IES Institute of Education Sciences

What Works Clearinghouse

Using Bayesian Meta-Analysis to Explore the Components of Early Literacy Interventions APPENDICES

September 2023
WWC 2023008
U.S. DEPARTMENT OF EDUCATION

A Publication of the National Center for Education Evaluation (NCEE) at IES

CONTENTS

TABLES

FIGURES

APPENDIX A. COMPONENTS OF EARLY LITERACY INTERVENTIONS

A.1. Details of the component coding process

The synthesis team used the coding protocol (Appendix E) to code the components of interventions. The coding protocol includes 81 components nested within 15 component domains, which were nested within five component types. A domain was coded for an intervention if any component under the domain was implemented. A type was coded for an intervention if any domain under the type was implemented. The main findings in this report focus on the 15 component domains because the number of studies in the analysis was too small to statistically differentiate the 81 components. The coding protocol with definitions of the components, domains, and types is included as Appendix E.

The coding protocol was built on a pilot taxonomy for coding components of beginning reading interventions developed for the Institute of Education Sciences on the Technical Assistance Supporting Evidence-Building and Use (TASEBU) project under contract number 91990020F0052 (Scher & Martinez, 2022) and tailored for this report in the following ways:

1. Modifications include adding a domain for building writing skills to the instructional practices, structures, and academic supports component type to reflect a broader early literacy focus. This new domain includes components that focus on writing quality and writing conventions.
2. The original TASEBU taxonomy included component approaches as a level of detail within individual components, but coding this level of detail was not feasible for this meta-analysis. Therefore, the modified protocol lists these component approaches as examples of how components might be implemented to aid accurate coding.
3. The synthesis team added an open-ended component option to each component domain to reflect the possibility that the studies may implement component domains in unanticipated ways.
4. The team removed sections to code the focal population and topic areas because all studies in the analysis are intended for students in grades K–3 and focused on early literacy.
5. The team modified the intervention delivery sections, although the data collected in these sections are not used in the meta-analysis.
6. The team adapted the instructions for using the coding form.

The focus of the coding was to determine the components of the intervention as it was implemented in each study. For example, because interventions are not always implemented the same way across studies, two studies of the same intervention might have implemented different components. The studies were independently coded by two trained coders. A reconciler worked to resolve discrepancies between the coded forms.

Because studies usually did not provide sufficient detail to accurately code the intervention's components, the synthesis team contacted study authors for additional information about how the intervention was implemented in their study. About half of study authors responded with this information. The team followed up with study authors with reminders and granted extensions when requested. Coders reviewed the information provided by study authors, but they generally prioritized information provided by the study authors when determining whether a component was present. Follow-ups were sent to study authors to resolve conflicting or unclear information for four studies of two interventions.

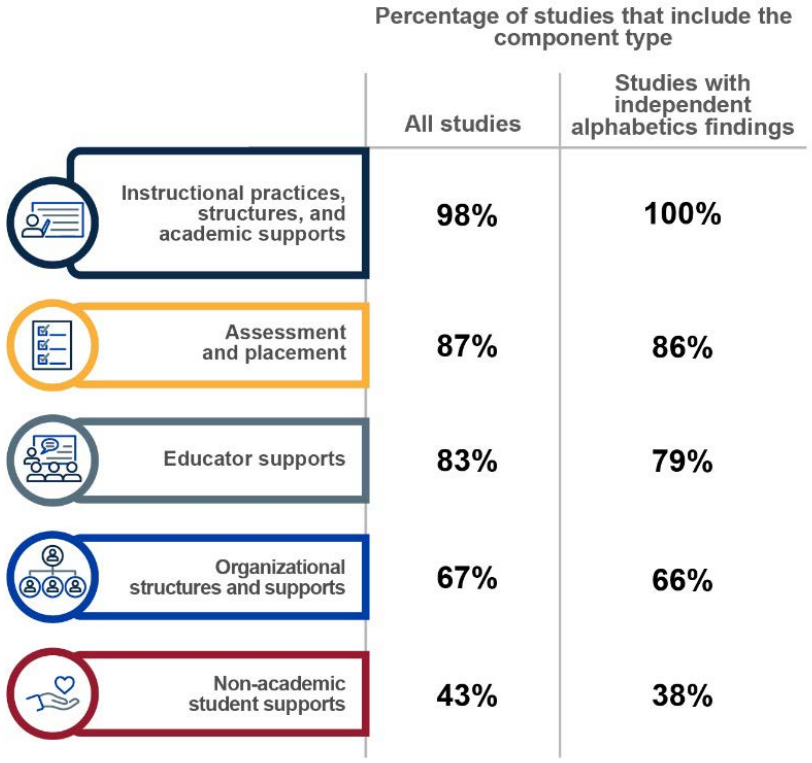
Although the coding process made the best use of the available information, significant uncertainty remains about which components were included for many of the interventions. This uncertainty likely decreases the precision with which the relationships between components and intervention impacts can be estimated. For example, not all study authors responded with information about their studies, and study authors were sometimes uncertain about whether some components were implemented. The synthesis team initially coded components in a way that reflected uncertainty on the presence of components. To do this, coders indicated how confident they were that a component was present. Coders could indicate levels of confidence corresponding to probabilities of 0, 0.25, 0.50, 0.75, or 1 that a component was present. Ultimately, these probabilities were recoded to eliminate the uncertainty, for three reasons. First, simulations of the meta-analysis showed that for the purpose of identifying relationships between the components and impacts, it is better for coders to err on the side of overconfidence (that is, to report probabilities of 0 or 1) than underconfidence (reporting a probability of 0.50). Second, when coders believed a component was likely present, they were highly confident that it was there (that is, coders reported probability 1 much more often than they reported probability 0.75). Third, when coders initially reported probabilities of 0.5 or 0.25, these tended to be revised down to 0 after querying the authors. For these reasons, only components coded with a probability of 1 were considered present in the interventions.

A.2. Implementation of early literacy component domains

The study team coded components in 54 studies of the effectiveness of early literacy interventions. These studies include 29 that are the focus of the results in the report–those that examined impacts on independent alphabetics outcomes–and another 25 that examined impacts on other literacy outcomes. All 54 studies are included in the Bayesian meta-analysis model.

The early literacy interventions examined in this report are generally complex, with most studies including multiple types of components. The types included in the 29 studies that examined impacts on independent alphabetics outcomes are similar to those included in the full sample of 54 studies. For the studies that examined findings for independent alphabetics outcome measures, Figure 4 in the report shows how common each of the five component types are. Figure A1 shows how common each of the component types are for the full sample of 54 studies along with the same percentages for the main sample from Figure 4.

Figure A1. Implementation of component types in early literacy interventions



Notes: The first column in this figure includes the 54 studies of early literacy interventions included in the meta-analysis, and the second column includes the 29 studies of early literacy interventions that examined impacts on independent alphabetics outcomes.

Source: Authors' coding of components in early literacy studies.

Similarly, the implementation of the individual components in the main sample of 29 studies is similar to how they were implemented in all 54 studies. Tables 1-5 in the report describe how common each of the 15 component domains and selected components are in the 29 studies with independent alphabetics outcomes. The tables that follow report how common the component domains and components are in the full sample of 54 studies. The tables also include less commonly implemented components that were not included in Tables 1-5, for both the sample of 29 studies with independent alphabetics outcomes and for the full sample of 54 studies. Due to the relatively small number of studies compared to the number of components, this synthesis is not able to statistically differentiate among the 81 components.

A.2.1. Instructional practices, structures, and academic supports

Instructional practices to build alphabetics skills

The domain for instructional practices to build alphabetics skills includes 13 components. This component domain was implemented in 89 percent of the studies in the full sample and all 29 of the studies with findings for independent alphabetics outcome measures. Table A1 reports how often each component was observed in studies that implement any component in this component domain.

Table A1. Implementation of the instructional practices to build alphabetics skills component domain

Component	Percentage of studies that have the component	
	Among 48 studies with any component in this domain	Among 29 studies with any component in this domain and independent alphabetics findings
Engaging in explicit/direct instruction	92	90
Teaching letter names and letter sound relations	83	90
Teaching students to recognize and manipulate segments of sound in speech	79	83
Teaching blending and sound-spelling patterns in words	73	79
Linking letter-sound relationship knowledge to phonemic awareness	73	76
Reading decodable words in isolation and in text	71	76
Teaching to read regular and irregular high-frequency words	65	66
Introducing common sound-spelling patterns	63	62
Teaching to recognize common word parts	56	52
Introducing phonetically irregular words	52	62
Handwriting programs to call attention to letter shape	35	34
Working with rhyming words	33	31
Other instructional practices to build alphabetics skills	15	21

Instructional practices to build reading accuracy and fluency skills

The domain for instructional practices to build reading accuracy and fluency skills includes five components. This component domain was implemented in 76 percent of the studies in the full sample and 79 percent of the studies with findings for independent alphabetics outcome measures. Table A2 reports how often each component was observed in studies that implement any component in this component domain.

Table A2. Implementation of instructional practices to build reading accuracy and fluency skills component domain

Component	Percentage of studies that have the component	
	Among 41 studies with any component in this domain	Among 23 studies with any component in this domain and independent alphabetics findings
Providing opportunities for oral reading practice with feedback	95	91
Supporting oral reading through modeling, scaffolding, and feedback	83	91
Teaching self-monitoring and self-correcting reading skills	83	91
Encouraging daily reading of connected text	66	70
Other instructional practices to build reading accuracy and fluency skills	17	26

Instructional practices to build comprehension skills

The domain for instructional practices to build comprehension skills includes 12 components. This component domain was implemented in 78 percent of the studies in the full sample and 76 percent of the studies with findings for independent alphabetics outcome measures. Table A3 reports how often each component was observed in studies that implement any component in this component domain.

Table A3. Implementation of instructional practices to build comprehension skills component domain

Component	Percentage of studies that have the component	
	Among 41 studies with any component in this domain	Among 22 studies with any component in this domain and independent alphabetics findings
Teaching how to use a specific reading comprehension strategy	67	64
Guiding students through focused discussion on the meaning of the text	62	64

Component	Percentage of studies that have the component	
	Among 41 studies with any component in this domain	Among 22 studies with any component in this domain and independent alphabetics findings
Purposefully selecting texts for comprehension development and knowledge-building	62	55
Creating opportunities for students to see themselves as successful readers	60	50
Explicit instruction with gradual release of responsibility	57	55
Helping students discover the purpose and benefits of reading	48	45
Providing peer collaboration/cooperative learning opportunities in reading comprehension activities	48	41
Modeling and explaining how to identify and use a text's organizational structure	45	41
Giving students reading choices	36	36
Metacognitive awareness for comprehension monitoring	33	32
Using multiple-strategy formats	24	27
Other instructional practices to build comprehension skills	10	14

Instructional practices to support vocabulary and relevant background knowledge development

The domain for instructional practices to support vocabulary and relevant background knowledge development includes five components. This component domain was implemented in 74 percent of the studies in the full sample and 66 percent of the studies with findings for independent alphabetics outcome measures. Table A4 reports how often each component was observed in studies that implement any component in this component domain.

Table A4. Implementation of instructional practices to support vocabulary and relevant background knowledge development component domain

Component	Percentage of studies that have the component	
	Among 40 studies with any component in this domain	Among 19 studies with any component in this domain and independent alphabetics findings
Providing instruction to build relevant vocabulary and background knowledge	90	89
Teaching academic vocabulary in context	68	74
Engaging in conversations to support inferential knowledge	65	68
Engaging students in developing narrative language skills	53	47
Other instructional practices to support vocabulary and relevant background knowledge development	10	5

Instructional practices to build writing skills

The domain for instructional practices to support writing skills includes three components. This component domain was implemented in 41 percent of the studies in the full sample and 34 percent of the studies with findings for independent alphabetics outcome measures. Table A5 reports how often each component was observed in studies that implement any component in this component domain.

Table A5. Implementation of instructional practices to build writing skills component domain

Component	Percentage of studies that have the component	
	Among 22 studies with any component in this domain	Among 10 studies with any component in this domain and independent alphabetics findings
Providing instruction to develop writing quality	68	60
Teaching writing conventions	50	40
Other instructional practices to build writing skills	32	50

A.2.2. Non-academic student supports

Social-emotional learning strategies

The domain for social-emotional learning strategies includes three components. This component domain was implemented in 31 percent of the studies in the full sample and 31 percent of the studies with findings for independent alphabetics outcome measures. Table A6 reports how often each component was observed in studies that implement any component in this component domain.

Table A6. Implementation of social-emotional learning strategies component domain

	Percentage of studies that have the component	
Component	Among 17 studies with any component in this domain	Among 9 studies with any component in this domain and independent alphabetics findings
Teaching self-management skills	76	89
Support development of a growth mindset	71	56
Other social-emotional learning strategies	24	22

Parent outreach and involvement

The domain for parent outreach and involvement includes four components. This component domain was implemented in 30 percent of the studies in the full sample and 21 percent of the studies with findings for independent alphabetics outcome measures. Table A7 reports how often each component was observed in studies that implement any component in this component domain.

Table A7. Implementation of parent outreach and involvement component domain

	Percentage of studies that have the component	
Component	Among 16 studies with any component in this domain	Among 6 studies with any component in this domain and independent alphabetics findings
Parent meetings and conferences to discuss learning and growth	56	50
Programming to encourage parent involvement in reading activities	50	50
Providing books to families	50	50
Other parent outreach and involvement	44	22

A.2.3. Organizational structures and supports

Programming to support improvements in school climate

The domain for programming to support improvements in school climate includes four components. This component domain was implemented in four studies, all of which examined impacts on independent alphabetics outcome measures. These four studies represent 7 percent of the studies in the full sample and 14 percent of the studies with findings for independent alphabetics outcome measures. Table A8 reports how often each component was observed in studies that implement any component in this component domain. We did not observe any anti-bullying programming in the studies, or any other school climate component not otherwise covered by the component taxonomy.

Table A8. Implementation of programming to support improvements in school climate component domain

Component	Percentage of studies that have the component	
	Among 4 studies with any component in this domain	Among 4 studies with any component in this domain and independent alphabetics findings
Schoolwide efforts to promote positive behaviors	100	100
Implementing multitiered systems of support around school climate and behaviors	25	25
Implementing anti-bullying programming	0	0
Other programming to support improvements in school climate	0	0

Providing out-of-school time supports

The domain for providing out-of-school time supports includes five components. This component domain was implemented in 17 percent of the studies in the full sample and 17 percent of the studies with findings for independent alphabetics outcome measures. Table A9 reports how often each component was observed in studies that implement any component in this component domain.

Table A9. Implementation of providing out-of-school time supports component domain

Component	Percentage of studies that have the component	
	Among 9 studies with any component in this domain	Among 5 studies with any component in this domain and independent alphabetics findings
Summer programming	67	40
Partnering with community organizations	44	40
Tutoring outside of school time	33	60
After school programming	33	40
Other provision of out-of-school time supports	0	0

Providing opportunities for reducing the ratio of students to teachers

The domain for opportunities for reducing the ratio of students to teachers includes four components. This component domain was implemented in 52 percent of the studies in the full sample and 52 percent of the studies with findings for independent alphabetics outcome measures. Table A10 reports how often each component was observed in studies that implement any component in this component domain.

Table A10. Implementation of opportunities for reducing the ratio of students to teachers component domain

Component	Percentage of studies that have the component	
	Among 28 studies with any component in this domain	Among 15 studies with any component in this domain and independent alphabetics findings
Scheduling small-group literacy blocks	89	93
Using trained assistants or paraprofessionals in literacy blocks	64	67
Reducing class sizes	18	33
Other opportunities for reducing the teacher/student ratio	7	7

A.2.4. Educator supports

Professional development for teachers

The domain for professional development for teachers includes five components. This component domain was implemented in 76 percent of the studies in the full sample and 76 percent of the studies with findings for independent alphabetics outcome measures. Table A11 reports how often each component was observed in studies that implement any component in this component domain.

Table A11. Implementation of professional development for teachers component domain

Component	Percentage of studies that have the component	
	Among 41 studies with any component in this domain	Among 22 studies with any component in this domain and independent alphabetics findings
Supporting instructional practices	93	91
Supporting the link between student assessment and practice	49	55
Training in technology use	32	45
Utilizing literacy coaches	32	36
Other teacher professional development related to the intervention	24	36

Implementing professional learning communities

The domain for implementing professional learning communities includes five components. This component domain was implemented in 19 percent of the studies in the full sample and 21 percent of the studies with findings for independent alphabetics outcome measures. Table A12 reports how often each component was observed in studies that implement any component in this component domain.

Table A12. Implementation of professional learning communities component domain

Component	Percentage of studies that have the component	
	Among 10 studies with any component in this domain	Among 6 studies with any component in this domain and independent alphabetics findings
Building grade-level professional learning communities	50	50
Providing common planning/prep time	40	50

Component	Percentage of studies that have the component	
	Among 10 studies with any component in this domain	Among 6 studies with any component in this domain and independent alphabetics findings
Developing cross-functional professional learning communities that meet on a regular basis	30	33
Building multigrade-level vertical teams to support curricular alignment	20	33
Other professional learning community activities related to the intervention	20	0

Providing instructional support materials

The domain for providing instructional support materials includes five components. This component domain was implemented in 69 percent of the studies in the full sample and 69 percent of the studies with findings for independent alphabetics outcome measures. Table A13 reports how often each component was observed in studies that implement any component in this component domain.

Table A13. Implementation of providing instructional support materials component domain

Component	Percentage of studies that have the component	
	Among 37 studies with any component in this domain	Among 20 studies with any component in this domain and independent alphabetics findings
Accessing and using sample lesson plans provided by developer	81	85
Accessing and using curricular guides and developer handbooks	68	70
Accessing and using instructional logs	46	50
Accessing and using sample lesson plans provided by instructional staff	38	40
Other instructional support materials	24	30

A.2.5. Assessment and placement

Testing and screening

The domain for testing and screening includes four components. This component domain was implemented in 81 percent of the studies in the full sample and 83 percent of the studies with findings for independent

alphabetics outcome measures. Table A14 reports how often each component was observed in studies that implement any component in this component domain.

Table A14. Implementation of testing and screening component domain

Component	Percentage of studies that have the component	
	Among 44 studies with any component in this domain	Among 24 studies with any component in this domain and independent alphabetics findings
Implementing universal screening for students in grades K-3	70	75
Formative assessments through curriculum-based measurement	61	71
Using centralized software platforms to monitor student progress	27	21
Other testing and screening	16	8

Student placement

The domain for student placement includes four components. This component domain was implemented in 67 percent of the studies in the full sample and 66 percent of the studies with findings for independent alphabetics outcome measures. Table A15 reports how often each component was observed in studies that implement any component in this component domain.

Table A15. Implementation of student placement component domain

Component	Percentage of studies that have the component	
	Among 36 studies with any component in this domain	Among 19 studies with any component in this domain and independent alphabetics findings
Grouping small literacy groups based on homogeneous student skill level	78	95
Implementing a multitiered system of support or response-to-intervention framework to identify students in need of different levels of support	58	53
Providing literacy instruction to heterogeneous skill groups	19	21
Other student placement	14	11

APPENDIX B. DATA FROM THE WHAT WORKS CLEARINGHOUSE'S DATABASE OF REVIEWED STUDIES

The primary data source was the What Works Clearinghouse's (WWC's) database of reviewed studies, which the synthesis team downloaded from the WWC website (https://ies.ed.gov/ncee/wwc/StudyFindings) in May 2021.

B.1. About the WWC

The WWC's study review process is thoroughly documented in the WWC Standards and Procedures Handbook, currently in version 5.0 (WWC, 2022), including how studies are identified for WWC review, criteria for which studies are eligible for review, the research standards used to review studies, and ways the WWC reports findings and synthesizes evidence. Previous versions of WWC handbooks are also on the WWC website.

The meta-analysis includes high-quality studies based on WWC research standards

The WWC standards assign one of three research ratings to each study based on the quality of the research design used to measure the impact of an intervention. Studies that receive one of the highest two research ratings are included in this synthesis:

- ***Meets WWC Standards Without Reservations***. This is the highest possible rating, which the WWC gives to studies that can provide a high degree of confidence that the intervention caused the measured impacts. For example, randomized controlled trials, a study design in which study participants are randomly assigned to receive the intervention or to a comparison group that does not receive the intervention, can receive this rating.
- ***Meets WWC Standards With Reservations***. Studies receiving this rating provide a lesser degree of confidence that the intervention caused the measured impacts. The WWC has reservations about some randomized controlled trials when they are executed with certain flaws and about all quasi-experimental studies, in which the intervention and comparison conditions are not formed by randomly assigning study participants.

Studies that receive the lowest WWC rating are not included in this synthesis:

- ***Does Not Meet WWC Standards***. This is the lowest possible rating, which the WWC gives to studies that provide a low degree of confidence that the intervention caused the measured impacts. Studies with major flaws in their research designs receive this rating.

B.2. Inclusion criteria for studies

The synthesis team retained all 54 studies of early literacy interventions in the WWC database–including the 29 studies that examined impacts on independent alphabetics outcomes that are the focus of the report–that met the seven eligibility criteria for this meta-analysis. To be included in the meta-analysis, the studies must:

1. Include students in grades K-3 (ages 5-8) but may also include students in other grades
2. Measure the impact of an intervention specifically designed to improve student literacy skills. For example, this meant excluding interventions that focused on improving a broader set of skills, such as *Teach for America*, *eMINTS*, *Teacher Advancement Program*, and charter schools
3. Be published within the past 10 years, so that the studies include those with the most relevant context to today's educational environment

4. Report at least one finding for the intervention's impact on literacy outcomes
5. Include at least one calculated impact estimate (effect size)
6. Meet WWC standards with or without reservations under version 2.1 (or later) of the WWC procedures and standards

The analysis is limited to interventions and studies contained in the WWC database, which are those the WWC has rigorously studied and reviewed. This may affect the generalizability of the findings from this synthesis because (1) only early literacy interventions that have been rigorously studied are represented in the analysis, but these interventions may not be representative of all early literacy interventions; (2) among the rigorous studies of early literacy interventions, some have not been reviewed by the WWC and would therefore not be represented in the findings; and (3) some studies of early literacy interventions that have been conducted are never published. The meta-analysis model includes features to address the second and third issues (see Appendix C).

B.3. Preparing the WWC data for analysis

To address some missing information in the WWC database on the 54 studies included in the meta-analysis, the synthesis team added some intervention names and grade levels. Also, the number of clusters was incorrectly reported for three studies. The synthesis team obtained the correct numbers from the WWC's original study review records.

The 54 studies included 326 findings, but the synthesis team excluded 85 findings because (1) it was not possible to calculate an effect size and standard error for them, (2) they were for subgroups of students in grades other than K–3, or (3) they were for subgroups of students other than grade level, such as race, ethnicity, or gender. For example, a study might report both a full-sample finding and findings for students by subgroup. In this case, the meta-analysis would include only the full-sample finding. This report's findings reflect those for the samples included in the studies. It is not possible to make precise statements about how effective the interventions or their components are for different groups of students because studies do not always provide subgroup findings for these student groups.

The synthesis team classified all outcome measures into the outcome domains in the Study Review Protocol that the WWC has used to review studies under the WWC Standards and Procedures Handbook, version 4.1. The Study Review Protocol describes each outcome domain. The synthesis team reclassified outcome measures in some studies into domains from the Study Review Protocol, because WWC reviews of studies conducted before this review protocol used different sets of domains. The synthesis team also incorporated into the study database a data element differentiating between outcome measures developed by intervention developers, by researchers, and independently of the developer or researcher (Walsh et al., 2023). The dataset and code used for the analysis are available on the report website.

B.4. Calculating effect sizes and standard errors to use in the meta-analysis

The synthesis team used the effect sizes and their standard errors to estimate the Bayesian meta-analytic model.

Impacts of interventions are measured using the Hedges' *g* effect size, which is a standardized measure of the impact of an intervention that can be synthesized across outcome measures and studies. For the outcome measures in this synthesis, the impact estimates are measured in units of student-level standard deviations. The

synthesis team performed some calculations to address missing or incorrect effect sizes. For three of the 241 findings, the WWC improvement index had an opposite sign from the effect size (for example, an improvement index of 3 and an effect size of -0.08). In these cases, the outcome measure is reverse-coded so that more negative values on the outcome measure scale are more favorable. Therefore, the synthesis team reversed the sign of the effect size. Next, the WWC-calculated effect size was missing for 36 of the 241 findings, so the synthesis team instead used the study-calculated effect size.

Because the WWC database did not include standard errors for the effect sizes, the synthesis team instead used p-values and sample sizes from the WWC database to calculate standard errors for each effect size to use in the meta-analysis. For 192 findings with a WWC-calculated p-value that was positive from individual-level assignment studies, the synthesis team first calculated a z-statistic based on the standard normal distribution and the p-value. Next, to obtain the standard error, the team divided the reported effect size by the absolute value of the z-statistic. For a few findings with p-values very close to 1, this calculation resulted in implausible standard errors, which were discarded. These cases were included with other findings from individual-level assignment studies with no WWC-calculated p-value or a p-value of 0. For these 17 findings, the synthesis team calculated the standard error using the individual-level assignment formula in E.2.2 from the Supplement to the What Works Clearinghouse Procedures Handbook, version 4.1, with the R2 assumed to be 0. The remaining 32 findings were from cluster-level assignment studies. To address some errors in the p-values for these findings, the synthesis team calculated all standard errors in these studies using the cluster-level assignment formula in E.2.2 from the same supplement, with the R2 assumed to be 0.

B.5. Overview of the WWC study data

The report includes findings based on the 29 studies with findings for independent alphabetics outcomes, but the data used to estimate the meta-analytic model also include findings from an additional 25 studies that report findings for other types of literacy outcome measures, including those that were created by the developers of the intervention or by the study authors and outcome measures in literacy domains other than alphabetics.

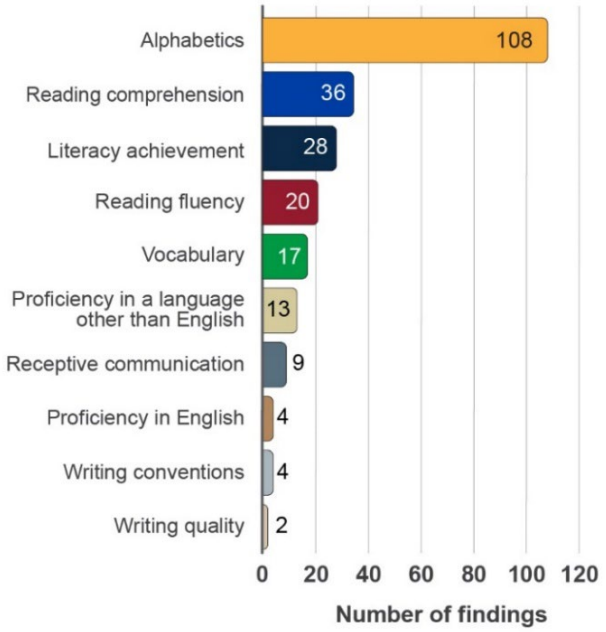
The studies examined a broad range of outcome measures spanning 10 different literacy domains (Figure B1). The rationale for focusing the findings on studies with independent outcome measures in the alphabetics domain is that (1) the associations between some component domains and intervention impacts may vary for different literacy outcome domains and (2) there were too few findings in other outcome domains besides alphabetics to precisely distinguish different associations for each domain. Collectively, the studies most frequently analyzed findings in the alphabetics domain (45 percent), which include phonics, phonemic awareness, phonological awareness, and letter identification. The next most common type of outcome measure was reading comprehension (15 percent). The greater focus on alphabetics outcomes in the studies reflects the emphasis on alphabetics skills in the interventions: 89 percent of the studies include a component related to building alphabetics skills. The WWC Study Review Protocol provides more information about each literacy domain.

Of the 241 findings shown in Figure B1, 83 percent are on measures independent of the study authors or intervention developers, 9 percent were developed by intervention developers, and 8 percent were developed by researchers. The independent measures are typically standardized measures designed to have broader application beyond the specific intervention. The prevalence of outcome domains for the 241 findings in the study data, including both independent and non-independent outcome measures, shown in B1, are similar to those for the 200 findings in 49 studies with independent outcome measures.

The 54 studies of early literacy interventions included in the meta-analysis were published between 2011 and 2019. About three-quarters of the studies are indexed in the Education Resources Information Center (ERIC) database. Study sample sizes ranged from 22 students to 22,583 students. When possible, the synthesis team included findings that included only students in grades K-3, but the reviewed findings did not always disaggregate by grade level. Therefore, some findings also included students in grades 4 and 5.

Thirty-seven studies meet WWC standards without reservations and 17 studies meet WWC standards with reservations. Almost 40 percent of studies were reviewed as part of WWC systematic reviews, including for practice guides (33 percent), intervention reports (7 percent), and one study included in a rapid review on distance learning interventions. Other studies were reviewed by the WWC because they were funded by IES or cited in IES grant competitions, could inform IES performance measures (such as the number of studies funded by IES that meet WWC standards), or were identified for review as single study reviews because the WWC believed they would be of particular interest or for other reasons. Because 60 percent of studies were not reviewed as part of WWC systematic review efforts, the study sample may include a higher percentage of studies with favorable findings than would have otherwise been identified and may not be representative of all studies on all literacy interventions. The meta-analysis attempts to control for possibility of the higher percentage of studies with favorable findings.

Figure B1. Number of findings by outcome domain in early literacy studies



Notes: The figure includes 241 findings from 54 studies of early literacy interventions.

Source: Authors' calculations based on data from the WWC's database of reviewed studies.

The studies examined the effectiveness of a combined 45 interventions. In most cases, the early literacy interventions were examined in just one study in the analysis, but three interventions–*Early Reading Intervention*, *individualized student instruction*, and *Spheres of Proud Achievement in Reading for Kids*–were each examined in three studies, and three other interventions–*Targeted Reading Intervention*, *Reading Recovery*, and *Success for All*–were each examined in two studies. Forty-seven percent of the interventions are unbranded practices or bundles of components. Table B1 describes the studies, interventions, participants, and outcomes that are included in the study data set. The table also indicates with an asterisk on the citation whether the study is among the 29 that examined impacts on independent alphabetics outcomes, which are the focus of the report. The references include the full citations for these studies and links to the WWC study page for each.

Table B1. Studies of early literacy interventions included in the meta-analysis

Study	Intervention	WWC rating	WWC purpose of review	Grade levels	Number of students	Number of findings
Amendum et al., 2011	Targeted Reading Intervention	MSWR	Practice guide	K-1	167	1
Baker et al., 2017*	GraphoGame Spanish	MSWOR	Single study review	1	78	4
Baker et al., 2013	Read aloud intervention	MSWOR	Practice guide	1	123	4
Baker et al., 2015*	Transition lessons	MSWOR	Single study review	1	78	6
Borman et al., 2019	Descubriendo La Lectura	MSWOR	Single study review	1	142	15
Buckingham et al., 2012*	Meeting Initial Needs In Literacy	MSWR	Practice guide	2	22	4
Cann et al., 2015	SmartStart K-3 Plus program	MSWOR	Department-funded evaluation	3	1,461	7
Case et al., 2014*	Tier 2 reading instruction	MSWOR	Practice guide	1	123	5
Chambers et al., 2011*	Small-group computer-assisted tutoring	MSWR	Practice guide	1	280	3
Cheatham et al., 2014*	Independent practice of multiple-criteria text	MSWR	Practice guide	2	62	2
Connor, Morrison, Fishman, et al., 2011	Individualized student instruction	MSWR	Grant competition	3	448	1
Connor, Morrison, Schatschneider, et al., 2011*	Individualized student instruction	MSWOR	Grant competition	1	396	1
Connor et al., 2013*	Individualized student instruction	MSWOR	Grant competition	1-3	568	6
Coyne et al., 2013*	Early Reading Intervention	MSWR	Practice guide	K	162	5
Coyne et al., 2018	Extended vocabulary instruction	MSWOR	Department-funded evaluation	K	1,440	3
Denton et al., 2013*	Responsive reading instruction (adaptation)	MSWOR	Practice guide	2	72	5
Dombek et al., 2017	Content-Area Literacy Instruction	MSWOR	IES performance measure	K-3	418	3
Duff et al., 2012	Reading Intervention	MSWR	Practice guide	K	59	1
Duff et al., 2014	Reading and Language Intervention	MSWOR	Practice guide	K	52	4

Study	Intervention	WWC rating	WWC purpose of review	Grade levels	Number of students	Number of findings
Fuchs et al., 2017 (COMP)	Reading comprehension	MSWOR	Department-funded evaluation	3	40	1
Fuchs et al., 2017 (WM COMP)	Reading comprehension plus working memory training	MSWOR	Department-funded evaluation	3	39	1
Gilbert et al., 2013*	RTI prevention model	MSWOR	Practice guide	1	212	4
Gunn et al., 2011*	Read Well Kindergarten	MSWR	Practice guide	K	1,405	7
Hagans & Good, 2013*	Phonological awareness intervention	MSWOR	Practice guide	1	50	1
Hill & Lenard, 2016	Achieve3000®	MSWR	Intervention report	2-5	22,583	4
Hooper et al., 2013	Process Assessment of the Learner	MSWOR	Single study review	3	129	1
Jacob et al., 2015*	Reading Partners	MSWOR	Single study review	2-3	602	3
Johnson et al., 2017	Dialect Awareness	MSWOR	Department-funded evaluation	2-4	76	3
Jones, 2014	Spheres of Proud Achievement in Reading for Kids	MSWOR	Department-funded evaluation	K-2	529	1
Jones, 2015*	Spheres of Proud Achievement in Reading for Kids	MSWOR	Grant competition	K-2	387	6
Jones & Lander, 2014	Spheres of Proud Achievement in Reading for Kids	MSWOR	Grant competition	K-2	273	1
Kloos et al., 2019	MindPlay Virtual Reading Coach	MSWR	Distance learning rapid review	2	94	2
Konstantopoulos, 2018*	HEROES	MSWR	Grant competition	K-4	490	2
Kuchle et al., 2018*	Pennsylvania dyslexia screening and early literacy intervention	MSWR	Grant competition	K	5,570	6
Little et al., 2012*	Early Reading Intervention	MSWR	Practice guide	K	90	5
Madden & Slavin, 2017	Tutoring With Alphie	MSWR	Grant competition	1-2	317	2
May et al., 2013	Reading Recovery®	MSWOR	Single study review	1	866	1
May et al., 2016	Reading Recovery®	MSWOR	Grant competition	1	6,888	4

Study	Intervention	WWC rating	WWC purpose of review	Grade levels	Number of students	Number of findings
Nakamoto et al., 2015	Arts for Learning Lessons Project	MSWOR	Department-funded evaluation	2, 3	3,074	4
Quint et al., 2015*	Success for All®	MSWOR	Intervention report	K-4	8,762	14
Ransford-Kaldon et al., 2013	Leveled Literacy Intervention	MSWOR	Intervention report	K-2	320	10
Ross et al., 2017*	Tutoring with the Lightning Squad	MSWOR	Grant competition	1-3	150	3
Savage et al., 2013*	ABRACADABRA Web-Based Literacy Program	MSWOR	Practice guide	K-2	74	8
Schechter et al., 2015	Blended learning	MSWOR	Grant competition	1-2	83	3
Simmons et al., 2011*	Early Reading Intervention	MSWR	Practice guide	K	206	11
Slavin et al., 2011*	Transitional bilingual education	MSWOR	Grant competition	1-3	247	12
Smith et al., 2016*	Enhanced Core Reading Instruction (ECRI)	MSWOR	IES performance measure	1	764	9
Steele et al., 2017	Dual-language immersion programs	MSWOR	IES performance measure	1-3	1,625	4
Tracey et al., 2014*	Success for All®	MSWR	Intervention report	1	886	5
Tse & Nicholson, 2014*	Phonics-enhanced Big Book reading	MSWOR	Practice guide	2	48	3
Vadasy & Sanders, 2010*	Supplemental phonics-based instruction	MSWOR	Practice guide	1-3	89	5
Vernon-Feagans et al., 2013*	Targeted Reading Intervention	MSWR	Grant competition	K-1	630	11
Wood et al., 2018	Bridging for Language Outcomes in the Classroom	MSWOR	Department-funded evaluation	K-1	258	2
Zvoch & Stevens, 2013*	Summer literacy intervention	MSWOR	Single study review	K-1	93	2

Notes: In addition to the reason for review listed in the table, the WWC conducted separate reviews of Quint et al. (2015) for a department-funded evaluation and Vernon-Feagans et al. (2013) for a grant competition. For each of these studies, the data set for the meta-analysis includes findings from both reviews because the individual reviews were not comprehensive. The references include full citations for each study and links to their WWC study pages.

Source: Authors' calculations based on data from the WWC database of reviewed studies.

MSWOR = meets WWC standards without reservations; MSWR = meets WWC standards with reservations.

* = examined impacts on independent alphabetics outcomes

APPENDIX C. THE BAYESIAN META-ANALYTIC MODEL

This report applies a synthesis methodology new to the What Works Clearinghouse (WWC) to explore the components included in early literacy interventions in high-quality effectiveness studies reviewed by the WWC. The Bayesian approach is a new way to characterize and report on the evidence for the WWC, but the findings from this meta-analysis are not aligned with definitions of evidence from the U.S. Department of Education. As such, the findings from this report are exploratory. The exploratory Bayesian meta-regression model developed for this synthesis relates the component domains in early literacy interventions to impacts on literacy outcomes. Results from this model provide information about three areas of interest: (1) which interventions examined in high-quality effectiveness studies have positive effects on literacy outcomes? (2) which component domains have positive associations with impacts on literacy outcomes? and (3) to what extent can the component domains examined in this synthesis explain variation in the intervention effects?

The model, and the detailed description that follows, is based on one used in Deke et al. (2022).

C.1. Why use a Bayesian model?

Compared with non-Bayesian meta-analyses, the Bayesian approach uses (1) a more realistic distribution of intervention impacts and (2) a more understandable assessment of uncertainty. Taken together, these advantages mean that the Bayesian model has a greater potential of identifying components that are truly related (though not necessarily *causally* related) to intervention effects compared to a fixed effects meta-regression.

A more realistic distribution of intervention impacts. More realistic models have greater potential to produce accurate findings. Researchers often use simplistic statistical models because more realistic models can be too complex to estimate. For example, it is common in meta-analysis for researchers to assume that intervention impacts (the true, rather than estimated, effects of interventions) follow the bell-shaped normal distribution simply because that assumption makes estimating statistical models easier. Yet it may be reasonable in some contexts to hypothesize that intervention impacts are skewed toward more favorable effects—meaning that although most effects are small, large favorable effects are more common than large unfavorable effects. If that hypothesis is correct, methods that rely on an assumption that intervention impacts follow the normal distribution may yield inaccurate findings.

Bayesian models can accommodate more realistic complexities because they are estimated using a powerful computational algorithm called Markov Chain Monte Carlo (MCMC). Originally developed to solve problems in nuclear physics, MCMC reduces (but does not eliminate) the need to make simplifying assumptions. For example, the meta-regression model in this synthesis is not based on an assumption that intervention effects follow the normal distribution. The synthesis model uses a more flexible distribution (the skewed generalized *t*-distribution) that includes the normal distribution as a special case but also allows for many other possibilities. So, although the model is still based on an assumption, it is a less restrictive assumption.

A more understandable assessment of uncertainty. Savvy decision makers know researchers cannot provide definitive answers and that there will always be some uncertainty due to statistical errors. For example, in an experimental evaluation of a reading tutoring program, it is possible, just by random chance, that students assigned to the intervention group perform better on the outcome due to luck. If the evaluation reports an estimated impact on test scores of 0.15 standard deviations, that does not mean the program definitely improved test scores by 0.15 standard deviations; the true impact might have been higher or lower.

Statistical significance has often been used to communicate confidence in research findings, but the American Statistical Association warns that statistical significance does not mean what most people think it means and can lead to overconfidence in research findings (Greenland et al., 2016; Wasserstein & Lazar, 2016). Specifically, the dichotomous nature of statistical significance can suggest either that something worked or did not work, with little room for nuance. People misinterpret statistical significance to mean that there is a small chance that an impact estimate is due to random chance as opposed to a genuine effect. That is wrong—by itself, statistical significance provides no information about the probability that an impact estimate is due to random chance as opposed to a genuine effect.

With Bayesian methods, researchers can calculate the probability that an intervention genuinely had a favorable effect or that the effect falls within a range that matters to the decision maker. Continuing the tutoring example, it is possible to calculate the probability that the impact on reading test scores was at least greater than zero (or some other value that is relevant to the decision maker), given the estimated impact, standard error, and prior evidence about how common it is for education interventions to have effects of various magnitudes. A decision maker might decide it is worth implementing the program if there is a 90 percent chance that it did no harm and at least a 50 percent chance that it improved test scores by at least 0.10 standard deviations. Without Bayesian methods, researchers cannot report probabilities on the effect of the tutoring program. Because the findings in this report are exploratory, decision makers should not rely on the reported percentages in the findings to make decisions. Instead, this report aims to assess the promise of the new synthesis methodology to potentially provide these recommendations in the future.

C.2. Description of the model

The outcome variable in this model is the impact estimates (effect sizes) reported in studies reviewed by the WWC. The meta-regression model has a multilevel structure with individual impact estimates nested within studies, and studies nested within interventions.

While the report focuses on findings on alphabetics outcomes in 29 studies, the meta-analysis model also draws from 25 additional studies of early literacy interventions with findings on other literacy outcomes. Relationships between component domains and intervention effects on a range of literacy outcomes estimated from these additional studies inform the estimated relationships with measures of alphabetics. In a frequentist analysis, the findings from these 25 additional studies might be excluded from the analysis. However, in a Bayesian analysis, including these additional findings improves the precision of the relationships estimated from the meta-analysis and avoids the problem of multiple comparisons (Berry & Hochberg, 1999; Gelman et al., 2013; Gelman et al., 2012). The Bayesian method uses partial pooling, also known as shrinkage, in which information from multiple sources is combined. The statistical benefits of this approach were once called Stein's Paradox, because the fact that an estimate can be improved using external information can seem counterintuitive (Efron and Morris 1977; Stein 1956).

Using the data described in Appendices A and B, this synthesis models the impact estimates as follows:

$$y_i \sim N\left(\theta_i + \beta_j s_i + \omega_j RES,\ s_i\right) \quad \text{(C.1)}$$

In this equation, y_i is the reported impact estimate (Hedges' *g*) and s_i is the reported standard error, in effect size units, of finding i. The impact estimate is modeled as coming from a normal sampling distribution. The

mean of this distribution has three components, θ_i, $\beta_j s_i$, and $\omega_j RES$, that influence these intervention impact estimates:

- ***True intervention impacts,*** θ_i, are the impacts that y_i seeks to estimate. Statistical errors that influence the impact estimate but do not represent genuine intervention effects are reflected in the standard error of the impact estimate, s_i, which governs the sampling distribution for an impact estimate. The distinction between y_i and θ_i is subtle, and this report uses careful language to refer to these concepts. When the intent is to refer to the estimated impacts, y_i, this report always refers to them as estimated. When the intent is to refer to the true impacts, θ_i, this report always refers to them as impacts.
- ***Potential statistical bias in impact estimates from file-drawer bias*** arises when researchers calculate multiple impact estimates but only report the most favorable estimate. This source of bias is most prevalent in small studies. Smaller studies that report findings with larger standard errors tend to have more favorable impact estimates (see Marks-Anglin & Chen, 2020, for an overview). One possible explanation for this correlation between impact estimate and standard errors is that some researchers may calculate multiple impact estimates but only report the most favorable estimate–a form of reporting bias sometimes referred to as file-drawer bias. Opportunities for selective reporting of findings are greater in smaller studies, which have larger standard errors, because they tend to have more variability across multiple estimates. This behavior, therefore, has the potential to explain why the magnitude of reporting bias in small studies tends to be larger. Using methods described in Deke et al. (2022), the model includes an adjustment, $\beta_j s_i$, for small-study effects. The term β_j is described below.
- ***Potential statistical bias in impact estimates from the quality of the research design.*** The term ω_j reflects bias particular to studies that meet WWC standards with reservations–that is, studies that are randomized controlled trials with one or more concerns that lead to the rating, such as having high attrition, or are quasi-experimental designs. The term RES is an indicator for whether a study meets WWC standards with reservations. The distribution of quasi-experimental bias (ω) is informed by Zurovac et al. (2021) and described in Deke et al. (2022).

A pair of equations model the true impacts (θ):

$$\theta_i = \theta^{Int} + \theta^{Gr}_{g[i]} + \theta^{Outc}_{o[i]} + \theta^{Indep}_{i[i]} + \theta^{Purp}_{p[i]} + \theta^{Inv}_{i} + \theta^{Stu}_{i} + \theta^{Find}_{i} + \theta^{C}_{i} \quad \text{(C.2)}$$

$$\theta^{C}_{i} = \theta^{Type}_{t[i]} + \theta^{Domain}_{d[i]} + \theta^{Type\times Outc\times Indep}_{t[i]o[i]s[i]} + \theta^{Domain\times Outc\times Indep}_{d[i]o[i]s[i]} + \theta^{NType}\sum t[i] + \theta^{NDomain}\sum d[i] + \theta^{NComp}\sum c[i] \quad \text{(C.3)}$$

In equation C.2, θ^{Int} is an intercept representing the impact of an average intervention; $\theta^{Gr}_{g[i]}$ is a set of grade-level random effects; $\theta^{Outc}_{o[i]}$ is a set of outcome domain random effects; $\theta^{Indep}_{i[i]}$ indicates with three categories whether an outcome measure was created by the researcher, the developer, or independently of the researcher and developer; θ^{Purp}_{i} is the purpose of the WWC review; θ^{Inv}_{i} is a set of random effects for each intervention; θ^{Stu}_{i} and θ^{Find}_{i} are sets of random effects capturing the idiosyncratic effect specific to each study and finding; and θ^{C}_{i} represents the contributions of intervention components to intervention impacts.

Because previous literature has found variation in performance gains by outcome and grade (Bloom et al., 2008), the model allows intervention effects, θ_i, to vary by grade, $\theta_{g[i]}^{Gr}$, and outcome domain, $\theta_{o[i]}^{Outc}$. The model also allows both the mean and variance of impacts to vary by a variable, $\theta_{i[i]}^{Indep}$, indicating whether an outcome was created by the researcher, the developer, or independently of the researcher and developer. Impacts measured using developer- and researcher-created outcomes tend to be larger than those for independent outcomes, such as established standardized tests (Wolf & Harbatkin, 2022). The categories used for grade levels, outcome domains, and outcome measure independence, as well as the categories for all indicator variables in the model, are described in Section C.3.

In addition to the file-drawer bias issue discussed above, the model aims to address another source of potential reporting bias that arises because the impact studies include only those the WWC has reviewed, and the WWC has not exhaustively reviewed all early literacy research. For example, the WWC has conducted systematic reviews on a number of topics within early literacy. The WWC will have thoroughly reviewed research on these topics, at least up to the point in time when the review was conducted. But the WWC reviewed 60 percent of the studies in the meta-analysis outside of systematic reviews. It is possible that studies with more favorable findings may have been more likely to have been selected for review. To address this, the analysis allows the purpose of the WWC review, θ_i^{Purp}, to influence intervention impacts. Additionally, some types of review might be more susceptible to selective reporting of findings, so the WWC purpose of review is also included in the file-drawer bias adjustment discussed below.

The set of random effects capturing the overall effect associated with each intervention, θ_i^{Inv}, is intended to measure the impact of common components of interventions that are implemented across studies of the intervention but are unmeasured by the study. The random effects in θ_i^{Stu} and θ_i^{Find} capture the idiosyncratic effect specific to each study and finding, such as characteristics of the study sample or other contextual factors.

The final term in Equation C.2, θ_i^{C}, reflects the influence of the observed intervention components on intervention impacts. This term is modeled in Equation C.3, which includes $\theta_{t[i]}^{Type}$ and $\theta_{d[i]}^{Domain}$, two sets of nested random effects for component types and component domains; $\theta_{t[i]o[i]s[i]}^{Type\times Outc\times Indep}$, which is a set of random effects for combinations of component types, outcome domains, and outcome independence; and $\theta_{d[i]o[i]s[i]}^{Domain\times Outc\times Indep}$, which is a set of random effects for combinations of component domains, outcome domains, and outcome independence. Because the number of findings in most outcome domains is small and because there are relatively few outcome measures created by the researcher or the developer, for $\theta_{t[i]o[i]s[i]}^{Type\times Outc\times Indep}$ and $\theta_{d[i]o[i]s[i]}^{Domain\times Outc\times Indep}$, the outcome domains are placed into three groups and the independence measure is coded dichotomously. This differs from the more flexible coding of outcome domain and independence in Equation C.2. For these interaction terms, independence is coded dichotomously so that researcher-created and developer-created measures are both considered non-independent; one group of outcome domains includes only the alphabetics domain; a second group includes reading comprehension, reading fluency, receptive communication, and vocabulary; and a third group includes literacy achievement, the proficiency domains, and the writing domains. The measures in this third group of outcome domains tend to be broader measures of academic achievement than the measures in the other two groups. Equation C.2 also includes $\theta^{NType}\sum t[i]$,

$\theta^{NDomain} \sum d[i]$, and $\theta^{NComp} \sum c[i]$, where θ^{NType}, $\theta^{NDomain}$, and θ^{NComp} are linear relationships between (1) the total number of component types, $\sum t[i]$, component domains, $\sum d[i]$, and components, $\sum c[i]$, present in the study and (2) the impacts of early literacy interventions, θ_i. Most findings in the studies included in the model will be linked to each of the multiple component types and component domains implemented in the study's intervention.

Estimates of the nested random effects in $\theta^{Type}_{t[i]}$ and $\theta^{Domain}_{d[i]}$ provide the associations between each component type or domain and intervention impacts. For example, these estimates provide information about how likely it is that instructional practices to build alphabetics skills are positively associated with the effects of interventions on student literacy outcomes. The interaction terms, $\theta^{Type\times Outc\times Indep}_{t[i]o[i]s[i]}$ and $\theta^{Domain\times Outc\times Indep}_{d[i]o[i]s[i]}$, capture the fact that relationships between components and impacts may differ based on the characteristics of the outcome measure. Estimates of these random effects provide a different set of associations for each group of outcome domain and for independent or non-independent measures. For example, these estimates provide information about how likely it is that instructional practices to build alphabetics skills are positively associated with the effects of interventions on independent alphabetics outcomes, or another type of outcome measure. However, in reality, because the majority of findings in the study sample are based on measures in the alphabetics domain, the findings reported in Appendix D.4 for other domains are not substantially different as the Bayesian approach uses information about the full study sample in estimating the relationships for each outcome domain. Linear relationships for the total number of component types, component domains, and components reflect the possibility that interventions that are more complex may tend to be more or less effective.

A final equation, C.4, governs the small-study effects (β):

$$\beta_i = \beta^{Int} + \beta^{Eric}_{e[i]} + \beta^{Res}_{r[i]} + \beta^{Purp}_{p[i]} \quad \text{(C.4)}$$

These small-study effects, which enter into the model as an interaction with the standard error of the impact estimate, s_i, are included to address statistical bias in impact estimates from file-drawer bias. In addition to an intercept term, β^{Int}, these effects can be influenced by whether a study is indexed in ERIC, $\beta^{Eric}_{e[i]}$, the study rating, $\beta^{Res}_{r[i]}$, and the reason the WWC reviewed the study, $\beta^{Purp}_{p[i]}$. For example, if the WWC is more likely to review studies when they are indexed in ERIC, then the file-drawer bias problem could tend to be smaller in those studies. As noted above, both the study rating and the reason the WWC reviewed the study also enter separately into the model as main effects in Equations C.1 and C.2 that are not related to the standard error of the estimate. However, only early literacy interventions that have been rigorously studied are represented in the analysis, but these interventions may not be representative of all early literacy interventions.

The model is fit using MCMC as implemented in the software Stan (Gelman et al., 2015), using four separate Markov chains, each with 1,000 burn-in iterations that are discarded and 1,000 iterations from which the posterior is calculated. The control parameters were left at their defaults, with the exception of Stan's *adapt_delta* parameter, which we increased to 0.95 to prevent divergent transitions from occurring. Autocorrelation, as measured by effective sample size, was excellent, with all but two of the 1,133 estimated parameters achieving an effective sample size (that is, the number of effectively independent posterior draws) over 100, a common benchmark. The two parameters were the estimated skew of the finding and intervention

random effects, with effective sample sizes of 32 and 90, respectively. Likewise, convergence, as measured by the Gelman-Rubin diagnostic (Gelman and Rubin 1992), $\hat{R}$, was excellent, with just one parameter exceeding 1.1, a common benchmark. This occurred for the estimated skew of the finding random effects, for which the $\hat{R}$ was 1.101.

The rest of this appendix is devoted to explaining everything about Equations C.2-C.4, including the notation and descriptions of each element. In particular, Section C.4 describes the prior distributions used in the Bayesian meta-analytic model.

C.3. Notation

The notation used to describe this model conforms to standard practice in the field (Gelman & Hill, 2007). However, the notation may be confusing for readers unfamiliar with Bayesian hierarchical modeling. In non-Bayesian models, it is common to represent subgroup variables using so-called "indicator variables" (also known as "dummy variables"). For example, one might relate an outcome to indicators of grade level using the equation:

$$y_i = \alpha + \beta_1 G1_i + \beta_2 G2_i + \beta_3 G3_i + \epsilon_i \quad \text{(C.5)}$$

In this equation, the variables *G1*, *G2*, and *G3* would be binary indicators of whether observation *i* belongs to the subgroup *grade 1*, *grade 2*, or *grade 3*, with *kindergarten* as the omitted category.

With Bayesian models, it is common to use a more compact notation, similar to vector notation sometimes used in other models. In this example, the same relationship between impact and grade levels would be described using the equation:

$$y_i = \alpha + \beta^{Gr}_{g[i]} + \varepsilon_i \quad \text{(C.6)}$$

In this equation, β^{Gr} is a vector of parameters that are restricted to sum to zero, and the subscript $g[i]$ denotes the grade level g that observation i belongs to, replacing the indicator variables for coding subgroup membership. Note that an "omitted category" is not needed in a hierarchical Bayesian model in which the vector of parameters is shrunken toward zero (Gelman & Hill, 2007).

C.4. Regression covariates and indexing

Every impact estimate recorded in the WWC database can have a variety of different characteristics. The list in Table C1 describes the indexing structure used to represent these characteristics that are included in the model described by Equations C.2-C.4 above.

Table C1. Subgroups and indicators included in the meta-analytic model

Subgroup indicators	Description
$g[i]$	The grades (kindergarten to grade 3) associated with finding i
$o[i]$	The outcome domain associated with finding i; outcome domains are: • Alphabetics • Literacy achievement • Proficiency in a non-English language • Proficiency in English • Reading comprehension • Reading fluency • Receptive communication • Vocabulary • Writing conventions • Writing quality For the interaction terms with the dichotomous independence indicator and intervention component domains or types, the three outcome domain groups are: • Alphabetics • Reading comprehension, reading fluency, receptive communication, or vocabulary • Literacy achievement, proficiency in a non-English language, proficiency in English, writing conventions, or writing quality
$i[i]$	The independence of the outcome associated with finding i. The three values are: • Outcome measure created by the developer of the intervention • Outcome measure created by the researchers • Independent outcome measure For the interaction terms with outcome domain groups and intervention component domains or types, independence was coded dichotomously as: • Outcome measure created by the developer of the intervention or by the researchers • Independent outcome measure
$p[i]$	The purpose of the WWC review associated with finding i. The seven purpose types are: • Department-funded evaluation • Distance learning rapid review • Grant competition • Institute of Education Sciences (IES) performance measure • Intervention report • Single study review • Practice guide
$t[i]$	Which of the five component types are associated with finding i
$d[i]$	Which of the 15 component domains are associated with finding i

Subgroup indicators	Description
$e[i]$	Whether finding i is reported within a publication that is indexed in the Education Resources Information Center (ERIC) database or not
$r[i]$	Whether finding i meets WWC standards without reservations (a high-quality randomized controlled trial) or meets WWC standards with reservations (typically a randomized trial with high attrition or a quasi-experimental design)

C.5. Parameters and priors

Each parameter in the model is informed by the WWC study data and a prior. The prior may be (a) informed by data from another source, such as data from broader groups of studies reviewed by the WWC that might include those in topic areas besides early literacy; (b) weakly informative, meaning that general assumptions about effect size distributions are made; or (c) hierarchical, meaning partly estimated by the relationships observed within the data. Weakly informative priors require the strongest assumptions, so these are avoided when one of the other two alternatives is feasible either because relevant information is available in other data sources, or the study data set includes relevant information. In the case of the overall intercept, the choice of prior is inconsequential.

An informative prior should be used only when the available information from another source is relevant to the parameter being estimated. Informative priors for some parameters in the meta-analysis are drawn from Deke et al. (2022), which estimated these parameters using WWC-reviewed studies across multiple topic areas in addition to early literacy. For other parameters, the meta-analysis draws informative priors from Zurovac et al. (2021), who estimated parameters related to non-experimental bias. The meta-analysis also uses an informative prior for estimating degrees of freedom based on the recommendation in Juarez and Steel (2010).

The meta-analysis uses several hierarchical priors. For example, a hierarchical prior is used to estimate whether studies tend to have larger impacts in some outcome domains. For outcome domains with fewer measures in the study data set, it is possible to observe a large average impact by chance that is larger than the range across the other outcome domains. In this case, the model will tend to adjust the estimate for that domain downwards in line with the other outcome domains.

Table C2 lists the parameters in the model along with the prior distribution and the type of prior for each. The parameters are organized in the table based on whether they contribute to (1) θ_i, the true impact estimated by finding i (Equation C.2); (2) θ^C, the overall effect of all types, domains, and components associated with each intervention (Equation C.3); (3) β_i, the multiplier on the standard error of finding i (s_i), to account for small-study effects, which is how the model addresses potential bias due to the file-drawer problem (Equation C.4); (4) measuring bias in impact estimates due to the research design; or (5) a parameter shared by multiple components of the model (called a hyperparameter).

Table C2. Prior distributions for all model parameters

Parameter	Description	Prior distribution	Type of prior
Components of θ_i, the true impact estimated by finding i			
θ^{Int}	The overall effect (that is, the intercept) common to all findings in this analysis	$N(0,1)$	Weakly informative
θ^{Gr}	The set of random effects of grades studied in finding i. If a finding is reported as pertaining to multiple grades, the average of all applicable random effects is used (this is achieved using dummy variables in the Stan code to handle this contingency). They are constrained to sum to zero: $\theta^{Gr}_{K} + \theta^{Gr}_{First} + \theta^{Gr}_{Second} + \theta^{Gr}_{Third} = 0$.	$N(0,\sigma^{Gr})$ with $\sum \theta^{Gr} = 0$	Hierarchical
σ^{Gr}	The standard deviation of the grade-level random effects	$N^{+}(0,\sigma_{\sigma})$	Hierarchical
θ^{Outc}	The set of random effects of the outcome domain of the finding. They are constrained to sum to zero: $\theta^{Outc}_{Alph} + \theta^{Outc}_{General} + \ldots + \theta^{Outc}_{WritQual} = 0$.	$N(0,\sigma^{Outc})$ with $\sum \theta^{Outc} = 0$	Hierarchical
σ^{Outc}	The standard deviation of the outcome domain random effects	$N^{+}(0,\sigma_{\sigma})$	Hierarchical
θ^{Indep}	A set of random effects for the independence of the outcome in finding i	$N(0,\sigma^{Indep})$	Hierarchical
σ^{Indep}	The standard deviation of the outcome independence random effects	$N^{+}(0,\sigma_{\sigma})$	Hierarchical
θ^{Purp}	A set of random effects for the purpose of the WWC review associated with finding i	$N(0,\sigma^{Purp})$	Hierarchical
σ^{Purp}	The standard deviation of the WWC purpose of review random effects	$N^{+}(0,\sigma_{\sigma})$	Hierarchical

Parameter	Description	Prior distribution	Type of prior
θ^{Find}	A set of random effects capturing the idiosyncratic effect specific to each finding	$SGT\left(0,\sigma^{Find},\lambda^{Find},\nu^{Find},2\right)$ The weakly informative priors for the skew and degrees of freedom parameters (below) pull the distribution toward symmetry and thin tails (that is, toward the normal distribution). The random effects are heteroskedastic, with standard deviation of σ_i modeled as follows, varying with the independence of the outcome measure: $\ln\left(\sigma_i^{Find}\right) = \tau^{Int} + \tau_{i[i]}^{Indep}$	Hierarchical
τ^{Int}	The overall variation across idiosyncratic findings effects	$N\left(\log\left(0.19\right),0.25\right)$	Informed by Deke et al. (2022)
τ^{Indep}	The variation in idiosyncratic findings effects within each of the three categories of independence of the outcome measure, constrained to sum to zero	$N\left(0,0.5\right)$	Weakly informative
λ^{Find}	The skew parameter for the idiosyncratic findings effects	$\mathrm{B}^{*}\left(1.5,1.5\right)$ The distribution is rescaled (by multiplying all values by 2 and then subtracting 1) to have support between -1 and 1; the rescaled $Beta$ distribution has a mean of 0 and a standard deviation of 0.5.	Weakly informative
ν^{Find}	The degrees of freedom parameter for the idiosyncratic findings effects	$\Gamma\left(2,0.1\right)$ This parameterization implies a mean of 20 with a standard deviation of about 14.	Informed by Juárez & Steel (2010)
θ^{Stu}	A set of random effects capturing the overall effect associated with each study	$SGT\left(0,\sigma^{Stu},\lambda^{Stu},\nu^{Stu},2\right)$ The random effects are homoskedastic, with a single standard deviation, σ^{Stu}	Hierarchical
σ^{Stu}	The standard deviation of the study random effects	$\Gamma\left(11.02,66.44\right)$	Informed by Deke et al. (2022)

Parameter	Description	Prior distribution	Type of prior
λ^{Stu}	The skew parameter for the study random effects	$\mathrm{B}^{*}(1.5,1.5)$	Weakly informative
ν^{Stu}	The degrees of freedom parameter for the study random effects	$\Gamma(2,0.1)$	Informed by Juárez & Steel (2010)
θ^{Inv}	A set of random effects capturing the overall effect associated with each intervention	$SGT(0,\sigma^{Inv},\lambda^{Inv},\nu^{Inv},2)$ The random effects are homoskedastic, with a single standard deviation, σ^{Inv}	Hierarchical
σ^{Inv}	The standard deviation of the intervention random effects	$\Gamma(11.02,66.44)$	Informed by Deke et al. (2022)
λ^{Inv}	The skew parameter for the intervention random effects	$\mathrm{B}^{*}(1.5,1.5)$	Weakly informative
ν^{Inv}	The degrees of freedom parameter for the intervention random effects	$\Gamma(2,0.1)$	Informed by Juárez & Steel (2010)
Components of θ^{C}, the overall effect of all types, domains, and components associated with each intervention			
θ^{Type}	A set of random effects capturing the effect of each component type, with each finding receiving the effect of all component types that are applicable to it	$N(0,\sigma^{Type})$	Hierarchical
σ^{Type}	The standard deviation of the component type random effects	$N^{+}(0,\sigma_{\sigma})$	Hierarchical
θ^{Domain}	A set of random effects capturing the effect of each component domain. As above, each finding receives the effect of all component domains that are applicable to it.	$N(0,\sigma^{Domain})$	Hierarchical
σ^{Domain}	The standard deviation of the component domain random effects	$N^{+}(0,\sigma_{\sigma})$	Hierarchical
$\theta^{Type\times Outc\times Indep}$	A set of random effects for combinations of component type, each of three groups of outcome domains, and dichotomous outcome independence. Each finding receives the interaction effects of all component types that are applicable to it.	$N(0,\sigma^{Type\times Outc\times Indep})$	Hierarchical

Parameter	Description	Prior distribution	Type of prior
$\sigma^{Type\times Outc\times Indep}$	The standard deviation of the component type-outcome domain-independence random interaction effects	$N^{+}\left(0,\sigma_{\sigma}\right)$	Hierarchical
$\theta^{Domain\times Outc\times Indep}$	A set of random effects for combinations of component domain, each of three groups of outcome domains, and dichotomous outcome independence. Each finding receives the interaction effects of all component domains that are applicable to it.	$N\left(0,\sigma^{Domain\times Outc\times Indep}\right)$	Hierarchical
$\sigma^{Domain\times Outc\times Indep}$	The standard deviation of the component domain-outcome domain-independence random interaction effects	$N^{+}\left(0,\sigma_{\sigma}\right)$	Hierarchical
θ^{NType}	Linear coefficient on the number of component types present in the study, standardized to have a mean of 0 and a standard deviation of 1	$N\left(0,0.1\right)$	Weakly informative
$\theta^{NDomain}$	Linear coefficient on the number of component domains present in the study, standardized to have a mean of 0 and a standard deviation of 1	$N\left(0,0.1\right)$	Weakly informative
θ^{NComp}	Linear coefficient on the number of components present in the study, standardized to have a mean of 0 and a standard deviation of 1	$N\left(0,0.1\right)$	Weakly informative
Components of β_i, the multiplier on s_i, the standard error of finding i, to account for small-study effects			
β^{Int}	The overall small-study coefficient (the intercept of its regression)	$N\left(0.5,0.5\right)$	Informed by Deke et al. (2022)
β^{Eric}	A set of additive effects to allow for different publication types having smaller or larger small-study effects. These are constrained to sum to zero: $\beta_{Yes}^{ERIC}+\beta_{No}^{ERIC}=0$.	$N\left(0,0.5\right)$	Weakly informative
β^{Res}	A set of additive effects to allow findings which meet WWC standards with or without reservations to have different small-study effects. These are constrained to sum to zero: $\beta_{Without}^{Res}+\beta_{With}^{Res}=0$.	$N\left(0,0.5\right)$	Weakly informative

Parameter	Description	Prior distribution	Type of prior
β^{Purp}	The set of random effects of each of the seven purpose of review types	$N\left(0,\sigma_{\beta_{Purp}}\right)$	Hierarchical
$\sigma_{\beta_{Purp}}$	The standard deviation of the purpose of review random effects	$N(0,0.5)$	Weakly informative
Research design bias			
ω	A set of random effects capturing the bias associated with each study that meets WWC standards with reservations	$N(0,0.12)$	Informed by Zurovac et al. (2021)
Additional common hyperparameter			
σ_{σ}	A standard deviation parameter common to several prior distributions	$\Gamma\left(3.96,12\right)$	Informed by Deke et al. (2022)

Notes:

$N(\mu,\sigma)$ indicates a normal distribution with mean μ and standard deviation σ.

$N^{+}(\mu,\sigma)$ indicates a normal distribution truncated below at zero (that is, positive-only) with location μ and scale σ.

$SGT(\mu,\sigma,\lambda,v,p)$ indicates a skewed generalized t distribution with location μ, scale σ, skewness λ, degrees of freedom v, and kurtosis parameter p.

$B^{*}(\alpha,\beta)$ indicates a beta distribution with shape parameters α and β, rescaled (by multiplying all values by 2 and then subtracting 1) to have support between -1 and 1. The $B^{*}(1.5,1.5)$ distribution has a mean of 0 and a standard deviation of 0.5.

$\Gamma(k,\beta)$ indicates a gamma distribution with shape parameter k and rate parameter β.

APPENDIX D. ADDITIONAL RESULTS

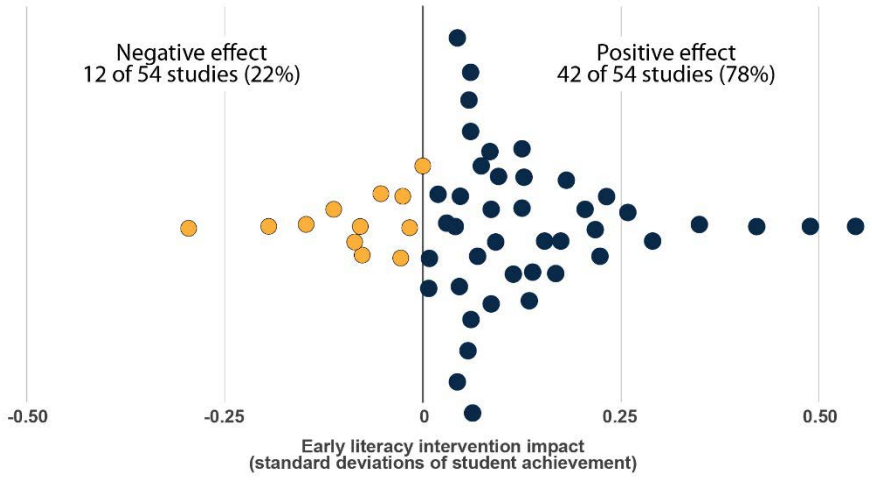
The report presented findings based on the 29 studies that reported impacts on independent alphabetics findings. However, the relationships estimated in the meta-analysis for independent alphabetics outcomes are also informed by relationships for the other types of literacy outcomes, as described in Appendix C. This appendix contains additional findings from the Bayesian meta-regression analysis, based on all 54 studies that reported impacts on any literacy outcome, including the following:

- The distribution of impacts in early literacy studies overall and by outcome type
- The implementation of component domains in interventions that had positive or negative effects based on all outcomes (not just independent alphabetics outcomes)
- Variation in unexplained impacts of early literacy interventions
- The extent to which the associations between component domains and intervention impacts change when considering different outcome domains
- Associations between the individual components and intervention impacts

D.1. The distribution of impacts in early literacy studies overall and by outcome type

The meta-analysis found that most early literacy interventions have positive impacts across the full set of student literacy outcomes, after accounting for several sources of error in impact studies. For example, across all early literacy outcome measures, interventions examined in 78 percent of studies in the meta-analysis report positive impacts of early literacy interventions, as shown in Figure D1. These impacts that are aggregated across all outcome measures, are slightly more favorable than those for independent alphabetics outcomes only, for which 72 percent of impacts were positive, shown in Figure 5 in the report.

Figure D1. Effects of early literacy interventions on all literacy outcomes



Notes: Results are based on 241 findings from 54 studies of early literacy interventions included in the meta-analysis. Effect sizes are based on adjustments applied in the meta-analysis to address statistical error in impact estimates, file-drawer bias, and the quality of the research design , as described in Appendix C.2. The range of impact estimates in the table reflects variation in impacts across grade levels, outcome and outcome domains, and interventions. Each dot in the figure represents a study.

Source: Authors' calculations based on data from the What Works Clearinghouse's database of reviewed studies.

However, readers should use caution when comparing the distributions of impacts in Figures 5 and D1. Not all studies included findings for all types of outcome measures. In particular, 29 of the 54 studies included independent alphabetics findings. When looking at the findings for only one type of outcome measure, the interventions or other contextual factors in the studies that include those findings may not be representative of the full sample. In that case, differences in the distributions for two groups of measures might be due to differences in the impact estimates for those outcome measure types or to differences in the interventions and contextual factors in those studies.

To focus on differences due to the outcome measures, Table D1 reports distributions for different groups of outcome measures that hold constant the studies and interventions. To do this, the distribution in each row of the table is calculated by applying an adjustment to the overall distribution of impacts to account for the mean and standard deviation of the impacts on outcome measures in the group that are estimated in the meta-analytic model. Because this approach retains all findings in each row, the interventions and studies included in each row are the same. Using this approach, 66 percent of the findings (as opposed to studies) are positive, as shown in first row of Table D1. It remains true that most early literacy interventions have favorable impacts on student literacy outcomes when looking at most types of outcome measures, including outcomes that were created independently of developers and researchers, and outcomes in different domains.

Table D1. Distribution of impacts by outcome domain and measure type

Outcome measure type	Percentage of impacts that are positive	Mean	Standard deviation	10th percentile	25th percentile	50th percentile	75th percentile	90th percentile
All measures	66%	0.07	0.20	-0.13	-0.05	0.05	0.17	0.28
Independent	61%	0.03	0.17	-0.18	-0.06	0.03	0.13	0.24
Researcher	87%	0.19	0.20	-0.04	0.07	0.16	0.30	0.44
Developer	91%	0.24	0.20	0.02	0.13	0.22	0.35	0.49
Independent alphabetics	70%	0.06	0.16	-0.14	-0.03	0.06	0.15	0.24
Independent reading comprehension	61%	0.02	0.16	-0.17	-0.05	0.03	0.11	0.21
Independent reading fluency	66%	0.03	0.16	-0.16	-0.04	0.04	0.12	0.22
Independent vocabulary	35%	-0.05	0.16	-0.24	-0.13	-0.05	0.04	0.14
Independent literacy achievement	54%	0.01	0.16	-0.21	-0.09	0.01	0.10	0.20

Notes: Results are based on 241 findings from 54 studies of early literacy interventions. The model accounted for statistical errors that influence impact estimates, including estimation error, bias associated with studies conducted using a quasi-experimental design, and file-drawer bias, which would otherwise lead to observing a broader range of impact estimates.

The range of impact estimates in the table reflects variation in impacts across grade levels, outcome and outcome domains, and interventions. The distribution in each row is calculated by applying an adjustment to the overall distribution of impacts to account for the mean and standard deviation of the impacts on outcome measures in the group that are estimated in the meta-analytic model. Because the number of studies is small, the table reports distributions at the findings level instead of aggregating to the study level. This means that the percentage of findings that are positive reported in this table differ from the percentage of studies with positive aggregated impacts shown in Figures 5 and D1.

The five rows for individual outcome domains exclude five other outcome domains with fewer than 15 findings, as shown in Figure B1. The bottom two rows of the table together with the independent alphabetics row represent mutually exclusive groups of findings and represent all 200 independent findings in the studies included in the meta-analysis.

Source: Authors' calculations based on data from the What Works Clearinghouse's database of reviewed studies.

Distribution of impacts of interventions on independent and non-independent measures. The intervention impacts on independent outcomes were smaller than the intervention impacts on non-independent outcomes. For example, as shown in Table D1, just 61 percent of impacts on independent outcomes are positive, compared to 87 percent for measures created by study authors and 91 percent for measures created by intervention developers. Also, for independent measures, 90 percent of impacts are *less* than 0.24 standard deviations, and 90 percent are *greater* than -0.18 standard deviations, representing a 90-10 percentile range of 0.42 standard deviations. This range is 0.48 standard deviations for measures created by study authors, meaning that these measures are less tightly distributed—that is, impacts on measures created by study authors are spread over a wider range. The range is 0.53 for measures created by intervention developers. These differences between impacts on independent and non-independent measures are consistent with findings in prior research (Wolf & Harbatkin, 2022).

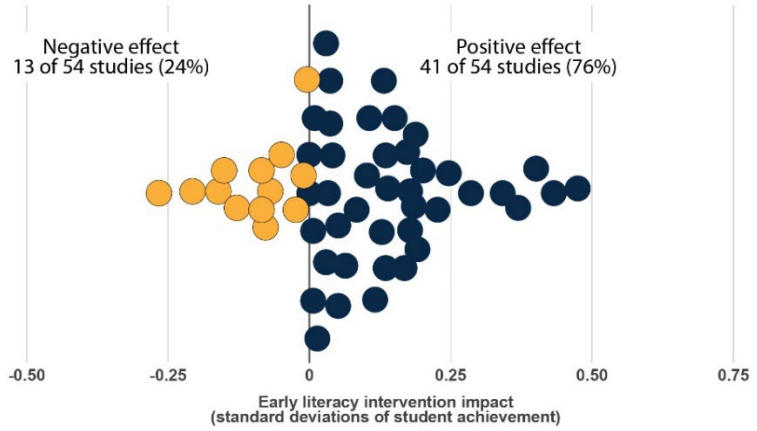
Distribution of impacts of interventions on outcome measures in different domains. Impacts of the interventions tend to be more favorable for independent alphabetics outcomes compared to independent outcomes in the other literacy domains. For example, 70 percent of intervention impacts on independent alphabetics outcomes are positive, while 61 percent are positive for independent reading comprehension outcomes, 66 percent are positive for independent reading fluency outcomes, and just 35 percent are positive for independent vocabulary outcomes. One possible explanation for the less favorable impacts of interventions on independent vocabulary outcomes is that the interventions were more likely to emphasize other skills. For example, fewer of the 54 studies included components related to building background knowledge (40 studies) compared to those related to building alphabetics skills (48 studies). However, there were only 17 independent vocabulary outcomes in the meta-analysis, so the mean and standard deviation of the impacts used to measure the distribution are likely imprecisely measured. Other possible explanations for the larger impacts of early literacy interventions on alphabetics outcomes compared to impacts on other outcomes include the relative simplicity of the measures (Nation & Snowling, 1997) and the relative simplicity of some alphabetics instruction (Paris, 1995). Table D1 excludes rows for five other outcome domains with fewer than 15 findings, as shown in Figure B1.

D.2. Variation in unexplained impacts of early literacy interventions

The application of the new methodology found that only 9 percent of the variation in impacts on independent alphabetics outcomes is explained by the intervention component types and domains. Alternatively, the component types and domains explain 10 percent of the variation in impacts on all literacy outcomes. The variation in impacts of interventions on all literacy outcomes remaining after removing variation that is explained by the meta-analytic model, including variation due to component types and domains of interventions, outcome measure type, outcome domains, and grade level, is shown in Figure D2. Therefore, Figure D2 shows the share of the variation in intervention impacts that is due to other factors not examined in this synthesis, which may include intervention components that this synthesis did not measure, contextual factors, and implementation quality.

In removing variation due to outcome measures, it is necessary to select a scale for the impact estimates. Therefore, the impacts in Figure D2 have been transformed to reflect the most common type of outcome measures in the analysis—independent alphabetics outcomes. This is done using an approach similar to the one used to report the distribution of impacts for independent alphabetics outcomes in Table D1.

Figure D2. Magnitudes of unexplained intervention impacts on all literacy outcomes in early literacy studies



Notes: Results are based on 241 findings from 54 studies of early literacy interventions included in the meta-analysis. The model accounted for statistical errors and other factors that influence impact estimates, and for grade, outcome domain and independence, and WWC purpose of review. Each dot in the figure represents a finding. The 100 dots in the figure are a stratified random subsample of all 241 findings. The impacts have been transformed to reflect the impact of interventions on independent alphabetics outcomes.

Source: Authors' calculations based on data from the What Works Clearinghouse's database of reviewed studies.

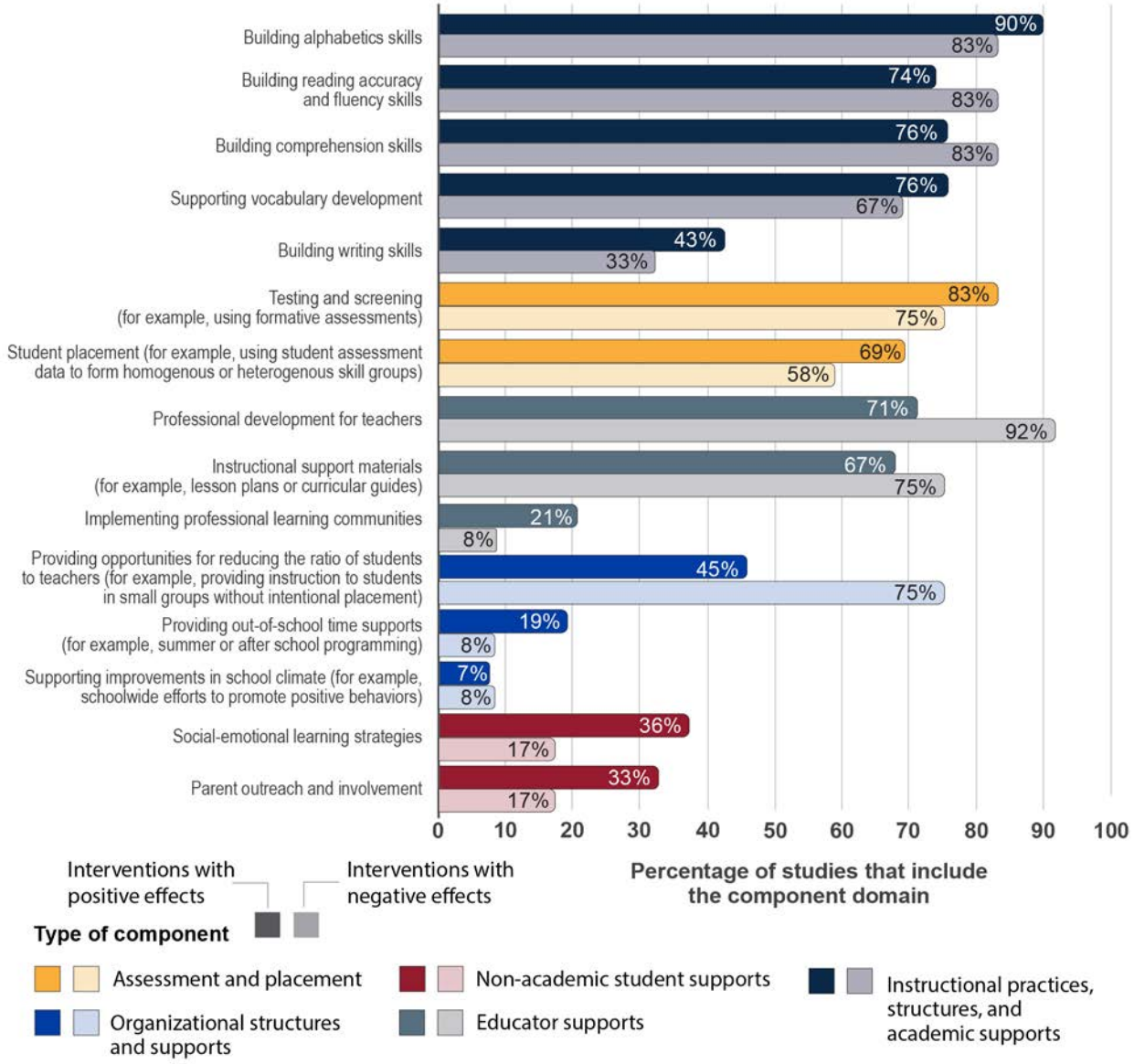
D.3. The implementation of component domains in interventions that had positive or negative effects based on all outcome measures

The application of the new methodology found that some component domains are more likely to be implemented in interventions that had positive effects on independent alphabetics outcomes, as shown in Figure 6 in the report. These descriptive results are generally similar when looking at interventions that has positive effects on any literacy outcomes, not just independent alphabetics outcomes. Figure D3 reports how often each component domain was implemented in each of two groups of studies based on the effect sizes reported in Figure D1. The top, darker bars in each pair report on implementation in the 42 studies that examined interventions that had positive effects on literacy outcomes, and the bottom, lighter bars report on implementation in the 12 studies that examined interventions that had negative effects on those outcomes.

Similar to the results in Figure 6 for independent alphabetics outcomes, the testing and screening component domain and student placement component domain are present in more interventions that had positive effects compared to interventions that had negative effects when applying the methodology to include findings based all literacy outcomes. Also similar to the findings in the report, component domains for professional development, instructional support materials, and providing opportunities for reducing the ratio of students to teachers were implemented more often in interventions that had negative effects compared to interventions that had positive effects.

Just as with the results in Figure 6, these descriptive differences in implementation between the two groups of interventions are suggestive of which component domains might be positively associated with intervention effects. However, multiple component domains were implemented together in interventions, so Figure D3 does not answer which component might be, by itself, positively associated with intervention effects.

Figure D3. Component domains included in early literacy interventions that had positive or negative effects on any literacy outcomes



Notes: The figure includes 54 studies of early literacy interventions included in the meta-analysis. Interventions with positive effects are those in the 42 studies in Figure D1 with positive impacts on any literacy outcomes. Interventions with negative effects are those in the 12 studies in Figure D1 with negative impacts on any literacy outcomes.

Source: Authors' coding of components in early literacy studies.

D.4. The extent to which the associations between component domains and intervention impacts change when considering different outcome domains

The synthesis investigated the possibility that the ranking of component domains might change when considering only outcomes in certain domains. To do this, the meta-analysis included an interaction between the component domain and three groups of outcome domains, as well as whether the outcome is independent or non-independent, as described in Appendix C. This interacted analysis allows the ranking of component domains to change based on the outcome measure type. The three groups of outcome domains are (1) only the alphabetics domain; (2) reading comprehension, reading fluency, receptive communication, and vocabulary; and (3) literacy achievement, the proficiency domains, and the writing domains. The 10 outcome domains are grouped for this analysis because most outcome domains included fewer than 20 findings, as shown in Figure B1. The study team selected these groups so that each would include at least 30 outcome measures and so that the measures in the third group of outcome domains would tend to be broader measures of academic achievement than the measures in the other two groups. There are other ways to group the outcome domains that could potentially be more meaningful for some purposes. However, the meta-analysis team did not explore other groupings because the results presented below indicate that it is not possible, based on the study data, to make precise statements about how much the ranking of component domains change based on the outcome measure type.

Considering different groups of outcome domains besides alphabetics using the interacted model did not change which four component domains are the most likely to have positive relationships with intervention impacts. Considering all the literacy outcomes together also did not change these four component domains. These four component domains, in the top four rows of Table D2, consistently have the largest positive associations with intervention effects. Some differences in the probabilities are evident across the columns of Table D2. For example, the component domains in the non-academic student supports type have more positive associations with the outcomes in the last column—including outcomes in domains for literacy achievement, proficiency in English or a non-English language, and writing conventions or quality, which tend to be broader measures of academic achievement—compared to independent measures of alphabetics or fluency and comprehension.

However, these results and patterns may be driven more by the limitations of the study data than empirically similar patterns across the outcome domains. In a Bayesian meta-analysis, results for any particular subgroup of studies will be "shrunken" towards the results for the full sample. This means that similar results across subgroups, such as those in Table D2, may be the result of this shrinkage rather than genuinely similar relationships. There are two specific reasons that the study data set might limit what can be learned from this analysis. First, the studies in the meta-analysis are not optimally designed for this purpose. Ideally, to provide the most information about the relative rankings of the component domains across different outcome domains, the studies would examine interventions with only one component domain—or at least relatively few in varied combinations—and report impact findings on measures in each outcome domain. The actual dataset is far from optimal in this way. In particular, alphabetics outcome measures are used only in the studies that include alphabetics component domains. As a consequence, the analysis cannot confidently distinguish between two explanations for larger impacts in one outcome domain relative to another: (1) that impacts are generally larger in the outcome domain no matter what the intervention entails or (2) that impacts are larger in the outcome domain because measures in the domain are well-aligned with the skills targeted by the interventions. Given the data limitations, the interacted analysis relies heavily on the Bayesian model's hierarchical priors, which results

in giving the first explanation roughly 2.5 times more weight than the second explanation, based on findings data from other outcome domains where these explanations can be better distinguished empirically.

A second reason the findings in the study data set limit what can be learned from this analysis arises because there are relatively few findings and studies in each outcome domain, with the exception of the alphabetics domain. The synthesis team estimates that at least 30 to 50 additional studies would be needed to begin to be able to make confident statements about similarities and differences in the ranking of component domains for different outcome domains. This rough estimate is based on comparing how the rankings of two specific component domains for two sets of outcome domains each change when adding additional simulated studies to the dataset. The two component domains selected by the team are the highest- and lowest-ranked component domains in Figure 7 based on their relationships with independent alphabetics outcomes, which are the testing and screening component domain and the reducing ratios of students to teachers. The two groups of outcome domains selected by the team are independent measures in the alphabetics domain and the group of outcome domains that includes reading comprehension, reading fluency, vocabulary, and receptive communication. To approximate the number of additional studies that might be needed to make precise claims about the relative rankings of the two component domains across the two sets of outcome domains, the team added copies of relevant studies in the meta-analysis, but for which the meta-analysis team replaced the findings for measures in the second group of outcome domains with findings that are more favorable for interventions with the reducing ratios of students to teachers component domain. By doing so, a synthesis across the existing studies and the additional studies with altered findings will tend to produce a higher ranking for the reducing ratios of students to teachers component domain.

Table D2. Probabilities that the associations between component domains and intervention impacts are positive when considering different outcome domains

		Probability a component domain is positively associated with early literacy intervention impacts:				
Component domain	Component type	Independent alphabetics (Figure 7)	All outcomes	All independent outcomes	Independent reading comprehension, reading fluency, receptive communica-tion, and vocabulary	Independent literacy achievement, proficiency in English or a non-English language, and writing conventions or quality
Testing and screening	Assessment and placement	87%	88%	90%	84%	78%
Student placement		82%	88%	85%	85%	78%
Social-emotional learning strategies	Non-academic student supports	74%	81%	78%	69%	84%
Parent outreach and involvement		61%	75%	66%	66%	80%

Component domain	Component type	Probability a component domain is positively associated with early literacy intervention impacts: Independent alphabetics (Figure 7)	All outcomes	All independent outcomes	Independent reading comprehension, reading fluency, receptive communication, and vocabulary	Independent literacy achievement, proficiency in English or a non-English language, and writing conventions or quality
Building reading accuracy and fluency skills	Instructional practices, structures, and academic supports	49%	37%	39%	40%	41%
Building alphabetics skills		45%	36%	36%	40%	38%
Supporting vocabulary development		43%	35%	32%	37%	37%
Building writing skills		42%	34%	36%	36%	35%
Building comprehension skills		33%	28%	29%	35%	30%
Implementing professional learning communities	Educator supports	43%	39%	53%	49%	42%
Instructional support materials		29%	27%	30%	35%	31%
Professional development for teachers		21%	22%	16%	26%	23%
Providing out-of-school time supports	Organizational structures and supports	26%	25%	36%	26%	28%
Supporting improvements in school climate		28%	23%	38%	28%	28%
Providing opportunities for reducing the ratio of students to teachers		16%	15%	17%	21%	19%

Notes: Results are based on 241 findings from 54 studies of early literacy interventions.

Source: Authors' calculations based on data from the What Works Clearinghouse's database of reviewed studies.

Specifically, the team added altered copies of the 37 studies in the meta-analysis that examined (1) interventions with either or both of the focal component domains and (2) impacts of those interventions on either or both of the groups of outcome domains. The altered findings for the second group of measures were generated by re-centering them within groups based on whether the intervention implemented the focal component domains to appear more similar to the alphabetics findings, but in reverse. For example, findings for the second group of measures in studies that implemented the testing and screening component domain were re-centered to have the mean among findings for the first group of measures in studies that implemented the reducing student-teacher ratios component domain. This ensured that the altered findings had overall differences exactly the opposite of what was observed for alphabetics, while preserving the natural variation in impact estimates. In doing this, adding one altered copy of each of the 37 studies led to the reducing ratios of students to teachers

component domain being the 6th-highest ranked outcome domain, rather than the 15th based on its relationship to the altered impacts on independent outcome in the group of outcome domains that includes reading comprehension, reading fluency, vocabulary, and receptive communication. This suggests that if there is a true difference between the strength of the relationship for a component domain between two groups of outcome measures, it would be possible to measure this difference with a dataset that included, roughly, about 80 to 100 studies (roughly, about 30 to 50 additional studies) so long as the studies examined interventions with the focal component domains and impacts on the focal outcome measures.

D.5. Associations between the individual components and intervention impacts

This synthesis cannot determine reliably whether some of the 81 individual components within the 15 component domains had stronger associations with intervention effectiveness than others. The meta-analysis is estimated at the level of component domain because the study sample size is too small to reliably differentiate such a large number of components. Although it is not possible to precisely measure the relationships between any individual component and intervention impacts, the Bayesian meta-analysis can estimate how large these relationships might be in general. To maximize the study sample size, the synthesis team examined the associations between individual components and intervention impacts for all literacy outcomes and did not restrict to independent alphabetics outcomes only.

The relationships of the 81 individual components with impacts of early literacy interventions are most likely very small. The estimated relationships for each component range from -0.007 to 0.008 standard deviations. These effect sizes are small in part because they are imprecisely estimated, and the meta-analysis shrinks the effect sizes to account for the imprecision. However, the meta-analytic findings suggest that the estimated effects of individual components would generally remain small even if estimated with additional data and greater precision. The meta-analysis results provide an estimate of the standard deviation of the component relationships, which measures how dispersed the relationships would be if they could be measured precisely. The estimated standard deviation of the component relationships is 0.018 standard deviations of student achievement, with a 90-percent credible interval of 0.002 to 0.042. A standard deviation within the range of the 90-percent credible interval is consistent with component relationships that are generally smaller than 0.05 standard deviations. An effect size of 0.05 standard deviations, which the results suggest would be among the largest component relationships, would be about half of the 0.09 effect size estimated for including the testing and screening component domain in Figure 7. Because the components are nested in the component domains, their relationships with intervention impacts will generally be smaller than the relationships for component domains reported in Figure 7.

APPENDIX E. COMPONENT CODING PROTOCOL

The synthesis team used and built upon a taxonomy previously developed for another IES project (Technical Assistance Supporting Evidence-Building and Use under contract number 91990020F0052; Scher & Martinez, 2022) with input from early literacy experts and intervention developers. Additional information about the IES's efforts to identify intervention components is available on the Components page of the IES website for its Standards for Excellence in Education Research (SEER). The synthesis team used the coding protocol to code the components of interventions and asked study authors to use the protocol and instructions below to code the components of the interventions examined in their studies.

Coding instructions and guide for describing components of early literacy interventions

The WWC is interested in the components of [Name of intervention] as implemented in the study described in the following citation:

[List of citations]

Instructions: Please use this coding guide to document the components of the intervention in your study indicated above.

- If your study has more than one intervention of interest to the WWC, you should complete one coding guide for each intervention.
- As you make your way through the coding guide, please refer to the instructions in each section as well as to the attached *GLOSSARY_ WWC_Early_Literacy_Components.docx* for specific definitions of intervention components.
- Please also see two attached examples of completed coding guides for your reference, one for a hypothetical curriculum and one for a hypothetical narrow practice.
- Complete both sections in this coding guide: (1) *Intervention Components by Type* (pages 2-20) and (2) *Intervention Delivery* (pages 21-23).

Background: This coding guide was developed to assess and refine the nomenclature related to early literacy. By nomenclature, we mean a common language for researchers, program developers, program providers, and practitioners to describe intervention components. A common language can lead to a better understanding of the similar characteristics of interventions and facilitate consistent comparisons across different interventions. The nomenclature is intended to be used to describe the components of an intervention. Documenting the components of early literacy interventions in this way could support efforts to investigate a broad range of research questions including, but not limited to, examining which of these components (or bundles of components) are associated with improved outcomes in early literacy.

Intervention components by type

Instructions: In this section, please code specific components of the intervention, by indicating which components are applicable and adding corresponding notes (when needed).

The components of interventions are organized in a hierarchical taxonomy using four categories of increasing granularity, although not all components will have an associated component approach:

Component categories		
Component type	→	Identifies the nature of a particular intervention
↓		
Component domain	→	A mechanism to organize similar sets of intervention components under a unifying theme
↓		
Component	→	Particular intervention characteristic
↓		
Component approach	→	Specific mechanisms or strategies that are employed to implement a particular component

Component approaches are listed in the following component coding tables as examples only. **The WWC is not currently coding component approaches for early literacy interventions.** But having examples of the approaches may help you to better understand and more accurately code the components in interventions.

The following tables are organized around five specific component types:

- Instructional practices, structures, and academic supports
- Non-academic student supports
- Organizational structures and supports
- Educator supports
- Assessment and placement

When completing this coding guide, please consider that the WWC is primarily interested in learning about the components of an intervention **as it was implemented in the environment in which it was studied**. In some cases, an intervention may have been subsequently updated or revised after a study was conducted. Alternatively, only some, as opposed to all, components of an intervention may have been implemented in a study. The WWC will want to understand what occurred in the study, rather than the intentions of those who have designed or implemented the intervention or how the intervention might have been implemented elsewhere. **Please describe the components of the intervention as it was implemented in the specific study cited above on page 1.** When you are uncertain about the implementation in the study, please make your best judgment about what components were likely included.

Please use the "**NOTES"** column in each of the five component type tables to provide:

1. Source information, when it is available, including citations and page numbers. It is acceptable to point to any source that describes the implementation of the intervention in the study, whether peer reviewed or not. We also recognize that the requested information will not always be documented publicly, and appreciate you sharing your knowledge and experience implementing the intervention.
2. Whether a component was modified in any way, such as any components of a branded intervention that were altered; for example, you might note that the intervention was intended to include 4 professional development sessions but only 3 were offered.
3. Differing terminology that the intervention uses that is aligned with a particular component domain, or component
4. Additional related components that fall within the domain
5. Any uncertainty in whether a component was included in the study and **who we might contact to learn more**

COMPONENT TYPE 1: INSTRUCTIONAL PRACTICES, STRUCTURES, AND SUPPORTS

Instructional practices, structures, and supports encompass academic-focused programmatic characteristics (such as course materials or curricular focus), teaching strategies aimed at improving academic skills (for example, using a pedagogical approach aimed at supporting the development of a particular reading skill), changes in the structures of academic delivery (for example, providing practice opportunities for small groups or pairs of students), or academic supports (for example, tutoring). In cases where both academic and non-academic elements cannot be disentangled, an intervention component should be considered part of this Instructional practices, structures, and supports category.

Please mark all component domains and components that apply.

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Instructional practices to build alphabetics skills	☐ Working with rhyming words		
	☐ Engaging in explicit/direct instruction		
	☐ Teaching students to recognize and manipulate segments of sound in speech	• Introducing familiar larger segments of speech (words) and drawing attention to smaller sounds segments • Demonstrating how sentences can be broken into words and then words into smaller words • Demonstrating and practicing how words can be broken into syllables and even small units (onsets or rimes) • Teaching to isolate and manipulate or substitute phonemes • Using Elkonin sound boxes • Using sorting pictures	

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Teaching letter names and letter-sound relations	• Presenting consonants and short vowel sounds represented by single letters • Introducing (through sounding not memorizing) consonant blends (e.g., fl, sm, st) and two-letter consonant diagraphs (e.g., sh, th, ch) • Teaching long vowels with silent e • Introducing two-letter vowels (e.g., ea and ou) • Introducing letters in upper and lowercase • Using pictures and stories that incorporate the sound of the letter • Reviewing previously learned letter sounds and using them in familiar contexts	
	☐ Link letter-sound relationship knowledge to phonemic awareness	• Using letter tiles or magnetic letters to build or change words based on sounds • Gradually adding more advanced words to support understanding of phonemic patterns (e.g., changing “can” to “cane” or “fat” to “flat”)	
	☐ Handwriting programs to call attention to letter shape		

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Teaching blending and sound-spelling patterns in words	• Starting with simple, familiar consonant-vowel-consonant words and demonstrate how to blend, then allow students to apply independently • Instructing on chunking sounds (read from left to right but adding sound from previous sound) • Instructing how to sound out words (saying the sound of each letter or letter combination one by one and the combining) • Manipulating letters by, for example, using a pocket chart with letter tiles, magnetic letters or Elkonin sound box, or moving tiles together as they are read • Encouraging students to minimize the schwa sounds • Encouraging students to check whether a word they produced sounds familiar when they try to pronounce it	
	☐ Introducing common sound-spelling patterns	• Sorting word cards into groups • Practicing writing words with similar patterns • Using Elkonin sound boxes to build words with specific sound-spelling patterns	

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Teaching to recognize common word parts	• Developing understanding of morphology and the knowledge of meaningful word parts • Teaching about suffixes, contractions, prefixes and basic roots of words • Practicing writing words or manipulating word parts • Practicing building words by adding or removing prefixes and suffixes • Teaching a word analysis strategy • Teaching how to adjust vowel sounds to achieve a recognizable word. • Detecting parts of words that are known	

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Teaching to read regular and irregular high-frequency words	• Teaching irregular words holistically (whole words rather than combinations of sounds) • Practicing reading words frequently until they are easily recognizable • Using flashcards to teach new words • Selecting high frequency words from text • Integrating high frequency words into text • Creating a word wall of high frequency words • Practicing high frequency words outside of regular literacy instruction • Small group practice of high frequency words on flashcards • Presenting lists of high frequency words and practicing writing them • Instructing on how to predict or guess a word based on context cues, background knowledge or memory	
	☐ Introducing phonetically irregular words	• Identifying phonetically irregular (or previously called non-decodable) words in advance in text • Limiting the number of phonetically irregular words introduced at any given time	
	☐ Reading decodable words in isolation and in text	• Using connected text passages to practice letter sounds • Using wordlists to practice letter sounds	
	☐ Other instructional practices to build alphabetics skills (specify in NOTES)		

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Instructional practices to build reading accuracy and fluency skills	☐ Encouraging daily reading of connected text	• Sustained silent reading (e.g., 20 minutes per day) in classroom • Requiring daily reading at home • Increasing the amount of reading over time • Encouraging voluntary reading	
	☐ Supporting oral reading through modeling, scaffolding and feedback	• Providing constructive feedback and support • Modeling use of effective word-reading strategies • Using instructional-level text with examples of recently taught sound-spelling patterns • Providing tailored prompts for strategies ("look for parts you know", "sound it out"). • Reducing prompts and supports as reading becomes more independent	
	☐ Teaching self-monitoring and self-correcting reading skills	• Modeling self-monitoring and self-correction (e.g., "Fix It" game) • Pausing when there are errors, provide support and have students reread the sentence with corrective feedback and support. • Reading a sentence as the student did to help identify an error. • Using scaffolds less frequently to promote independent self-monitoring	

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Providing opportunities for oral reading practice with feedback	• Modeling expression and phrasing • Introducing students to punctuation • Individual oral reading with instructor support • Individualized reading with a computerized reading device with the pace appropriate for students' reading rate • Peer or partner reading/shared reading • Choral reading in small groups with monitoring • Echo reading where more experienced reader begins • Alternating reading where students take turns • Simultaneous reading • Individual oral reading with a recording device, with teacher feedback provided later. • Repeated oral reading of the same texts • Practicing reading at a natural pace, with pauses, and with expression	
	☐ Other instructional practices to build reading accuracy and fluency skills (specify in NOTES)		

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Instructional practices to build comprehension skills	☐ Teaching how to use a specific reading comprehension strategy	• Instructing how to activate prior knowledge or predict • Instructing how to draw inferences • Instructing how to use monitoring, clarifying, or fix-up strategies • Instructing how to use questioning strategies • Instructing how to use summarizing/retelling strategies • Instructing how to use visualization strategies	
	☐ Using multiple-strategy formats	• Incorporating concept-oriented reading instruction • Incorporating informed strategies for learning • Incorporating reciprocal teaching methods • Incorporating transactional strategy instruction	
	☐ Explicit instruction with gradual release of responsibility		
	☐ Modeling and explaining how to identify and use a text's organizational structure	• Adapting complexity based on age/capacity of students • Developing simple mnemonics/charts/graphics to help identify elements of structure • Using familiar stories to model and explain how to identify and connect parts of narrative texts • Providing instruction on common structures of informational texts	

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Guiding students through focused discussion on the meaning of text	• Discussing concepts of locate and recall • Discussing how to integrate and interpret text • Discussing how to critique and evaluate text • Developing and presenting thought-provoking discussion questions • Asking follow-up questions and facilitating discussions • Having students lead structured, small-group discussions	
	☐ Purposefully selecting texts for comprehension development and knowledge-building	• Incorporating multiple genres (e.g., literary, informational) • Choosing texts of high quality with richness and depth of ideas and information • Choosing text aligned students' reading ability and the instructional activity • Using texts that support the purpose of a particular lesson (e.g., a lesson on text structure) • Selecting culturally responsive texts that are aligned with the students' diverse backgrounds	
	☐ Helping students discover the purpose and benefits of reading	• Modeling and discussing how the ability to read affects daily life • Displaying students' work or reading-related posters and signs • Creating classroom library or literacy centers • Incorporating hands-on activities that exemplify specific themes in readings • Choosing texts that are relevant to students	

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Creating opportunities for students to see themselves as successful readers		
	☐ Giving students reading choices	• Allowing students to choose from a variety of activities or centers • Permitting students to choose the order in which they complete work • Encouraging students to think of questions that lead them to texts they may be interested in • Allowing students to choose how to respond to a text • Giving students a choice in where they can read • Allowing students to choose from a selection of instructional texts.	
	☐ Providing peer collaboration/cooperative learning opportunities in reading comprehension activities	• Students summarizing paragraphs together • Students engaged in active listening • Students taking turns in making predictions	
	☐ Metacognitive awareness for comprehension monitoring	• Building skills to know when a student understands what they are reading • Developing skills to troubleshoot when they have trouble comprehending text (restating, looking back through text)	
	☐ Other instructional practices to build comprehension skills (specify in NOTES)		

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Instructional practices to support vocabulary and relevant background knowledge development	☐ Providing instruction to build relevant vocabulary and background knowledge	• Providing explicit vocabulary instruction through provision of definitions • Providing indirect vocabulary instruction through exposure to a range of words during reading • Using multimedia methods using media such as graphic representations • Using association methods to support finding connections to words students already know • Repeated exposure to vocabulary words, including in different contexts • Pre-instruction of vocabulary words prior to reading • Restructuring text materials by substituting easy for hard words	
	☐ Engaging in conversations to support inferential knowledge	• Using open-ended questions to challenge students to make inferences in narrative and informational text • Asking increasingly complex questions to support critical thinking around author motivation and choices • Modeling how to provide reasoned answers • Designating a small group conversation leader	

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Engaging students in developing narrative language skills	• Focusing on the elements of narrative language • Identifying and teaching complex grammatical structures and syntax • Modeling and providing practice opportunities for developing narrative language skills • Supporting development of skills through scaffolding (modeling, prompting and practice) • Asking students to predict or summarize stories or factual information	
	☐ Teaching academic vocabulary in context	• Explicitly teaching a small group of words each week by defining and providing examples • Identifying and explicitly teaching a common set of words across grade level teams aligned with reading selection and curriculum standards • Providing extended opportunities to use academic vocabulary words • Integrating academic language skills throughout instruction (for example, in social studies, science or read-aloud)	
	☐ Other instructional practices to support vocabulary and relevant background knowledge development (specify in NOTES)		

Component type: INSTRUCTIONAL PRACTICES, STRUCTURES, AND ACADEMIC SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Instructional practices to build writing skills	☐ Providing instruction to develop writing quality	• Providing explicit instruction to writing effective, clear and well-organized text. • Providing instruction on particular kinds of writing such as narrative, exposition, or argument. • Teaching complexity or variation in words or sentence structure used, quality or richness of ideas, use of appropriate genre elements, organization of ideas, elaboration of ideas, style or voice, effectiveness of a story or argument.	
	☐ Teaching writing conventions	• Using written language in terms of syntax (sentence structure) or morphology (word inflections). • Teaching language mechanics (such capitalization and punctuation), word usage, grammar, or spelling.	
	☐ Other instructional practices to build writing skills (specify in notes)		

COMPONENT TYPE 2: NON-ACADEMIC STUDENT SUPPORTS

Non-academic student supports encompass non-academic efforts to support students and help them learn and achieve. In early literacy, this includes components such as teaching self-management skills or supporting parent involvement.

Does the program/intervention provide non-academic student supports? Yes/No

If "no," proceed to component type 3 below.

If "yes," please mark all component domains and components that apply.

Component type: NON-ACADEMIC STUDENT SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Social-emotional learning strategies	☐ Support development of a growth mindset	• Incorporating read-alouds that tell stories of overcoming challenges • Conducting activities where students practice using phrases that promote growth mindset • Displaying growth-mindset vocabulary in the classroom • Modeling growth mindset • Providing feedback using prompts that provide opportunities for self-evaluation • Providing feedback that focuses on effort and progress	
	☐ Teaching self-management skills	• Supporting the development of impulse control • Supporting the development stress management skills • Supporting the development self-discipline • Supporting the development self-motivation • Supporting the development goal-setting skills • Supporting the development of organizational skills	

Component type: NON-ACADEMIC STUDENT SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
	☐ Other social-emotional learning strategies (specify in NOTES)		
☐ Parent outreach and involvement	☐ Parent meetings and conferences to discuss learning and growth		
	☐ Programming to encourage parent involvement in reading activities	• This includes a broad array of parent outreach programming including virtual meetings, whole-school curriculum nights or assemblies, or other programming such as home visiting programs.	
	☐ Providing books to families	• Directly providing books to families to support reading at home.	
	☐ Other parent outreach and involvement (specify in NOTES)		

COMPONENT TYPE 3: ORGANIZATIONAL STRUCTURES AND SUPPORTS

Organizational structures and supports encompass broader administrative efforts (for example, setting district literacy goals and targets or building districtwide partnerships around literacy). These administrative efforts may occur within a school, community, district, or state.

Does the program/intervention encompass broader administrative efforts? Yes/No

If "no," proceed to component type 4 below.

If "yes," please mark all component domains and components that apply.

Component type: ORGANIZATIONAL STRUCTURES AND SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Programming to support improvements in school climate	☐ Schoolwide efforts to promote positive behaviors	• Planning for school climate improvements • Engaging stakeholders (such as family-school partnerships) • Collecting and reporting school climate data (surveys, focus groups, interviews) • Monitoring and evaluating school improvement efforts	
	☐ Implementing multi-tiered systems of support around school climate and behaviors		
	☐ Implementing anti-bullying programming		
	☐ Other programming to support improvements in school climate (specify in NOTES)		
	☐ After school programming		

Component type: ORGANIZATIONAL STRUCTURES AND SUPPORTS			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Providing out-of-school time supports	☐ Tutoring outside of school time		
	☐ Summer programming		
	☐ Partnering with community organizations		
	☐ Other provision of out-of-school time supports (specify in NOTES)		
☐ Providing opportunities for reducing the teacher/student ratio	☐ Reducing class sizes		
	☐ Scheduling small group literacy blocks		
	☐ Using trained assistants or paraprofessionals in literacy blocks		
	☐ Other opportunities for reducing the teacher/student ratio (specify in NOTES)		

COMPONENT TYPE 4: EDUCATOR SUPPORTS

Educator supports encompass direct supports provided to educators to support student learning, including professional development activities or materials to support program implementation such as teacher guides or sample lessons provided by a curriculum developer.

Does the program/intervention encompass educator support? Yes/No

If "no," proceed to component type 5 below.

If "yes," please mark all component domains and components that apply.

Component type: EDUCATOR SUPPORTS		
Component domain	**Component**	**NOTES**
☐ Professional development for teachers	☐ Supporting instructional practices	
	☐ Supporting the link between student assessment and practice	
	☐ Training in technology use	
	☐ Utilizing literacy coaches	
	☐ Other teacher professional development related to the intervention (specify in NOTES)	
☐ Implementing professional learning communities (PLCs)	☐ Building grade-level PLCs	
	☐ Building multi-grade level vertical teams to support curricular alignment	
	☐ Developing cross-functional professional learning communities that meet on a regular basis	
	☐ Providing common planning/prep time	
	☐ Other PLC activities related to the intervention (specify in NOTES)	
	☐ Accessing and utilizing curricular guides and developer handbooks	

Component type: EDUCATOR SUPPORTS		
Component domain	**Component**	**NOTES**
☐ Instructional support materials	☐ Accessing and utilizing sample lesson plans provided by developer	
	☐ Accessing and utilizing sample lesson plans provided by instructional staff	
	☐ Accessing and utilizing instructional logs	
	☐ Other instructional support materials (specify in NOTES)	

COMPONENT TYPE 5: ASSESSMENT AND PLACEMENT

Assessment and placement components include activities involving student assessments including those used for placement as well as those designed to help students identify their strengths and weaknesses. In the table below, please indicate assessment as activities that were implemented as part of the specific early literacy intervention (do not indicate any assessment activities in place that were not part of the intervention).

Does the program/intervention encompass assessment and placement components? Yes/No

If "no," proceed to the outcome section below.

If "yes," please mark all component domains and components that apply.

Component type: ASSESSMENT AND PLACEMENT			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Testing and screening	☐ Implementing universal screening for students in grades K-3	• Conducting screening at the beginning of the school year • Conducting a mid-year screening • Conducting adapted assessments for English learners • Providing assessment accommodations for students with disabilities	
	☐ Formative assessments through curriculum-based measurement		
	☐ Using centralized software platforms to monitor student progress		
	☐ Other testing and screening (specify in NOTES)		

Component type: ASSESSMENT AND PLACEMENT			
Component domain	**Component**	**Examples of component approach**	**NOTES**
☐ Student placement	☐ Implementing a multi-tiered system of support or response-to-intervention framework to identify students in need of different levels of supports	• Creating a building-level team for screening and progress monitoring • Providing core academic instruction for all youth. • Providing secondary supports for youth at risk for poor learning outcomes, • Providing supplemental instruction to small groups, • Providing intensive intervention for those who require additional supports. • Implementing regular progress monitoring for students at elevated risk or require additional supports • Using data to drive decision-making to identify students' needs for additional supports.	
	☐ Grouping small literacy groups based on homogeneous student skill level		
	☐ Providing literacy instruction to heterogeneous skill groups		
	☐ Other student placement (specify in NOTES)		

Intervention delivery

Instructions: This section captures elements of intervention delivery, including the setting, duration and frequency, instructor or provider, and group size. Please respond by providing the best available information about how the intervention was delivered in the study. For each section, please indicate how the intervention delivery differed across components. Use the "NOTES" section below to provide additional information or summarize any other aspects of how intervention delivery varied.

Setting

1. Where was the intervention implemented? (check all that apply)
   - ☐ Early childhood care center
   - ☐ Pre-K-12 school (check all that apply)
     - ☐ Public
     - ☐ Private
     - ☐ Parochial
     - ☐ Charter
     - ☐ Magnet
     - ☒ Title 1
   - ☐ Afterschool
   - ☐ Community-based organization
   - ☐ Other, specify: _____

2. Did the setting vary by component?
   - ☐ No.
   - ☐ Yes. Explain how the setting varied:

Duration and frequency

1. What was the duration of the intervention (in weeks, months, or school years) from beginning to end? _____
   - ☐ Duration unknown
   - ☐ Other, specify: _____

2. How frequently was the intervention delivered to students? (for example, "instruction was provided 2.5 hours per day 5 days/week with grades 1-2")
   _____
   _____

3. How frequently was the intervention delivered to teachers? (for example, "professional development was provided to teachers in 4 2-hour sessions over one summer")
   _____
   _____

4. Did the duration or frequency vary by component?
   - ☐ No.
   - ☐ Yes. Explain how the duration or frequency varied:

Instruction

1. Who provided instruction or support? (*check all that apply*)
   - ☐ Teacher
   - ☐ Paraprofessional
   - ☐ Tutor
   - ☐ Counselor or psychologist
   - ☐ Parent
   - ☐ Peer
   - ☐ Technology-based delivery (e.g., computer, app, phone)
     - ☐ Supervised by an instructor
     - ☐ Unsupervised
   - ☐ Researcher
   - ☐ Unknown
   - ☐ Other, specify: _____

2. Did instruction vary by component?
   - ☐ No.
   - ☐ Yes. Explain how instruction varied:

Group size

1. In what size groups was the intervention delivered? (*check all that apply*)

   - ☐ Broad program (implementation is coordinated across multiple schools or settings)
   - ☐ Schoolwide
   - ☐ Whole classroom or large group (> 6 students)
   - ☐ Small group (<=6 students)
   - ☐ Individual (1:1 support)
   - ☐ Group size unknown
   - ☐ Other, specify: _____

2. Did group size vary by component?
   - ☐ No.
   - ☐ Yes. Explain how group size varied:

Implementation challenges: Were any significant implementation challenges encountered during the study?

☐ No

☐ Unknown/unsure

☐ Yes, specify issues:

Additional notes related to intervention delivery: