhood Education topic area protocol does not include sociobehavioral outcomes.

Effectiveness Summary

The WWC review of *Let's Begin with the Letter People*® for the Early Childhood Education topic area includes child outcomes in six domains: oral language, print knowledge, phonological processing, early reading and writing, cognition, and math. The three studies of *Let's Begin with the Letter People*® that meet WWC evidence standards reported findings in three of the six domains: (a) oral language, (b) print knowledge, and (c) phonological processing. The findings below present the authors' estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Let's Begin with the Letter People*® on preschool children. 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 oral language domain

Three studies that meet WWC standards with reservations reported findings in the oral language domain.

Assel et al. (2007) analyzed the effectiveness of *Let's Begin with the Letter People*® on oral language outcomes using the Expressive Vocabulary Test (EVT) and the Preschool Language Scale, Fourth Edition (PLS-4) Auditory Comprehension subscale. The authors reported differences between *Let's Begin with the Letter People*® and the comparison group within program type (Head Start, Title I, and universal prekindergarten) rather than across the combined program types. WWC analyses of the Assel et al. (2007) data show that the effect sizes for the *Let's Begin with the Letter People*® group for these measures are neither statistically significant nor substantively important according to WWC criteria (that is, an effect size of at least 0.25). The WWC characterizes these study findings as an indeterminate effect.

Fischel et al. (2007) analyzed the effectiveness of *Let's Begin with the Letter People*® on children's oral language outcomes using the Peabody Picture Vocabulary Test, Third Edition (PPVT-III). The authors reported that differences between *Let's Begin with the Letter People*® and the comparison group on these measures were not statistically significant and, according to WWC criteria, were not large enough to be considered substantively important. The WWC characterizes these study findings as an indeterminate effect.

The PCER Consortium (2008, Chapter 6) analyzed the effectiveness of *Let's Begin with the Letter People*® on child outcomes in oral language using the PPVT-III and the Test of Language Development–Primary: III (TOLD-P:3). The authors reported that differences between *Let's Begin with the Letter People*® and the comparison group on these measures were not statistically significant and, according to WWC criteria, were not large enough to be considered substantively important. The WWC characterizes these study findings as an indeterminate effect.

Thus, for the oral language domain, three studies showed indeterminate effects. This results in a rating of no discernible effects, with a medium to large extent of evidence.

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

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the three studies that reported findings, the estimated impact of the intervention on outcomes in the *oral language* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Medium to large	Three studies that included 879 children in 103 classrooms reported evidence of effectiveness in the *oral language* domain.

Summary of effectiveness for the print knowledge domain

Three studies that meet WWC standards with reservations reported findings in the print knowledge domain.

Assel et al. (2007) analyzed the effectiveness of *Let's Begin with the Letter People*® on children's print knowledge using the Woodcock-Johnson III (WJ-III), Letter-Word Identification subtest. The authors reported differences between *Let's Begin with the Letter People*® and the comparison group within program type (Head Start, Title I, and universal prekindergarten) rather than across the combined program types. WWC analyses of the Assel et al. (2007) data show that the effect size for the *Let's Begin with the Letter People*® group is not statistically significant nor substantively important according to WWC criteria (that is, at least 0.25). The WWC characterizes this study finding as an indeterminate effect.

Fischel et al. (2007) analyzed the effectiveness of *Let's Begin with the Letter People*® on children's print knowledge outcomes using the Woodcock-Johnson Psycho-Educational Battery–Revised (WJ-R) Letter-Word Identification and Dictation subtests, the Letters Known and Print Conventions subtests of the storybook assessment developed for the Head Start Family and Child Experiences Study (FACES), and the Get Ready to Read! (GRTR) screening instrument. The authors reported significant differences favoring *Let's Begin with the Letter People*® for the Dictation, Print Conventions, and GRTR measures. After correcting for the mismatch between the unit of assignment and the unit of analysis, the WWC does not find this difference to be statistically significant. However, WWC analyses of the Fischel et al. (2007) data show that the difference between the *Let's Begin with the Letter People*® group and the comparison group on Letters Known (0.34), the WJ-R Letter-Word Identification subtest (0.30), the WJ-R Dictation subtest (0.37), and the GRTR screen (0.33) is positive and substantively important (but not statistically significant), according to WWC criteria. The mean effect for the print knowledge domain, 0.31, was substantively important but not statistically significant. The WWC characterizes these study findings as a substantively important positive effect.

The PCER Consortium (2008, Chapter 6) analyzed the effectiveness of *Let's Begin with the Letter People*® on child outcomes in the print knowledge domain using the Test of Early Reading Ability, Third Edition (TERA-3) and the WJ-III Letter-Word Identification subtest. The authors reported that differences between the *Let's Begin with the Letter People*® group and the comparison group were not statistically significant and, according to WWC criteria, were not large enough to be considered substantively important. The WWC characterizes these study findings as an indeterminate effect.

Thus, for the print knowledge domain, one study showed substantively important positive effects and two studies showed indeterminate effects. This results in a rating of mixed effects, with a medium to large extent of evidence.

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

Rating of effectiveness	Criteria met
Mixed effects *Evidence of inconsistent effects.*	In the three studies that reported findings, the estimated impact of the intervention on outcomes in the *print knowledge* domain was positive and substantively important in one study and neither statistically significant nor large enough to be substantively important in the other two studies.
Extent of evidence	**Criteria met**
Medium to large	Three studies that included 881 children in 103 classrooms reported evidence of effectiveness in the *print knowledge* domain.

Summary of effectiveness for the phonological processing domain

One study that meets WWC standards with reservations reported findings in the phonological processing domain.

The PCER Consortium (2008, Chapter 6) analyzed the effectiveness of *Let's Begin with the Letter People*® on child outcomes in phonological processing using the Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) Elision subtest. The authors reported that differences between the *Let's Begin with the Letter People*® group and the comparison group were not statistically significant and, according to WWC criteria, were not large enough to be considered substantively important. The WWC characterizes these study findings as an indeterminate effect.

Thus, for the phonological processing domain, one study showed indeterminate effects. This results in a rating of no discernible effects, with a small extent of evidence.

Table 5. Rating of effectiveness and extent of evidence for the phonological processing domain

Rating of effectiveness	Criteria met
No discernible effects *No affirmative evidence of effects.*	In the one study that reported findings, the estimated impact of the intervention on outcomes in the *phonological processing* domain was neither statistically significant nor large enough to be substantively important.
Extent of evidence	**Criteria met**
Small	One study that included 184 children in 30 classrooms reported evidence of effectiveness in the *phonological processing* domain.

References

Studies that meet WWC evidence standards without reservations

None.

Studies that meet WWC evidence standards with reservations

Assel, M. A., Landry, S. H., Swank, P. R., & Gunnewig, S. (2007). An evaluation of curriculum, setting, and mentoring on the performance of children enrolled in pre-kindergarten. *Reading and Writing, 20*(5), 463–494.

Fischel, J. E., Bracken, S. S., Fuchs-Eisenberg, A., Spira, E. G., Katz, S., & Shaller, G. (2007). Evaluation of curricular approaches to enhance preschool early literacy skills. *Journal of Literacy Research, 39*(4), 471–501.

Preschool Curriculum Evaluation Research (PCER) Consortium. (2008, Chapter 6). Doors to Discovery and Let's Begin with the Letter People. In *Effects of preschool curriculum programs on school readiness* (pp. 85–98). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Study that does not meet WWC evidence standards

Hayes, K., Maddahian, E., & Fernandez, A. (2002). *An evaluation of pre-K reading programs* (Planning, Assessment, and Research Division Publication No. 137). Los Angeles, CA: Los Angeles Unified School District. The study does not meet WWC evidence standards because it uses a quasi-experimental design in which the analytic intervention and comparison groups are not shown to be equivalent.

Studies that are ineligible for review using the Early Childhood Education Evidence Review Protocol

Chambers, B., Cheung, A., Slavin, R. E., Smith, D., & Laurenzano, M. (2010). *Effective early childhood education programs: A systematic review*. Baltimore, MD: Johns Hopkins University, Center for Research and Reform in Education. Retrieved from http://www.bestevidence.org/word/early_child_ed_Sep_22_2010.pdf The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Fischer, K. M. (2012). *Efficacy study of the Letter People® programs 2009–2011*. Austin, TX: Abrams & Company Publishers, Inc. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Judkins, D., St. Pierre, R., Gutmann, B., Goodson, B., von Glatz, A., Hamilton, J., …Rimdzius, T. (2008). *A study of classroom literacy interventions and outcomes in Even Start* (NCEE 2008-4028). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. The study is ineligible for review because it does not examine an intervention implemented in a way that falls within the scope of the review—the intervention is bundled with other components.

Koontz, L. M. (2009). *Efficacy study of the Letter People® programs 2004–2009*. Waterbury, CT: Abrams & Company Publishers, Inc. The study is ineligible for review because it does not use a comparison group design or a single-case design.

Lehr, F. R. (2009). *Abrams learning trends: Research foundations for the Letter People® programs*. Waterbury, CT: Abrams & Company Publishers, Inc. The study is ineligible for review because it is a secondary analysis of the effectiveness of an intervention, such as a meta-analysis or research literature review.

Appendix A.1: Research details for Assel et al. (2007)

Assel, M. A., Landry, S. H., Swank, P. R., & Gunnewig, S. (2007). An evaluation of curriculum, setting, and mentoring on the performance of children enrolled in pre-kindergarten. *Reading and Writing, 20*(5), 463–494.

Table A1. Summary of findings

Meets WWC evidence standards with reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Oral language	52 classrooms/366 children	–3	nr
Print knowledge	52 classrooms/362 children	+2	nr

nr = not reported

Setting

The *Let's Begin with the Letter People*® study was conducted with children from 18 full-day preschools (52 classrooms: 25 *Let's Begin with the Letter People*® and 27 comparison) selected from Head Start and public preschool programs in the greater Houston, Texas area.[9]

Study sample

This randomized controlled study, conducted during the 2003–04 school year, included three groups: *Let's Begin with the Letter People*®, *Doors to Discovery*™, and a comparison group. Study authors recruited 32 Title I and non-Title I (universal) preschools and Head Start centers that included a total of 79 classrooms. Within each of the three school types (Title I, non-Title I, and Head Start), schools were randomly assigned to either the *Let's Begin with the Letter People*® intervention group, the *Doors to Discovery*™ intervention group, or a comparison group, with all classrooms within a preschool being assigned to the same intervention condition. Twelve schools in the study (25 classrooms) implemented *Let's Begin with the Letter People*®, 14 schools (27 classrooms) were assigned to implement the *Doors to Discovery*™ curriculum, and six schools (27 classrooms) were assigned to the comparison condition. In a second stage of random assignment, those schools assigned to *Let's Begin with the Letter People*® and *Doors to Discovery*™ were randomly assigned (within each curriculum) to two groups: one that received mentoring and one that did not receive mentoring. The number of children in the study classrooms at random assignment was not provided in the article or in author queries. However, the authors indicated that the consent rate was 65% in the intervention classrooms and 55% in the comparison classrooms. The authors also noted that at random assignment, 215 study children were in intervention group classrooms and 203 study children were in comparison classrooms.[10] Using the consent rate and the number of study children, the WWC estimates that the total sample of children included 705, with 300 in intervention classrooms and 405 in comparison classrooms. At baseline, 51% of children in the Head Start classrooms were Hispanic and 38% were African American; in the Title I classrooms, 53% of the children were Hispanic and 27% were Caucasian; and in the non-Title I classrooms, 71% of the children were Caucasian and 15% were classified as Other. Pretest data were collected prior to the implementation of the curriculum, and posttest data were collected at the end of the school year. The analysis sample for the *Let's Begin with the Letter People*® study included 366 children (182 *Let's Begin with the Letter People*® and 184 comparison).

Although the Assel et al. (2007) study used a randomized controlled trial design to assign schools to intervention or comparison conditions, the study had high attrition at the child

level and must demonstrate baseline equivalence between the intervention and comparison group samples of children used in the analyses of outcomes. An author query was conducted to obtain the study data necessary to establish equivalence at baseline (i.e., unadjusted means and standard deviations of the outcome measures for the intervention and comparison groups). Baseline equivalence was established for outcomes in the oral language and print knowledge domains but not for the phonological processing domain. The study also discusses differences in child outcomes for the groups that received mentoring compared with those that did not, but since the estimated differences are not presented in the paper, we do not present these analyses in this intervention report.

Intervention group

Intervention group teachers implemented *Let's Begin with the Letter People*®. The curriculum is organized into 26 thematic units, with teachers typically devoting at least one week for each unit. Fidelity to the curriculum was measured three times during the school year. At the first evaluation, 30% of teachers scored at high levels for curriculum fidelity (4 or 5 on a 5-point scale). By mid-year, 72% of teachers received high scores for curriculum fidelity.

A second intervention group was assigned to the *Doors to Discovery*™ curriculum; the effects of this intervention on the study sample are not discussed in this report.

Comparison group

Comparison group teachers used nonspecific curricula, which included a variety of curriculum materials that followed state guidelines for public preschool programs. Head Start comparison classrooms did not use a curriculum with a specified scope or sequence.

Outcomes and measurement

Outcome domains that met WWC evidence standards with reservations include oral language and print knowledge. Oral language was assessed with the PLS-4 Auditory Comprehension Subscale and EVT. Print knowledge was assessed with the WJ-III Letter-Word Identification subtest. For a more detailed description of the outcome measures presented in this intervention report, see Appendix B. In addition, the study authors assessed children in the phonological processing domain. Phonological processing was assessed with the Developing Skill Checklist (DSC) and the Rhyming section from the WJ-III Sound Awareness subtest. Children were assessed at the beginning and end of the school year.

Support for implementation

Intervention teachers were trained on *Let's Begin with the Letter People*® by the curriculum's publishing company during a 4-day workshop. Training took place in small groups and included instruction in all content areas. The mentors were senior-level trainers of the curriculum. Teachers in the mentoring condition received help from one of three senior-level trainers of the curriculum who served as mentors. Mentors met with teachers two times a month for about 1.5 hours, providing assistance in areas of lesson planning, curriculum components, and fidelity, among other topics. Mentors also identified and discussed areas of improvement for individual teachers. All teachers, regardless of mentoring condition, received three feedback sessions over the course of the school year, surrounding their implementation of the intervention.

Appendix A.2: Research details for Fischel et al., 2007

Fischel, J. E., Bracken, S. S., Fuchs-Eisenberg, A., Spira, E. G., Katz, S., & Shaller, G. (2007). Evaluation of curricular approaches to enhance preschool early literacy skills. *Journal of Literacy Research, 39*(4), 471–501.

Table A2. Summary of findings

Meets WWC evidence standards with reservations

Outcome domain	Sample size	Study findings	
		Average improvement index *(percentile points)*	Statistically significant
Oral language	21 classrooms/329 children	+1	No
Print knowledge	21 classrooms/335 children	+12	No

Setting

The *Let's Begin with the Letter People*® study was conducted with children from six full-day Head Start centers in southeastern New York state (four centers in Year 1; one additional center in Year 2; and one additional center in Year 3) during the 2001–02, 2002–03, and 2003–04 school years. All centers were part of the same Head Start grantee.

Study sample

This randomized controlled study included three groups: *Let's Begin with the Letter People*®, the *Waterford Early Reading*™ *Level One*, and a comparison group. Preschool classrooms in six centers were randomly assigned to one of the two curricular approaches overlaid onto their standard curriculum (*High/Scope*® *Educational Approach*), or to a business-as-usual comparison group that used only the *High/Scope*® *Educational Approach* over 3 years of the study.[11] In Year 1 of the study, nine classrooms participated, with three classrooms assigned to each group. In Year 2, 15 classrooms participated in the study, with five in each group. In Year 3, 11 classrooms participated, with four in each intervention group and three in the comparison group. In Years 2 and 3, some teachers from the intervention groups were randomly selected to continue in the study (remaining in that curriculum group), some teachers from the comparison group could opt for random assignment to one of the intervention groups, and some new teachers were randomly assigned to fill out the study groups.

Data from all three study years were analyzed together, yielding a sample of 35 study classrooms (12 *Let's Begin with the Letter People*® classrooms, 12 *Waterford Early Reading*™ *Level One* classrooms, and 11 comparison classrooms).[12] Thus, although the study used a randomized controlled trial design to assign classrooms to intervention or comparison groups, the study included some classrooms randomly assigned with unequal probabilities of assignment not accounted for in the analysis. As a result, the study must demonstrate baseline equivalence between the intervention and comparison group samples of children used in the analysis of outcomes. An author query was conducted to obtain the study data necessary to establish equivalence at baseline for the outcome measures (i.e., unadjusted means and standard deviations of the outcome measures for the intervention and comparison groups). Baseline equivalence was established from the data provided by the study authors.[13]

A total of 507 children participated in the study across all three conditions over the 3 years of the study. Children in the study sample had a mean age of 4 years, 4 months at the time of pretest. The sample of children included African-American (42%), Hispanic (41%), multiracial (8%), Caucasian (7%), and other race/ethnicity (2%). Approximately 14% of the sample were

Spanish-language dominant at Head Start entry. The evaluation of *Let's Begin with the Letter People*® included 21 of the 35 study classrooms, and the analysis sample included 336 children (186 *Let's Begin with the Letter People*® and 150 comparison).

Intervention group

The intervention group classrooms used the *Let's Begin with the Letter People*® curriculum overlaid onto the existing *High/Scope*® curriculum, which all programs had used for at least 10 years before the study. Teachers incorporated the activities, lessons, and materials from *Let's Begin with the Letter People*® into their daily classroom instruction, while continuing to follow the *High/Scope*® curriculum. *Let's Begin with the Letter People*® also includes take-home activity sheets designed to be completed with the help of an adult; however, the study did not assess the degree to which parents engaged their children in these activities.

A second intervention group was assigned to the *Waterford Early Reading™ Level One* curriculum; the effects of this intervention on the study sample are not discussed in this report.

Comparison group

The business-as-usual comparison group classrooms used the standard classroom curriculum (*High/Scope*®), which prescribes a daily routine (planning time, work time, cleanup time, time for recall, large-group time, small-group time, and outdoor play) and aligns with Head Start's performance standards, focusing on language, literacy, and other school readiness skills, such as numeracy, reasoning, problem solving, and decision making.

Outcomes and measurement

The outcome domains assessed were children's oral language and print knowledge. Oral language was assessed with the PPVT-III. Print knowledge was assessed with five measures: the WJ-R Letter-Word Identification and Dictation subtests, the Letter Knowledge and Print Conventions subtests of the storybook assessment developed for the Head Start FACES study, and the GRTR screening instrument.[14] Spanish-language dominant children were excluded by the authors from analyses involving the WJ-R Dictation subtest, as these children did not receive the English-language version of this measure at pretest. For the GRTR, Letter Knowledge, and Print Conventions assessments, the pretest administration for Spanish-language dominant children used Spanish instructions, with assessment questions administered in English. Pretests were administered in October/November, and posttests in May/June, corresponding to the Head Start academic year. For a more detailed description of these outcome measures, see Appendix B.

Support for implementation

Teachers and teacher assistants in the *Let's Begin with the Letter People*® group participated in a 3-day curriculum training each August conducted by a professional trainer from Abrams & Company (the developer and distributor of this curriculum). The trainer visited each classroom in the *Let's Begin with the Letter People*® condition in the fall and spring of each intervention year and provided individual feedback to teachers. Fidelity was measured during these visits using a checklist to assess the degree of implementation in two domains: classroom organization and teacher behavior. Implementation by all teachers in each year of the study was determined to be accurate and appropriate. Fischel et al. (2007) reported that additional training was offered by the trainer; however, details of the frequency, content, or degree of participation in these trainings were not provided. Teachers and assistants in the *Let's Begin with the Letter People*® group and the business-as-usual comparison group participated in a week-long in-service *High/Scope*® curriculum training at the beginning of the school year. Support was provided in the classroom by educational and child development specialists throughout the school year.

Appendix A.3: Research details for PCER Consortium, 2008, Chapter 6

Preschool Curriculum Evaluation Research (PCER) Consortium. (2008, Chapter 6). Doors to Discovery and Let's Begin with the Letter People. In *Effects of preschool curriculum programs on school readiness* (pp. 85–98). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Table A3. Summary of findings

Meets WWC evidence standards with reservations

		Study findings	
Outcome domain	Sample size	Average improvement index *(percentile points)*	Statistically significant
Oral language	30 classrooms/184 children	0	No
Print knowledge	30 classrooms/184 children	+4	No
Phonological processing	30 classrooms/184 children	–6	No

Setting

The *Let's Begin with the Letter People*® study was conducted with children from 30 full-day preschool classrooms (15 *Let's Begin with the Letter People*® and 15 comparison) selected from Head Start and public preschool programs in the greater Houston, Texas area.

Study sample

This randomized controlled study, conducted during the 2003–04 and 2004–05 school years, included three groups: *Let's Begin with the Letter People*®, *Doors to Discovery*™, and a comparison group. Study authors recruited 32 Title I and non-Title I preschools and Head Start centers that included a total of 79 classrooms. Within each of the three school types (Title I, non-Title I, and Head Start), schools were randomly assigned to either the *Let's Begin with the Letter People*® intervention group, the *Doors to Discovery*™ intervention group, or a comparison group, with all classrooms within a preschool being assigned to the same intervention condition. Twelve schools in the study (25 classrooms) were assigned to implement *Let's Begin with the Letter People*®, 14 schools (27 classrooms) implemented the *Doors to Discovery*™ curriculum, and six schools (27 classrooms) were assigned to the comparison condition.

Subsequent to randomization, teachers were provided with a description of the national PCER study; of those that opted to participate in the national PCER study during the 2003–04 school year, 45 were randomly selected (15 from each group). All 79 classrooms participated in the local investigator's pilot study during the first year. Following the pilot year, and prior to starting the national PCER study, one teacher (and her classroom) dropped out of the study, leaving 15 *Let's Begin with the Letter People*®, 14 *Doors to Discovery*™, and 15 comparison classrooms.

The evaluation of *Let's Begin with the Letter People*® included 30 of the 44 classrooms (15 *Let's Begin with the Letter People*® and 15 comparison, while the remaining 14 were assigned to *Doors to Discovery*™). Seven children (whose parents had provided consent to participate in the study) were randomly selected from each classroom at baseline for inclusion in the study.[15] The number of children participating in the study at baseline was 196 (100 *Let's Begin with the Letter People*® and 96 comparison). The parental consent rate was 65% for the intervention group and 55% for the comparison group.[16] At baseline, children in the study averaged 4.6 years of age; 55% were male; 43% were Hispanic, 30% were White, and 13% were African American. The analysis sample for the *Let's Begin with the Letter People*® study included 184 children (95 *Let's Begin with the Letter People*® and 89 comparison).

For the PCER Consortium (2008, Chapter 6) study, the *Let's Begin with the Letter People*® intervention had been in place for a full (pilot) year when the evaluation year started. Although the PCER Consortium (2008, Chapter 6) study used a randomized controlled trial design to assign classrooms to intervention or comparison conditions in the pilot year, the study analyzed data from the second year of implementation, when children who had been in the classrooms at random assignment had moved to kindergarten and a new class of children had replaced them. Thus, the study had high attrition at the child level and must demonstrate baseline equivalence between the intervention and comparison group sample of children used in the analyses of outcomes. An author query was conducted to obtain the study data necessary to establish equivalence at baseline for one outcome measure in each domain (i.e., unadjusted means and standard deviations of the outcome measures for the intervention and the comparison groups). The pretest data provided for each domain were used to establish baseline equivalence for the domain. Baseline equivalence was established from the data provided by the study authors. Baseline equivalence of the analytic sample of children in the two groups at the end of kindergarten was not available, so findings from the kindergarten follow-up are not reported.

Intervention group

Intervention group teachers implemented *Let's Begin with the Letter People*®. Each classroom's fidelity to the curriculum was rated on a four-point scale ranging from "not at all" (0) to "high" (3). The average score for *Let's Begin with the Letter People*® classrooms was 1.86 on this measure.

A second intervention group was assigned to the *Doors to Discovery*™ curriculum; the effects of this intervention on the study sample are not discussed in this report.

Comparison group

Comparison group classrooms used a variety of teacher-developed, nonspecific curricula reflecting the business-as-usual curricula for those classrooms. Comparison teachers' classrooms were rated with the same fidelity measure used in the *Let's Begin with the Letter People*® classrooms, which ranged from 0 to 3. The average score for the comparison classrooms using this measure was 1.0.

Outcomes and measurement

The outcome domains assessed were children's oral language, print knowledge, phonological processing, and math. Oral language was assessed with the PPVT-III and the TOLD-P:3 Grammatic Understanding subtest. Print knowledge was assessed with the TERA-3 and the WJ-III Letter-Word Identification and Spelling subtests. Phonological processing was assessed with the Pre-CTOPPP Elision subtest. For a more detailed description of these outcome measures, see Appendix B. Math was assessed with the WJ-III Applied Problems subtest, the Child Math Assessment–Abbreviated (CMA-A), and the Building Blocks Shape Composition task. The math outcomes are not reported because the findings for this domain do not meet WWC evidence standards due to a lack of baseline equivalence of the intervention and comparison children in the analytic sample.

Support for implementation

Teachers received curriculum training prior to the start of the 2003–04 school year. This was the second year of implementation of the intervention, and most of the teachers had been trained prior to the start of the 2002–03 school year. New teachers each received 12 hours of training, and returning teachers each received 6 hours of training. The research team collected site-specific curriculum fidelity data three times during the preschool year. All classrooms were observed using the Teacher Behavior Rating Scale in fall and spring of the preschool year.

Appendix B: Outcome measures for each domain

Oral language	
Expressive Vocabulary Test (EVT)	A nationally-standardized, individually-administered assessment of expressive vocabulary and word retrieval skills appropriate for ages 2.5 and up. Expressive vocabulary is measured through the use of labeling ("What do you see?") and synonym ("What is another word for __?") items (as cited in Assel et al., 2007).
Peabody Picture Vocabulary Test, Third Edition (PPVT-III)	A nationally-standardized, individually-administered assessment of children's receptive vocabulary, in which children demonstrate understanding of a spoken word by pointing to a picture that best represents the meaning (as cited in Fischel et al., 2007, and PCER Consortium, 2008, Chapter 6).
Preschool Language Scale, Fourth Edition (PLS-4) Auditory Comprehension Subscale	A nationally-standardized, individually-administered assessment of children's receptive vocabulary and their understanding of complex language forms (as cited in Assel et al., 2007).
Test of Language Development–Primary: III (TOLD-P:3) Grammatic Understanding subtest	A nationally-standardized, individually-administered assessment of children's ability to comprehend the meaning of sentences by selecting pictures that most accurately represent the sentence (as cited in PCER Consortium, 2008, Chapter 6).
Print knowledge	
Get Ready to Read! (GRTR) Screening test	A 20-question, nonstandardized screening test designed to measure emergent writing skills (identifying clearest writing exemplars), print knowledge (differentiating print from pictures, letter naming, and identifying letter sounds), and phonological awareness (rhyming, segmenting words, and deletion of sounds) (as cited in Fischel et al., 2007).
Letter Knowledge subtest	A test developed for use in the Head Start Family and Child Experiences Study (FACES) requiring the child to identify as many letters as possible from arrays of uppercase letters (as cited in Fischel et al., 2007).
Print Conventions subtest	A test developed for use in the Head Start FACES using a storybook, in which a child is asked to identify print conventions (for example, reading from left to right and top to bottom, location of page start, page turning) (as cited in Fischel et al., 2007).
Test of Early Reading Ability–III (TERA-3)	A nationally-standardized, individually-administered assessment of children's developing reading skills with three subtests: alphabet, conventions, and meaning (as cited in PCER Consortium, 2008, Chapter 6).[17]
Woodcock-Johnson III (WJ-III) Letter-Word Identification subtest	A nationally-standardized, individually-administered assessment of identification of letters and reading of words (as cited in Assel et al., 2007, and PCER Consortium, 2008, Chapter 6).
WJ-III Spelling subtest	A nationally-standardized, individually-administered assessment that assesses children's prewriting skills, such as drawing lines, tracing, and writing letters (as cited in PCER Consortium, 2008, Chapter 6).
Woodcock-Johnson–Revised (WJ-R) Dictation subtest	A nationally-standardized, individually-administered assessment that assesses children's prewriting skills, such as drawing lines and copying letters, as well as more advanced writing skills, including writing letters and phrases, capitalization, and punctuation (as cited in Fischel et al., 2007).
WJ-R Letter-Word Identification subtest	A nationally-standardized, individually-administered assessment of identification of letters and reading of words (as cited in Fischel et al., 2007).
Phonological processing	
Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) Elision subtest	An individually-administered assessment of children's ability to identify and manipulate sounds in spoken words, using word prompts and picture plates for the first nine items and word prompts only for later items (as cited in PCER Consortium, 2008, Chapter 6).

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Assel et al., 2007[a]								
Expressive Vocabulary Test (EVT)	Preschool children	52 classrooms/ 366 children	nr	nr	nr	–0.16	–6	> 0.05
Preschool Language Scale, Fourth Edition (PLS-4) Auditory Comprehension subscale	Preschool children	52 classrooms/ 366 children	nr	nr	nr	0.03	+1	> 0.05
Domain average for oral language (Assel et al., 2007)						**–0.07**	**–3**	**Not statistically significant**
Fischel et al., 2007[b]								
Peabody Picture Vocabulary Test, Third Edition (PPVT-III)	Preschool children	21 classrooms/ 329 children	85.66 (13.79)	85.36 (13.62)	0.30	0.02	+1	> 0.05
Domain average for oral language (Fischel et al., 2007)						**0.02**	**+1**	**Not statistically significant**
PCER Consortium, 2008, Chapter 6[c]								
PPVT-III	Preschool children	30 classrooms/ 184 children	nr	nr	nr	–0.08	–3	> 0.05
Test of Language Development–Primary: III (TOLD-P:3) Grammatic Understanding subtest	Preschool children	30 classrooms/ 184 children	nr	nr	nr	0.08	+3	> 0.05
Domain average for oral language (PCER Consortium, 2008, Chapter 6)						**0.00**	**0**	**Not statistically significant**
Domain average for oral language across all studies						**–0.02**	**–1**	**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 child outcomes, representing the average change expected for all children 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 child's percentile rank that can be expected if the child 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. nr = not reported. na = not applicable.

[a] For Assel et al. (2007), the effect sizes and *p*-values presented here were calculated by the WWC using data provided by the authors, adjusting for the pretest. Mean scores and differences are not reported in this table because the author-reported group means were not adjusted for the baseline pretest scores. The study did not report group differences or effect sizes. A correction for clustering was needed and resulted in a WWC-computed *p*-value of 0.90 for the *PLS-4* and a WWC-computed *p*-value of 0.49 for the *EVT*; therefore, the WWC does not find the results to be statistically significant. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important.

[b] For Fischel et al. (2007), the effect size presented here was calculated by the WWC using means, standard deviations, and intraclass correlation coefficients reported in the study as well as sample sizes provided by the authors, correcting for clustering and adjusting for the pretest. (The study-reported effect size did not adjust for either pretests or clustering.) The *p*-value presented here was reported in the original study. A correction for clustering was needed but did not affect whether the contrast was found to be statistically significant. The study did not report group differences. The WWC calculated the program group mean using a difference-in-differences approach (see the WWC Handbook) 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. This study is characterized as having an indeterminate effect because the effect is neither statistically significant nor substantively important.

[c] For PCER Consortium (2008, Chapter 6), the effect sizes and *p*-values presented here were reported in the original study (in Table A-13, based on an alternative estimation approach, analysis of covariance [ANCOVA], that included the baseline pretest). A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. Mean scores and differences are not reported in this table because the study-reported group means and differences were not adjusted for the baseline pretest scores. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important.

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

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
Assel et al., 2007[a]								
Woodcock-Johnson III (WJ-III) Letter-Word Identification subtest	Preschool children	52 classrooms/ 362 children	nr	nr	nr	0.04	+2	> 0.05
Domain average for print knowledge (Assel et al., 2007)						**0.04**	**+2**	**Not statistically significant**
Fischel et al., 2007[b]								
Get Ready to Read! (GRTR) Screen	Preschool children	335 children	nr	nr	nr	0.33	+13	0.01
Letter Knowledge subtest	Preschool children	334 children	nr	nr	nr	0.34	+13	0.06
Print Conventions subtest	Preschool children	334 children	nr	nr	nr	0.23	+9	0.04
Woodcock-Johnson–Revised (WJ-R) Letter-Word Identification subtest	Preschool children	284 children	nr	nr	nr	0.30	+12	> 0.05
WJ-R Dictation subtest	Preschool children	233 children	nr	nr	nr	0.37	+14	0.01
Domain average for print knowledge (Fischel et al., 2007)						**0.31**	**+12**	**Not statistically significant**
PCER Consortium, 2008, Chapter 6[c]								
Test of Early Reading Ability–III (TERA-3)	Preschool children	30 classrooms/ 183 children	92.94 (16.06)	92.76 (17.86)	0.18	0.02	+1	> 0.05
WJ-III Letter-Word Identification subtest	Preschool children	30 classrooms/ 184 children	108.72 (12.54)	106.04 (13.82)	2.68	0.10	+4	> 0.05
WJ-III Spelling subtest	Preschool children	30 classrooms/ 184 children	101.34 (13.01)	97.37 (12.63)	3.97	0.17	+7	> 0.05
Domain average for print knowledge (PCER Consortium, 2008, Chapter 6)						**0.10**	**+4**	**Not statistically significant**
Domain average for print knowledge 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 child outcomes, representing the average change expected for all children 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 child's percentile rank that can be expected if the child 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. nr = not reported. na = not applicable.

[a] For Assel et al. (2007), the effect size and *p*-value presented here were calculated by the WWC using data provided by the authors, adjusting for the pretest. Mean scores and the group difference are not reported in this table because the author-reported group means were not adjusted for the baseline pretest scores. The study did not report group differences or effect sizes. A correction for clustering was needed and resulted in a WWC-computed *p*-value of 0.85 for the *WJ-III Letter-Word Identification subtest*; therefore, the WWC does not find the results to be statistically significant. This study is characterized as having an indeterminate effect because the effect is neither statistically significant nor substantively important.

[b] For Fischel et al. (2007), the effect sizes presented here were calculated by the WWC using data provided by the authors, adjusting for pretests. (The study-reported effect sizes did not adjust for pretests.) The *p*-values presented here were reported in the original study. A correction for clustering was needed and resulted in a WWC-computed *p*-value of 0.11 for the *Letter Knowledge subtest*, a WWC-computed *p*-value of 0.16 for the *Print Conventions subtest*, and a WWC-computed *p*-value of 0.10 for the *WJ-R Dictation subtest*; therefore, the WWC does not find the results to be statistically significant. The study authors calculated outcome-specific intra-class correlations (ICCs) if the random effect of classroom was significant in their ANCOVA analysis. These ICCs are used when available; otherwise, the default ICC value of 0.20 is used, as specified in the WWC Handbook. A correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. Mean scores and differences are not reported in this table because the study-reported group means were not adjusted for the baseline pretest scores. The study did not report group differences. This study is characterized as having a substantively important positive effect because the average effect is positive and substantively important (exceeds 0.25), and there is no evidence of a negative effect.

[c] For PCER Consortium (2008, Chapter 6), the effect sizes, mean differences, and *p*-values presented here were reported in the original study (in Table 6.5, Table D-7a, and Table 6.5, respectively). A correction for clustering and multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. Adjustment for the baseline pretest scores was not required for this domain; therefore, the unadjusted means (and standard deviations) reported in the study (in Table C-7a) are presented. This study is characterized as having an indeterminate effect because the mean effect is neither statistically significant nor substantively important.

Appendix C.3: Findings included in the rating for the phonological processing domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			*p*-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
PCER Consortium, 2008, Chapter 6[a]								
Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) Elision subtest	Preschool children	30 classrooms/ 184 children	nr	nr	nr	–0.16	–6	> 0.05
Domain average for phonological processing (PCER Consortium, 2008, Chapter 6)						**–0.16**	**–6**	**Not statistically significant**
Domain average for phonological processing across all studies						**–0.16**	**–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 child outcomes, representing the average change expected for all children 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 child's percentile rank that can be expected if the child 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 the study's domain average was determined by the WWC. nr = not reported. na = not applicable.

[a] For PCER Consortium (2008, Chapter 6), the effect sizes and *p*-values presented here were reported in the original study (in Table A-13, based on an alternative estimation approach, analysis of covariance [ANCOVA], that included the baseline pretest). A correction for multiple comparisons was not needed. Mean scores and differences are not reported in this table because the study-reported group means and differences were not adjusted for the baseline pretest scores. This study is characterized as having an indeterminate effect because the effect is neither statistically significant nor substantively important.

Endnotes

[1] The descriptive information for this program was obtained from publicly available sources: the program's website (http://www.abramslearningtrends.com, downloaded February 2012) and the research literature (Assel et al., 2007; Fischel et al., 2007). The WWC requests developers review the program description sections for accuracy from their perspective. The program description was provided to the developer in March 2012, 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. The literature search reflects documents publicly available by December 2012.

[2] The previous report was released in September 2009. This report has been updated to include reviews of four studies that have been released since September 2009. Of the additional studies, four were not within the scope of the review protocol for the Early Childhood Education topic area. A complete list and disposition of all studies reviewed are provided in the references. The report includes reviews of all previous studies that met WWC evidence standards with or without reservations and resulted in a revised disposition of Assel et al. (2007), Fischel et al. (2007), and PCER Consortium (2008, Chapter 6). The revised disposition for all three studies is: *meets standards with reservations*. Assel et al. (2007) is a randomized controlled trial with high subcluster attrition that demonstrated equivalence for the analytic sample through response to an author query. Assel et al. (2007) meets standards with reservations in the current report; the study did not meet standards in the previous report due to high attrition and a lack of evidence of baseline equivalence. Fischel et al. (2007) is a randomized controlled trial with high subcluster attrition and unequal probabilities of selection not accounted for in the analysis that demonstrated baseline equivalence for the analytic sample. Fischel et al. (2007) meets standards with reservation in the current report; the study met standards without reservations in the previous report. PCER Consortium (2008, Chapter 6) is a randomized controlled trial with severe attrition that demonstrated baseline equivalence for the analytic sample. PCER Consortium (2008, Chapter 6) meets standards with reservations in the current report; the study met standards without reservations in the previous report. The revised dispositions are due to a change in the review protocol, particularly in baseline equivalence standards, as well as information received through author queries. The studies in this report were reviewed using the Evidence Standards from the WWC Procedures and Standards Handbook (version 2.1), along with those described in the Early Childhood Education review protocol (version 2.0). The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.

[3] Absence of conflict of interest: The PCER Consortium (2008, Chapter 6) study summarized in this intervention report was prepared by staff of one of the WWC contractors. Because the principal investigator for the WWC review of early childhood education is also a staff member of that contractor, the study was rated by staff members from a different organization, who also prepared the intervention report. The report was then reviewed by the principal investigator, a WWC Quality Assurance reviewer, and an external peer reviewer.

[4] 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 child-level improvement indices for all findings across the studies.

[5] The national PCER Consortium (2008) study conducted a rigorous efficacy evaluation of 14 preschool curricula. Twelve research teams implemented one or two curricula in preschool settings serving predominantly low-income children using an experimental design. For each team, preschools or classrooms were randomly assigned to the intervention curricula or comparison curricula, and the children were followed from prekindergarten through kindergarten. The studies each used a common set of measures with the cohort of children beginning preschool in the summer/fall of 2003. PCER Consortium (2008) summarized the details and results of each curriculum study.

[6] The number of classrooms (79 overall) is from the author query response dated June 13, 2012. Assel et al. (2007) and PCER Consortium (2008, Chapter 6) state that there were 76 classrooms overall.

[7] Information provided by the study authors, at the WWC's request.

[8] Information provided by the study authors, at the WWC's request.

[9] Information provided by the study authors, at the WWC's request.

[10] Information provided by the study authors, at the WWC's request.

[11] For the rating of effectiveness in this WWC intervention report, the WWC includes only the results comparing the *Let's Begin with the Letter People*® group to the business-as-usual comparison group. The WWC includes the *Waterford Early Reading*™ *Level One* versus the business-as-usual comparison in a separate *Waterford Early Reading*™ *Level One* intervention report. Both intervention groups used the studied intervention in conjunction with the *High/Scope*® curriculum, which was the standard curriculum used by the classrooms prior to the study.

[12] Information provided by the study authors, at the WWC's request.

[13] The previous review of Fischel et al. (2007) focused on the subset of new randomly-assigned classrooms and obtained data on those classrooms from an author query. This review focused on all classrooms analyzed for the study and was reviewed using WWC Evidence Standards (version 2.1), as described in the Early Childhood Education review protocol (version 2.0).

[14] The previous version of this intervention report included the Book Knowledge and Comprehension subtests from the storybook assessment developed for the Head Start FACES study. However, the authors report internal consistency values for these two measures that are below the minimum value for an outcome to be considered marginally acceptable under the Early Childhood Education review protocol (version 2.0).

[15] PCER Consortium (2008, p. 88) reported that eight children were selected from each classroom. In response to a query, the study authors noted that eight children were randomly selected for the site-specific study; however, only seven children were randomly selected for the PCER Consortium study.

[16] Information provided by the study authors, at the WWC's request.

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

Recommended Citation

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, June). Early Childhood Education intervention report: *Let's Begin with the Letter People®*. Retrieved from http://whatworks.ed.gov

WWC Rating Criteria

Criteria used to determine the rating of a study

Study rating	Criteria
Meets WWC evidence standards without reservations	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
Meets WWC evidence 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 evidence 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 evidence 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 students, the improvement index represents the gain or loss of the average student due to the intervention. As the average student starts at the 50th percentile, the measure ranges from –50 to +50.

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 subjects 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 investigators randomly assign eligible participants into 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.

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% ($p < 0.05$).

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

Please see the WWC Procedures and Standards Handbook (version 2.1) for additional details.