META-ANALYSIS

The State of Current Reading Intervention Research for English Learners in Grades K–2: a Best-Evidence Synthesis

Garrett J. Roberts[1] · Colby Hall[2] · Eunsoo Cho[3] ·
Brooke Coté[1] · Jihyun Lee[4] · Bingxin Qi[5] · Jacklyn Van Ooyik[1]

Abstract

This best-evidence synthesis reviews the past 20 years of rigorous reading intervention research to identify effective programs of instruction for Grade K–3 English Learners (ELs), as well as to determine the average effect of reading instruction on reading outcomes for this population. We identified 10 studies, all of which only included students in Grades K, 1, and/or 2. These studies evaluated the effects of seven programs, reporting 76 effect sizes. We did not identify any studies that included Grade 3 ELs. To guide researchers and practitioners, we describe each program and discuss associated effect sizes in foundational skills, fluency, comprehension, and oral language. *Proactive Reading*, delivered in a small group setting, produced some of the largest effects on foundational skills, fluency, and reading comprehension outcomes. *Sound Partners* was also shown to be effective even when delivered for a shorter duration, in a one-to-one setting. Finally, for practitioners and researchers aiming to improve oral language outcomes, *Early Vocabulary Connection*, delivered to small groups of students for 20 weeks, had the largest effects on oral language outcome measures. The primary limitation of this review was the small number of studies meeting the best-evidence synthesis criteria. Future research is needed to better understand the impact of reading interventions on reading outcomes for ELs in Grades 2–3 and the impact of meaning-focused intervention on reading outcomes. In particular, additional research is needed to identify interventions that have the potential to meaningfully improve reading comprehension and oral language outcomes for K–3 ELs.

Keywords Synthesis · Reading · Intervention · Elementary school · English language learner · Best-evidence

---

✉ Garrett J. Roberts
garrett.roberts@du.edu

Extended author information available on the last page of the article

In the fall of 2017, 15.9% of kindergarten (K) students in the USA were English Learners (ELs; Hussar et al., 2020). In the USA, the proportion of ELs with reading difficulties (RD) is significantly greater than the proportion of English-monolingual (EM) students with RD. This is due to a combination of factors, including the significant contribution of English language vocabulary, morphological awareness, and syntactic knowledge to English language reading comprehension (e.g., Adlof & Catts, 2015; Gottardo et al., 2018; Kieffer & Lesaux, 2012; Lipka & Siegel, 2012; Tong et al., 2014). Economic disadvantage is another contributing factor. In the USA, EL students are disproportionately subject to the stressors and deprivation associated with economic disadvantage. In particular, Spanish-speaking ELs in the USA are more likely than EM students to have family incomes below or near poverty levels (Fry & Gonzales, 2008; Hernandez et al., 2008) and parents with relatively low levels of education and literacy (Capps, 2005; Hernandez et al., 2008). They are also more likely to be enrolled in under-resourced, low-performing schools (Capps, 2005; Cosentino de Cohen et al., 2005). As a result, many ELs in the USA have fewer opportunities to access texts and educational experiences that contribute to academic English literacy acquisition.

Reading difficulties can have serious consequences. Longitudinal research demonstrates that children who struggle with reading in the primary grades are very likely to continue to struggle through high school (e.g., Adlof et al., 2010; Boscardin et al., 2008; Francis et al., 1996; McNamara et al., 2011) and are at higher risk of school dropout, incarceration, and unemployment (Daniel et al., 2006; Greenberg et al., 2007; Hernandez, 2011). Fortunately, there is evidence that early reading interventions are effective in preventing this cascade of negative outcomes. The primary grades (K–3) represent a unique window of opportunity for early reading intervention, with the effects of reading intervention tending to be larger when they are delivered earlier rather than later (Al Otaiba et al., 2009; Lovett et al., 2017; Wanzek et al., 2018).

Much is known about effective early literacy instruction for EM students with or at risk for RD in the primary grades. More than three decades of research demonstrate that interventions with a primary focus on explicit, systematic instruction in foundational word reading skills (including alphabetic principle, phonological awareness [PA], phonics, and reading fluency) and some instruction focused on meaning (both on word meaning and on local and global comprehension of connected text) are associated with positive effects on measures of reading achievement (e.g., Gersten et al., 2020; Wanzek et al., 2016, 2018) for this population. Such multi-component reading instruction for students in the primary grades has the potential to reduce not only the severity but even the incidence of RD (Fletcher et al., 2018; Torgesen, 2002). Al Otaiba et al. (2009) reviewed intervention research evidence demonstrating that effective core (i.e., tier 1) reading instruction can reduce the number of children with below-level reading achievement to around 5–7%, and supplemental tutoring can reduce these numbers further to 0.2–2%.

Although there is much research establishing an understanding of effective early reading instruction for EM students, less is known about the degree to which this approach is effective for primary grade ELs. ELs may encounter challenges in acquiring proficient English reading skills for a variety of reasons. The degree to which decoding and language comprehension skills in the early elementary grades predict reading comprehension in the later elementary grades appears to be very similar for both EL and EM students (Kieffer & Lesaux, 2012; Lesaux et al., 2007). However, research suggests that, at least by the time students progress to the upper elementary grades, reading comprehension problems in ELs are, on average, more strongly predicted by linguistic comprehension difficulties than by decoding difficulties

(Babayiğit, 2014; Cho et al., 2019; Grant et al., 2011; Proctor et al., 2005), and the association between language comprehension difficulties and reading comprehension difficulties is stronger for EL students than for EM students (Babayiğit, 2014; Cho et al., 2019). Based on this research, it stands to reason that approaches to reading intervention for primary grade ELs might be more effective in the long term if they target oral language proficiency to a greater degree than do literacy intervention programs designed for primary grade EM students. This reasoning is reflected in the Institute of Education Sciences (IES) practice guide on teaching academic content and literacy to ELs (Baker et al., 2014), which calls for the integration of vocabulary and oral language instruction into reading instruction. Therefore, it is surprising that recent systematic reviews have either not investigated reading intervention effects on oral language outcomes or have identified few research studies investigating the effects of interventions that focus on building oral language comprehension (Richards-Tutor et al., 2016).

Recent Systematic Reviews on Reading Instruction for ELs

Two recent meta-analyses (Ludwig et al., 2019; Richards-Tutor et al., 2016) and one best-evidence synthesis (Cheung & Slavin, 2012) reviewed reading intervention research for ELs, including ELs in Grades K–3. In the best-evidence synthesis conducted by Cheung and Slavin (2012), the authors identified 18 studies with the majority of participants in Grades K–3. While not all effect sizes reported were positive (ES range on combined reading outcomes, −0.38 to 0.68), there were promising findings for a number of programs, including Direct Instruction (DI; Adams & Engelmann, 1996), English Language and Literacy Acquisition (ELLA; Irby et al., 2010), and Success for All (Slavin et al., 2009), as well as for instructional approaches that incorporated cooperative learning, small-group instruction, and one-to-one (1:1) interventions.

The meta-analysis published by Richards-Tutor et al. (2016) reported consistently positive, statistically significant effects on foundational word reading outcomes for the seven included studies that delivered early reading interventions to EL participants with or at risk for reading disabilities (RD) in Grades K–1 (ES range, 0.58–0.91); however, they noted mixed effects of interventions in the two studies with participants in Grades 2–3 (ES range, −0.13–1.00). Richards-Tutor et al. identified three studies with participants in Grades K–3 that reported effects on vocabulary/oral language outcomes, not all of which were positive (ES range, −0.17–0.78); five studies with participants in Grades K–3 reported effects on reading comprehension outcomes, which were positive but varied in effect size (ES range, 0.06–1.00). The last meta-analysis, conducted by Ludwig et al. (2019), identified 20 studies with students in Grades K–3 and found that reading interventions had positive effects overall, with mean effects of 1.22, 0.80, and 0.50 on reading accuracy, fluency, and comprehension outcomes, respectively.

All three previous reviews articulated criteria for inclusion that systematically excluded studies that were included in the present review. For example, the best-evidence synthesis conducted by Cheung and Slavin (2012) only included studies evaluating the effects of English reading interventions on reading performance for Spanish-speaking ELs and excluded studies that did not (a) measure reading comprehension, (b) have a minimum of a 12-week duration, and/or (c) employ a rigorous research design (i.e., random assignment of students to conditions or matching/statistical adjustment for pretest differences). The mean effect of all studies K–6 in Cheung and Slavin (2012) was 0.23, although the authors did not further

disaggregate findings (e.g., grade, outcome type), except by program type. The present synthesis differs from Cheung and Slavin in that we did not require EL participants to be Spanish speakers; we also included studies with a range of reading outcomes, including foundational word reading skills (e.g., PA, phonics) outcomes and/or language comprehension (e.g., vocabulary knowledge, listening comprehension) outcomes that are known to be robust predictors of reading comprehension.

The meta-analysis conducted by Richards-Tutor et al. (2016) only included studies for which EL participants were identified as having or being at risk for RD. We did not impose this requirement because many ELs who are at risk for RD are not identified as such in early grades; it is difficult to guarantee that young EL students have been properly screened and identified for risk for reading difficulties (Hutchinson et al., 2004). When EL students are assessed only in English, their difficulties with language acquisition and early literacy skills may be overestimated and attributed to reading difficulties when they are instead explained by students' still-developing English language proficiency (Sullivan, 2011). Conversely, there is some evidence that ELs may be identified for special education later than their EM peers and thus may not receive appropriate educational supports in their first years of school (O'Connor et al., 2013; Samson & Lesaux, 2009). Due to the unreliable methods for determining and interpreting risk status, we noted when studies only included students with RD (see Table 1) but did not exclude studies when participants were not identified as having or being at risk for RD.

In contrast to Ludwig et al. (2019), we included studies that set out to measure reading outcomes beyond reading accuracy, reading fluency, and reading comprehension. Our decision to analyze effects on additional outcomes was based on previous research (Babayiğit, 2014; Cho et al., 2019; Grant et al., 2011; Proctor et al., 2005) and an IES practice guide (Baker et al., 2014) that suggest the importance of building oral language/vocabulary knowledge and listening comprehension in EL students prior to the upper elementary grades. That said, our review was more exclusive than Ludwig et al. (2019) in one way: like Cheung and Slavin (2012) and Richards-Tutor et al. (2016), who established inclusion criteria that excluded studies with less stringent research designs/methods, we set out to include in this review only studies that met best-evidence synthesis requirements (Cheung & Slavin, 2005; Slavin et al., 2008; Slavin & Lake, 2008), enabling us to identify and describe in detail interventions that were demonstrated to be effective by rigorous research.

Purpose and Research Questions

This best-evidence synthesis aims to provide researchers and practitioners with guidance in understanding current K–3 EL reading interventions evaluated by means of rigorous, high-quality research studies. We followed best-evidence synthesis procedures described by Slavin (1986) and refined in subsequent studies by Slavin and colleagues (e.g., Cheung & Slavin, 2005; Slavin et al., 2008; Slavin & Lake, 2008). To be included in this best-evidence synthesis, studies needed to have (a) a sample size of at least 15 participants per condition, (b) a duration of at least 12 weeks such that reading instructional programs could be considered practical/have the potential to affect reading outcomes, and (c) either random assignment of students or a quasi-experimental research design that used matching to ensure equivalence between groups at baseline/post-hoc statistical adjustments to account for pretest differences. After identifying studies that met these best-evidence criteria, we described the programs in detail. We analyzed the relations between program components and effects achieved by these programs in order to

provide meaningful guidance for primary grade teachers of students who are ELs. Finally, to further contextualize recommendations that emerge from this best-evidence synthesis, we identified the quality of the counterfactual condition (i.e., alignment of instruction to evidence-based practices). We did this in response to calls for better measurement and description of the counterfactual and its inclusion of evidence-based reading practices, which enables the field to achieve a better understanding of the relative effectiveness of interventions (Lemons et al., 2014; Scammacca et al., 2016).

We believed this best-evidence synthesis was warranted for the following reasons. First, as described above, too little is known about the content of effective instruction for primary grade ELs, and, in particular, about the ways reading interventions can improve EL students' language comprehension. Second and relatedly, within the last twenty years, there has yet to be a best-evidence synthesis that has investigated the effects of reading interventions delivered to ELs on a diverse range of reading outcomes (i.e., not just on reading comprehension outcomes, which was the only type of outcome evaluated by Cheung & Slavin, 2012) and including interventions that had a duration of less than one year (i.e., less than the duration of studies included by Cheung and Slavin (2005)). We aimed to build on findings reported in prior research (Cheung & Slavin, 2012; Ludwig et al., 2019; Richards-Tutor et al., 2016) by conducting this best-evidence synthesis (Cheung & Slavin, 2005; Slavin et al., 2008; Slavin & Lake, 2008) and describing the effects of rigorously researched reading instruction for Grade K–3 ELs. We asked:

1. What is the overall mean effect of reading programs identified in this best-evidence synthesis, and what is the mean effect of best-evidence programs on disaggregated word reading/foundational word reading skill (e.g., PA, phonics, nonword reading, word reading), passage reading fluency, oral language (e.g., vocabulary knowledge, listening comprehension), and reading comprehension outcomes?
2. Which reading programs demonstrated the largest overall reading effect sizes?
3. Which programs demonstrated the largest effects on disaggregated word reading/ foundational word reading skill, passage reading fluency, oral language, and reading comprehension?
4. What are the characteristics of the programs that demonstrated the largest reading effect sizes within each outcome domain (i.e., word reading/foundational word reading skill, passage reading fluency, oral language, reading comprehension) and overall?

Method

Search Strategy and Inclusion Criteria

To identify studies for inclusion in this meta-analysis, we searched peer-reviewed research articles published between January 1, 2000, and December 31, 2020, to represent recent research on reading instruction for K–3 ELs. The search procedure followed PRISMA guidelines (Liberatti et al., 2009) with a three-step process and each step being independently completed by a minimum of two members of the research team. First, we searched the electronic databases of ERIC and psycINFO with a combination of the following terms: (read* or "phonological awareness" or "phonemic awareness" or vocab* or fluen* or decod* or comprehend* or lit* or "language arts") and ("English language learners" or "English

second language" or "English as a second language" or "second language learn*" or lingual* or "language minority") and ("elementary education" or "elementary school" or "primary education" or kinder* or "grade 1" or "grade 2" or "grade 3"). Studies were included if they met the following criteria:

1. At least 90% of participants were identified as ELs. Various labels (e.g., limited English proficient, language minority) were accepted, as long as labels conveyed that English was not the first language for participating students. Studies with fewer than 90% EL participants were included if disaggregated data were provided for EL students (Hall et al., 2017); only disaggregated data were used to calculate effect sizes in these cases. We queried authors when studies reported disaggregated findings for EL students but did not include enough data to calculate an effect size.

2. Participants were in Grades K–3 (ages 5–9), or the sample mean age was within the targeted range.

3. Reading instruction focused on PA, phonics/word reading, vocabulary, fluency, comprehension, or a combination of these domains.

4. The primary language of instruction was English (i.e., greater than 50% of instruction was delivered in English). However, studies were included when first language supports were incorporated into English language instruction.

5. The setting was school-based, including school-based after-school tutoring.

6. The study had at least one calculable effect size on a reading or oral language outcome.

7. The study used an experimental or quasi-experimental group research design with assignment of participants to at least one treatment and one comparison condition (defined as a no-treatment control condition with or without exposure to materials or a weaker instructional condition that mirrored classroom practice). For studies that did not use random assignment, researchers employed a matching procedure and/or statistical adjustments for pretest differences when appropriate. In this latter case, effect sizes without statistical adjustments were excluded. Historical comparison conditions were not included (e.g., Leafstedt et al., 2004).

8. Studies had a minimum duration of 12 weeks and a minimum of 15 students per condition.

9. Studies were conducted in the USA and were either unpublished dissertations, published in a peer-reviewed journal, or published as a technical report.

Our search yielded 4441 articles. Of these, 4339 articles were excluded on the basis of disqualifying information reported in article abstracts or because they were duplicates. We reviewed the full text of 102 articles. Of these, 92 articles were excluded because they were conducted outside the USA ($k = 9$); did not include or did not disaggregate findings for K–3 students ($k = 5$); had an ineligible comparison condition (e.g., comparison condition did not include 90% ELs; $k = 40$); had fewer than 90% EL participants and findings were not disaggregated for ELs ($k = 4$); primarily delivered instruction in a language other than English ($k = 5$); presented follow-up data to an already-published article ($k = 7$); did not deliver a reading intervention ($k = 6$); represented an unpublished dissertation which was later published as part of a published article that was reviewed (and excluded) or included in this meta-analysis ($k = 3$); delivered an intervention that was less than 12 weeks ($k = 6$); included an inadequate sample size ($k = 5$); and for studies that did not use random assignment, there was no student matching with appropriate adjustments for pretest differences ($k = 2$).

After the initial search, we completed an ancestral search of the reference sections of prior relevant literature reviews (Cheung & Slavin, 2012; Ludwig et al., 2019; Richards-Tutor et al., 2016) as well as of studies meeting the inclusion criteria of the current meta-analysis. Finally, we completed a hand search of published articles in the following journals, all of which either published articles included in this synthesis or tend to publish relevant articles related to reading and/or ELs: *American Educational Research Journal*, *Bilingual Research Journal*, *Exceptional Children*, *Journal of Behavioral Education*, *Journal of Educational Psychology*, *Journal of Literacy Research*, *Learning Disabilities Research and Practice*, *Learning Disability Quarterly*, *Reading and Writing Quarterly*, *Reading Research Quarterly*, *Scientific Studies of Reading*, and *The Elementary School Journal*. This hand search began with articles published in 2015 and concluded with articles published by December 31, 2020. No further articles meeting inclusion criteria were identified. A total of ten published articles (Baker et al., 2016; Ehri et al., 2007; Foorman et al., 2018; McMaster et al., 2008; Nelson et al., 2011; Tong, Lara-Alecio, et al., 2008b; Vadasy & Sanders, 2010, 2011; Vaughn, Cirino, et al. 2006; Vaughn, Mathes, et al. 2006) met inclusion criteria and were coded.

Coding Procedures

Code sheets captured information about study and participant characteristics, study outcomes, and WWC determinants of study quality (IES, 2020; Vaughn et al., 2014). All coders were trained by the first author. Prior to coding, all coders independently coded articles until they obtained a minimum of 90% reliability with the first author in each code sheet section (e.g., study characteristics, study outcomes). Once coding began, all articles were independently double coded by the first author and one additional coder. When coding discrepancies arose, coders reached consensus through re-reading and discussion.

One set of articles warrant additional explanation as to their coding (Tong, Irby, et al., 2008a, Tong, Lara-Alecio, et al., 2008b). Findings from Tong, Lara-Alecio, et al., 2008b) Tong, Irby, et al., 2008a) reported on findings from years two and three of the same longitudinal study. We opted to include only the year 2 study (Tong, Lara-Alecio, et al., 2008b), because it is most similar to the other studies in this review as no other study exceeded a duration of one year. In Tong, Lara-Alecio, et al. (2008b), we compared the Structured English Immersion-Enhanced/ Experimental group to the Structured English Immersion-Typical/Control group. We did not investigate results for the Transitional Bilingual Education-Enhanced/Experimental or the Transitional Bilingual Education-Typical/Control groups, because the primary language of instruction for those groups was Spanish.

Coding of Study Characteristics and Outcomes

In coding study characteristics and outcomes, we used processes similar to those used in previous meta-analyses of reading intervention studies (Wanzek et al., 2016, 2018). Similar to Wanzek et al. (2016), we coded for intervention type by differentiating between studies of interventions that focused on foundational skills, those that were meaning-focused, and those that were multi-component. Foundational skill interventions focused on PA, phonics, and word/nonword reading accuracy or fluency with or without a simultaneous focus on building passage reading fluency. Meaning-based interventions focused on vocabulary or comprehension (i.e., vocabulary knowledge, reading comprehension, listening comprehension). Multi-component interventions included a combination of foundational skills and meaning-based

components. We used similar categories to describe reading outcome domains, differentiating between (a) word reading/foundational word reading skills measures (we refer to this category as "foundational skills" measures going forward, and it includes measures of letter naming, PA, phonics, and word/non-word reading accuracy or fluency), (b) passage reading fluency measures, (c) oral language comprehension (i.e., vocabulary, listening comprehension), and (d) reading comprehension outcomes.

    We coded comparison conditions based on the extent to which the counterfactual condition utilized evidence-based instruction. Based on a modified coding scheme designed by Scammacca et al. (2016), we developed three categories of counterfactual quality (i.e., high, moderate, low) and a *not defined* category. The counterfactual condition was considered *high quality* if *all* students received an evidence-based, commercially available reading curriculum or an instructional program that was described as evidence-based instruction in phonemic awareness, phonics, fluency, vocabulary, comprehension (National Reading Panel, 2000); *moderate quality* if *some* students received a research-based, commercially available reading curriculum or explicit instruction in phonemic awareness, phonics, fluency, vocabulary, comprehension; and *low quality* if *no* students received such a curriculum or instructional program. The counterfactual condition was coded *attention control* if it constituted a condition defined as a researcher-delivered weaker treatment in a program other than reading (e.g., math, social studies). Finally, the counterfactual condition was considered *not defined* if the counterfactual was not described in sufficient detail to be included in one of the categories outlined above. To the best of our knowledge, this is the first systematic review (i.e., best-evidence synthesis, meta-analysis) on EL reading instruction to code for counterfactual quality.

Effect Size Calculations

We calculated standardized mean differences between intervention and control groups estimated with Hedges' *g* (Hedges, 1981), using the reported posttest mean and standard deviation (*SD*) by conditions.

    For each study, we calculated mean effect sizes separately for different reading outcomes and for the combined outcome using a random effects model with restricted maximum likelihood estimator (REML) for variance estimation. There were two programs (*Proactive Reading* and *Sound Partners*) that were evaluated in more than one study and each study reported multiple effect sizes. In such cases, program-level effect sizes were estimated using robust variance estimator along with small sample size correction (RVE; Hedges et al., 2010; Tipton, 2015) with a rho of .80.[1] All the analyses were conducted using *metafor* (Viechtbauer, 2010) and *robumeta* (Risher, Tipton,, & Zhipeng, 2017) packages in R (R Core Team, 2020).

Results

Table 1 presents study characteristics for the ten studies included in this best-evidence synthesis. Program and within-study effect sizes are reported in Tables 2, 3, 4, 5, and 6. Studies encompassed 76 individual effect sizes and a combined sample of 2150 students. Three of the ten studies were published between 2000 and 2009. The remaining seven studies

---

[1] Note that standard errors may be unreliable when degrees of freedom is less than 4 in random effects model with RVE

Table 1 Study characteristics

Study	Design	N	Grade	At risk for RD	Dosage (hours)	Duration	Group size	Instructor	Curriculum	Content	Counterfactual quality
Baker et al. (2016)	RCT	78	1	Yes	30	12-week	SG	Teacher, IA	Transition Lessons	PA, P, F, V, C	High
Ehri et al. (2007)	QED[a]	126	1	Yes	75	6-month	1:1	Teacher, IA	Reading Rescue Intervention Model	PA, P, F, V, C	Moderate
Foorman et al. (2018)	QED[b]	1157	K–2	Yes	101.25	27-week	SG	Researcher-hired, IA	Sound Partners, Bridge of Vocabulary, Language in Motion	PA, P, F, V, C	High
McMaster et al. (2008)	QED[a]	40	K	No	30	18-week	WC	Teacher	K-PALS	PA, P	Moderate
Nelson et al. (2011)	RCT	210	K	No	33.3	20-week	SG	IA	Early Vocabulary Connections	P, V	Low
Tong, Lara-Alecio, et al. (2008b)	QED[b]	213	K–1	No	NR	2-year	Varied	Teacher, IA	Santillana Intensive English, STELLA, AOL, AOLS, EIR, communication games	PA, P, F, V, C	Moderate
Vadasy and Sanders (2010)	RCT	84	K	Yes	27.68	18-week	1:1	IA	Sound Partners	P	Not defined
Vadasy and Sanders (2011)	RCT	98	1	Yes	33.15	20-week	1:1	IA	Sound Partners	P	Not defined
Vaughn, Cirino et al. (2006)	RCT	96	1	Yes	98.7	8-month	SG	Teacher	Proactive Reading	PA, P, F, V, C	Not defined
Vaughn, Mathes et al. (2006)	RCT	48	1	Yes	96.55	7-month	SG	Teacher	Proactive Reading	PA, P, F, V, C	Not defined

Note. *RD*, reading disability; *SG*, small group; *IA*, instructional assistant; *PA*, phonemic awareness; *P*, phonics; *F*, fluency; *vocab*, vocabulary; *comp*, comprehension; *WC*, whole class; *MTSS*, multi-tiered systems of support; *STELLA*, Story Telling for English Language and Literacy Education; *AOL*, Academic Oral Language; *AOLS*, Academic Oral Language in Science (AOLS); *EIR*, Early Intervention in Reading

[a] No randomization; matching at student level

[b] Randomized at school level; analysis at student level; no matching

were published between 2010 and 2020. Studies were conducted in kindergarten ($k = 3$), Grade 1 ($k = 5$), Grades K–1 ($k = 1$), or Grades K–2 ($k = 1$). No studies included students in Grade 3, and therefore, findings reflect the effects of interventions for EL students in Grades K–2. Seven of the ten studies solely included students with or at risk for reading difficulties. Studies provided instruction to students 1:1 ($k = 3$), in small groups ($k = 5$), and as a whole class ($k = 1$). One study delivered instruction within a school-wide multi-tiered system of support framework with students receiving instruction in various grouping arrangements. Counterfactuals were rated as high quality ($k = 2$), moderate quality ($k = 3$), and low quality ($k = 1$). The low-quality counterfactual was interactive book reading (Nelson et al., 2011). Four studies did not describe the counterfactual condition. No study had a counterfactual condition rated as attention control.

Three studies implemented foundational skill–focused interventions (McMaster et al., 2008; Vadasy & Sanders, 2010, 2011). McMaster et al. (2008) evaluated the effects of K-PALS, which included both PA and phonics components. The two studies reported by Vadasy and Sanders (2010, 2011) implemented a version of *Sound Partners* (e.g., Vadasy et al., 2006; Vadasy et al., 2008; Vadasy & Sanders, 2012). The remaining seven studies implemented multi-component interventions. Six studies delivered a multi-component intervention with all five reading instructional components (i.e., PA, phonics, fluency, vocabulary, comprehension; Baker et al., 2016; Ehri et al., 2007; Foorman et al., 2018; Tong, Lara-Alecio, et al., 2008b; Vaughn et al., 2006a, b). Baker et al. (2016) implemented a researcher-designed vocabulary-focused intervention titled *Transition Lessons*. The instructional intervention delivered by Ehri et al. titled *Reading Rescue* used a sequence of lessons aligned to fiction and non-fiction books that increased with difficulty over time. Foorman et al., Vaughn, Cirino, et al., and Vaughn, Mathes, et al. all used scripted, commercially available curriculums. Multi-component instruction in the study reported by Tong, Lara-Alecio, et al. was delivered within a school-wide multi-tiered system of support framework and included *Santillan Intensive English* (Ventriglia & González, 2000), *Story Telling for English Language and Literacy Education* (STELLA; Irby et al., 2004), *Academic Oral Language* (AOL; in kindergarten only; Lakeshore Learning Materials, 1997), *Academic Oral Language in Science* (AOLS; in first grade; Lakeshore Learning Materials, 1997), *Early Intervention in Reading* (EIR; Mathes & Torgesen, 2005), and communication games with students receiving varying levels of support and curriculum components based on need. Nelson et al. (2011) were the only multi-component study that did not include all five coded reading components. In Nelson et al. (2011), the authors delivered a commercially available phonics and vocabulary-focused reading intervention titled *Early Vocabulary Connections* (Nelson & Vadasy, 2007).

Reading Program Descriptions

Early Vocabulary Connections

Nelson et al. (2011) was the only study to implement the researcher-designed *Early Vocabulary Connections*. This program focuses on early vocabulary knowledge development, targeting word meaning, phonology, and orthography in an integrated fashion. The curriculum introduces 184 decodable root words in a way that systematically builds student knowledge of grapheme-phoneme correspondences. Instructional activities teach word meanings while also reinforcing the application of decoding skills (Nelson et al., 2011). Activities include defining, blending, and spelling target words; reading decodable text; and using cloze procedures,

matching word meanings to pictures, and sentence production tasks to demonstrate understandings of target words. While this program was designed to be delivered in a variety of instructional settings, Nelson et al. (2011) chose to evaluate the effects of instruction in a small-group format (2–5 children per group) in which the activities listed above were delivered by paraeducators in 20-minute sessions. Nelson et al. (2011) had a mean effect of 0.38 for the *Early Vocabulary Connections* intervention. Table 2 presents the within-study effects.

K-PALS

One study utilized *K-PALS* (McMaster et al., 2008), implementing a supplemental peer-tutoring program for kindergarten students. This program emphasizes building phonemic awareness, knowledge of letter-sound correspondences, decoding, and fluency using both teacher-directed activities and activities performed by students in pairs. By designating reciprocal roles to each pair of students (which included a higher-performing and a lower-performing reader), the program helps students facilitate their peers' learning; students take turns being both coach and reader (D. Fuchs et al., 2001). McMaster et al. (2008) implemented *K-PALS* as designed. Students engaged in Sound Play and Sounds and Words activities during 20- to 30-minute sessions. Sound Play was teacher-directed and consisted of 5 phonemic awareness games, lasting approximately 5–10 minutes. During the remaining 15–20 minutes, reciprocal pairs engaged in Sounds and Words activities in which they practiced identifying sounds, reading words, and reading sentences. McMaster et al. (2008) had a mean effect of 0.29 for *K-PALS* instruction. Table 3 presents the within-study effects.

Proactive Reading

Proactive Reading (Mathes & Torgesen, 2005) consists of a scripted curriculum that facilitates teachers' provision of explicit instruction in phonemic awareness, word recognition, and comprehension strategies. Because efficient word identification is a primary focus of this program, students spend much of each session learning and reviewing letter-sound correspondences, sounding out and reading words, and reading decodable connected text. They are also provided with opportunities to spell words and learn/apply comprehension strategies (Mathes & Torgesen, 2005). *Proactive Reading* is designed to be delivered in a small-group format.

Two studies utilized *Proactive Reading* (Vaughn et al., 2006a, b). Each study enhanced the curriculum by incorporating oral language development activities during each session. Interventions in both studies were delivered by teachers to groups of 3 to 5 students, 5 days a week, 50 minutes each day, for an average of 115 sessions. Vaughn et al., 2006a, b) designated 10 minutes of each session to vocabulary, listening comprehension, and academic language development. This included teaching two or three vocabulary words connected to the text, posing questions about vocabulary and key ideas, supporting students in retelling stories, and engaging students in dialogue. Vaughn, Mathes, et al., 2006b) also devoted 10 minutes of each session to the development of oracy and vocabulary. Both studies (Vaughn et al., 2006a, b) included additional instructional practices identified as effective in improving academic outcomes for ELs in previous research, including the use of visuals/gestures, clarification of English word meanings, and opportunities for teachers and students to work together to elaborate on student responses. Vaughn et al., 2006a, b) reported a mean effect of 0.08 and

Table 2 Early Vocabulary Connection and within-study effect sizes

Reading rrogram	Study	Foundational skills			Fluency			Reading comprehension			Oral language			Combined		
		S	U	All	S	U	All	S	U	All	S	U	All	S	U	All
Early Vocabulary Connection	Nelson et al. (2011)	.23 (.15)	NA	.23 (.15)	NA	NA	NA	NA	NA	NA	.24 (.15)	.67 (.15)	.45 (.22)	.23 (.01)	.67 (.15)	.38 (.15)

Notes. *S*, standardized; *U*, unstandardized

Table 3 K-PALS and within-study effect sizes

Reading program	Study	Foundational skills			Fluency			Reading comprehension			Oral language			Combined		
		S	U	All	S	U	All	S	U	All	S	U	All	S	U	All
K-PALS	McMaster et al. (2008)	.16 (.06)	.44 (.18)	.35 (.13)	NA	.14 (.04)	.14 (.04)	NA	NA	NA	NA	NA	NA	.16 (.06)	.34 (.13)	.29 (.10)

Notes. *S*, standardized; *U*, unstandardized

0.65, respectively. The mean program-level effect for *Proactive Reading* was 0.35. Table 4 presents within-program and within-study effects for studies that examined the effects of *Proactive Reading*.

Reading Rescue

Reading Rescue, evaluated by Ehri et al. (2007), is a comprehensive intervention model in which objectives, instructor training, observational tools, and progress-monitoring assessments are integral components that accompany curriculum materials and instructional procedures. This program, guided both by a scope and sequence of skills and by students' performance on assessments, explicitly and systematically provides instruction in phonological awareness, phonics, fluency, vocabulary knowledge, and comprehension. Teacher analysis of students' assessment performance and written observations during instruction ensures that instruction targets individual students' needs (Ehri et al., 2007). The amount of time varied for each component, as it was dependent upon each student's needs as determined by assessments/observations. *Reading Rescue* is designed to be delivered 1:1, and Ehri et al. (2007) mostly did so, with the intervention provided by certified reading specialists, educators certified in other areas (e.g., counseling, math, and social work), and paraprofessionals. In this quasi-experimental study, we only reported effects on the one measure that could be adjusted for pretest differences, a standardized composite measure of foundational skills and reading comprehension ($g = 0.50$); this was the only effect size from Ehri et al. (2007) included in this best-evidence synthesis.

Sound Partners

Sound Partners (Vadasy et al., 2006; Vadasy et al., 2008; Vadasy & Sanders, 2012) is a scripted reading program in which students receive daily, 30-minute sessions of 1:1 instruction in foundational reading skills, including phonemic awareness, phonics, and spelling. Each lesson begins with the introduction of a new grapheme-phoneme correspondence, followed by review of previously taught correspondences. This is followed by phoneme decoding in which students orally blend phonemes corresponding to graphemes on the page. Irregular words are then introduced and practiced through a read, spell, and reread format. Spelling is also integrated into the letter sounds and irregular words activities. At the conclusion of the lesson, students engage in oral reading of decodable texts through independent, partner, or echo reading practices.

Three studies examined the effects of *Sound Partners* (Foorman et al., 2018; Vadasy & Sanders, 2010, 2011). Both Vadasy and Sanders studies evaluated effects when paraeducators delivered *Sound Partners* 1:1, 4 days a week, during 30-min sessions for 18 weeks. Foorman et al. (2018) implemented *Sound Partners* for 25–30 minutes daily using locally hired interventionists and school paraprofessionals. Foorman et al. (2018) supplemented *Sound Partners* with *Bridge of Vocabulary* (Montgomery, 2007) instruction, which was delivered three times per week for 15 minutes each session to build listening, speaking, and reading skills using manipulatives. In addition, Foorman et al. (2018) included a language instruction supplement titled *Language in Motion*. *Language in Motion* (Phillips, 2014) was delivered two times per week to support students' syntax and language comprehension by means of activities that incorporated manipulatives, stories, and games. Foorman et al. (2018), Vadasy and Sanders, (2010), and Vadasy and Sanders (2011) reported mean effect sizes of −0.01, 0.56,

Table 4 Proactive Reading and within-study effect sizes

Reading program	Study	Foundational skills			Fluency			Reading comprehension			Oral language			Combined		
		S	U	All	S	U	All	S	U	All	S	U	All	S	U	All
Proactive Reading	Mean effect	.54 (.30)	.40 (.37)	.50 (.36)	.26 (.06)	NA	.26 (.06)	.54 (.51)	NA	.54 (.51)	.06 (.27)	NA	.06 (.27)	.33 (.27)	.40 (.37)	.35 (.28)
	Vaughn, Cirino et al. (2006)	.19 (.10)	.06 (.30)	.16 (.09)	.29 (.03)	NA	.29 (.03)	.06 (.21)	NA	.06 (.21)	−.16 (.02)	NA	−.16 (.02)	.09 (.07)	.06 (.30)	.08 (.06)
	Vaughn, Mathes et al. (2006)	.92 (.17)	.79 (.21)	.87 (.12)	.17 (.01)	NA	.17 (.01)	1.09 (.35)	NA	1.09 (.35)	.36 (.20)	NA	.36 (.20)	.62 (.14)	.79 (.21)	.65 (.12)

Notes. *S*, standardized; *U*, unstandardized

and 0.20, respectively. The mean program-level effect for *Sound Partners* was 0.22. Table 5 presents the within-program and within-study effects for studies using *Sound Partners.*

Santillana Intensive English, STELLA and AOL(S)

Tong, Lara-Alecio, et al., 2008b) delivered reading instruction in both tier 2 and tier 3 contexts for student participants with the lowest levels of reading performance. The tier 2 intervention utilized *Santillana Intensive English* as one of three components. The tier 2 program was designed to be used in a variety of instructional settings. It involves content-based English language instruction, during which a math, science, or social studies topic is explored over a four-day period. Lessons focus on the development of vocabulary and listening comprehension, but also of foundational reading skills. In Tong, Lara-Alecio, et al., 2008b), students in kindergarten and first grade received 40 minutes of daily language instruction, provided by either teachers or paraprofessionals. Vocabulary words were introduced and practiced using flashcards, stories, and role-play conversations with pairs and/or small groups of students.

In addition to *Santillana Intensive English*, Tong, Lara-Alecio, et al., 2008b) incorporated *STELLA* and *AOL/AOLS* into their tier 2 interventions. *STELLA* was used for 25 minutes daily in kindergarten and 40 minutes daily in first grade to support the development of higher-order thinking skills during reading and discussion of culturally relevant texts. Additionally, 10 minutes of each day was spent developing academic oral language via the *AOL/AOLS* interventions in both kindergarten and first grade, with the emphasis on general academic language in kindergarten and on the academic language of science in first grade.

Tong, Lara-Alecio, et al., 2008b) also included a tier 3 intervention for the lowest-performing students. The tier 3 intervention included an additional 10 and 20 minutes of researcher-developed communication games targeting vocabulary and phonemic awareness for students in kindergarten and the first half of first grade, respectively. The tier 3 intervention in the second half of first grade was *EIR* level I, a commercially available intervention with explicit and systematic instruction in phonemic awareness, fluency, and comprehension. This study included two standardized measures of oral language with an oral language composite score outcome of $g = 0.002$. This effect size represents an aggregate for students who received the tier 2 and the combined tier 2 and tier 3 interventions.

Transition Lessons

Baker et al. (2016) was the only study to use the researcher-developed *Transition Lessons* with small groups of students who were receiving tier 2 instruction. These lessons were designed for Spanish-speaking students who were transitioning from reading instruction in Spanish to reading instruction in English. Within this program, there were two fundamental components: the development of academic language through content vocabulary instruction, read-alouds of stories, and instruction in comprehension strategies; and skill-building practice in phonemic awareness, letter-sound knowledge, and decoding (Baker et al., 2016). Scripted lessons were taught by teachers and instructional assistants within an explicit instruction framework that included teacher explanations and modeling and scaffolded opportunities to respond and receive corrective feedback. This study had a mean effect of $g = -0.02$. Table 6 presents the within-study effects for Baker et al. (2016).

Table 5 Sound partners and within study effect sizes

Reading Program	Study	Foundational Skills			Fluency			Reading Comprehension			Oral Language			Combined		
		S	U	All	S	U	All	S	U	All	S	U	All	S	U	All
Sound Partners	Mean Effect	.12 (.12)	.57 (.25)	.20 (.16)	NA	.53 (.36)	.53 (.36)	.16 (.13)	NA	.16 (.13)	−.004 (.04)	NA	−.004 (.04)	.14 (.11)	.55 (.30)	.22 (.17)
	Foorman et al. (2018)	−.04 (.06)	NA	−.04 (.06)	NA	NA	NA	.02 (.07)	NA	.02 (.07)	−.004 (.04)	NA	−.004 (.04)	−.01 (.03)	NA	−.01 (.03)
	Vadasy & Sanders (2010)	.32 (.29)	.83 (.23)	.49 (.24)	NA	.90 (.23)	.90 (.23)	.47 (.22)	NA	.47 (.22)	NA	NA	NA	.37 (.18)	.86 (.03)	.56 (.16)
	Vadasy & Sanders (2011)	.18 (.002)	.33 (.20)	.23 (.05)	NA	.18 (.20)	.18 (.20)	.10 (.20)	NA	.10 (.20)	NA	NA	NA	.16 (.03)	.25 (.08)	.20 (.04)

Notes. *S* standardized; *U* unstandardized

Table 6 Sound Partners and within-study effect sizes

Reading program	Study	Foundational skills			Fluency			Reading comprehension			Oral language			Combined		
		S	U	All	S	U	All	S	U	All	S	U	All	S	U	All
Transitions	Baker et al. (2016)	−.11 (.05)	.19 (.23)	−.04 (.08)	−.12 (.23)	.19 (.23)	.03 (.15)	−.09 (.23)	.19 (.23)	.05 (.14)	−.06 (.08)	.22 (.04)	.08 (.09)	−.09 (.03)	.22 (.04)	−.02 (.06)

Notes. *S*, standardized; *U*, unstandardized

Discussion

This best-evidence synthesis sets out to examine the effects of reading interventions on reading and language outcomes for K–2 EL students, as reported in current studies that employed rigorous research designs/methods. Seven programs were implemented across ten studies. Across 76 effect sizes there was an overall mean effect of 0.23, which is smaller than the effect sizes presented in Ludwig et al. (2019; ES range = 0.50–1.22) and the K–1 effect sizes in Richards-Tutor et al. (2016; 0.58–0.91) but not out of line with the effect size range (−0.13–1.00) reported by Richards-Tutor et al. (2016) for the two studies with participants in Grades 2–3. The 0.23 effect size we identified was the same effect size reported in Cheung and Slavin (2012; i.e., 0.23). The fact that the mean effect size for this best-evidence synthesis was smaller than ones reported in meta-analyses that allowed for a wider range of study quality in their inclusion criteria reflects the previous research finding (Cheung & Slavin, 2012; Hall et al., 2017; Scammacca et al., 2015; Slavin & Smith, 2009) that effect sizes associated with higher-quality studies tend to be smaller than those associated with lower-quality studies. In the present meta-analysis, mean effects were largest on fluency outcomes ($g = 0.30$), followed by reading comprehension ($g = 0.27$) and foundational skill outcomes ($g = 0.27$), and finally oral language ($g = 0.11$) outcomes. This fact that effects on oral language were relatively small is concerning, given the significant contribution of oral language to reading comprehension for ELs in the upper elementary grades (e.g., Babayiğit, 2014; Cho et al., 2019). This meta-analysis differed from previous meta-analyses (i.e., Cheung & Slavin, 2012; Ludwig et al., 2019) that did not report effects on oral language/vocabulary outcomes or did not include studies that investigated effects on oral language/vocabulary outcomes without also investigating effects on reading comprehension outcomes. However, even given our more expansive criteria for inclusion, we did not identify many studies that yielded practically significant immediate effects on oral language outcomes; in fact, only the study conducted by Nelson et al. (2011) fits this description. This finding highlights the need to develop and pilot interventions that show more promise in improving oral language outcomes for ELs, a need that has been articulated in previous reviews (Baker et al., 2014; Richards-Tudor et al., 2016). It may be necessary to provide multiple years of content-rich, academic language–focused intervention to see meaningful effects; nevertheless, research suggests that language comprehension is a uniquely powerful contributor to reading comprehension for EL students and this is therefore worth investigating in future research.

There were several programs that were found to be effective, although it is worth noting that the studies found to be the most effective had counterfactual qualities that tended to be lower or undefined. In terms of building foundational skills, the multi-component *Proactive Reading* curriculum (Vaughn et al., 2006a, b) produced a mean foundational skill effect of $g = 0.50$ after seven to eight months of instruction. This effect size was larger than that achieved by any other intervention that targeted foundational skills outcomes. Provision of 18 weeks of the foundational skill–focused *Sound Partners* (Vadasy & Sanders, 2010) produced a mean effect size of $g = 0.49$, although in the 20-week Vadasy and Sanders (2011) study of *Sound Partners*, the intervention produced a smaller mean foundational skill effect size of $g = 0.24$. *Proactive Reading* was designed to be delivered in a small group setting; *Sound Partners* was designed to be delivered in a 1:1 setting. Both *Proactive Reading* and *Sound Partners* had samples of students with or at risk for RD and an undefined counterfactual quality. The whole-class *K-PALS* intervention study (McMaster et al., 2008) had a comparable mean effect on foundational skill outcomes ($g = 0.38$) when delivered for 18 weeks and a counterfactual quality

rating of moderate; this intervention is thus a good option when instruction must be delivered in a whole-class setting.

Effects on fluency outcomes were largest for the 1:1 *Sound Partners* intervention (Vadasy & Sanders, 2010, 2011; $g = 0.53$) and for small-group *Proactive Reading* (Vaughn et al., 2006a, b; $g = 0.26$), both of which had a sample of students with or at risk for RD and an undefined counterfactual quality. The mean effect size ($g = 0.14$), when compared to a moderate-quality counterfactual condition, achieved by the whole-class *K-PALS* intervention (McMaster et al., 2008) was not much smaller than that achieved by small-group *Proactive Reading*. Furthermore, *K-PALS* may be more feasible as it was delivered for a shorter duration (18 weeks) than was *Proactive Reading* (7–8 months).

Of the studies that measured reading comprehension, all exclusively included students with or at risk for RD. Two long-duration, multi-component programs produced the largest effect sizes. *Proactive Reading* (Vaughn et al., 2006a, b) produced an effect size of $g = 0.54$ relative to an undefined counterfactual condition. *Reading Rescue* (Ehri et al., 2007) produced an effect size of $g = 0.50$, relative to a counterfactual condition in which some students received a research-based, commercially available reading curriculum or explicit instruction in phonemic awareness, phonics, fluency, vocabulary, comprehension (i.e., moderate quality rating). *Reading Rescue* had a duration of six months, and the *Proactive Reading* intervention was slightly longer at seven to eight months. The shorter duration phonics-only *Sound Partners* intervention yielded a similar effect of $g = 0.47$ in one study (Vadasy & Sanders, 2010), but a much smaller effect size of $g = 0.10$ in another (Vadasy & Sanders, 2011). It should be noted that, while foundational skills have been shown to contribute to reading comprehension, reading comprehension is the ultimate goal of reading instruction. We acknowledge that it is difficult to measure reading comprehension in very young students (e.g., in kindergartners) who are able to read only the simplest of texts. Still, future research with K–2 ELs should take care to measure the degree to which improvements in foundational skills transfer to reading comprehension when possible.

Again, the smallest effects were reported for oral language outcomes. Across the five programs with measures of oral language, four had small oral language effect sizes: *Transitions* ($g = 0.08$; Baker et al., 2016), *Santillan Intensive Language* as part of a school-wide system of support ($g = 0.002$; Tong, Lara-Alecio, et al., 2008b), *Sound Partners* as part of a multi-program package ($g = -0.002$; Foorman et al., 2018), and *Proactive Reading* ($g = .06$; mean effect from Vaughn, Cirino, et al., (2006) and Vaughn, Mathes, (2006)). Only *Early Vocabulary Connection* (Nelson et al., 2011), the intervention that dedicated the greatest percentage of instructional time to building oral language skills, produced meaningful effects on oral language, with an effect size of 0.24 and 0.67 on standardized and unstandardized measures, respectively (mean effect on oral language was $g = 0.45$). For those considering *Early Vocabulary Connection* (Nelson et al., 2011) as an intervention to support EL oral language outcomes, it is worth considering the fact that the counterfactual condition employed in the study reported by Nelson et al. (2011) consisted of interactive book reading (i.e., low rating) and the sample of students did not only include students with or at risk for RD.

In this best-evidence synthesis as in all syntheses, effect sizes need to be interpreted relative to the counterfactual conditions in place in included studies. Counterfactual conditions are dynamic conditions that can vary in quality over time, as evidence-based instructional best practices are disseminated (Lemons et al., 2014) and other contextual factors. Therefore, we described the quality of the counterfactual (high, moderate, low, or not defined) for all included studies. Both Baker et al. (2016) and Foorman et al. (2018) included a high-quality

counterfactual condition. The strength of the counterfactual condition in both of these studies may be related to the fact that they had the lowest overall mean effect sizes, ranging from $g = -0.01$ to -0.02. Three studies compared the effects of an intervention with a moderately rated counterfactual condition (Ehri et al., 2007; McMaster et al., 2008; Tong, Lara-Alecio, et al., 2008b), with mean effect sizes ranging from $g = 0.002$ to 0.50. It is worth noting that Ehri et al. (2007) reported an effect size of $g = 0.50$ on a standardized, composite measure of foundational reading skills and reading comprehension. This is among the largest effect sizes reported in any study and the largest reported in a study with a counterfactual condition rated as moderate or high quality. That said, because only one effect size meeting best-evidence synthesis criteria (e.g., pretest mean or similar information was provided so that it was possible to statistically adjust for pretest differences between groups in quasi-experimental studies) could be calculated for the Ehri et al. (2007) study, conclusions that can be drawn from this evidence are somewhat limited. Nelson et al. (2011) was the only study to have a low-quality counterfactual; it reported an effect size of $g = 0.38$ in favor of the intervention condition. The four studies with undefined counterfactual conditions had the largest range of mean effect sizes ($g = 0.08$ to 0.65). Descriptively, this synthesis does provide some evidence that aligns with observations made by Lemons et al. (2014) that reading effect sizes tend to decrease when the counterfactual condition is of higher quality (i.e., employing evidence-based practices).

Limitations and Future Research

This meta-analysis has limitations that suggest topics for future research. First, only ten studies over the past 20 years met the criteria for inclusion in this best-evidence meta-analysis, such that there was inadequate power to investigate the effects of potential moderator variables. The limited number of studies identified also resulted in the majority of the programs ($k = 5$) having only one study to demonstrate program-level findings. Furthermore, the remaining programs evaluated in more than one study (i.e., *Proactive Reading*, *Sound Partners*) often demonstrated significant variation in effect sizes within programs, across studies. This limited program-level research with effect size variation in similar studies of similar programs highlights the need for more research, including replication studies, evaluating the effects of reading instruction designed to improve outcomes for EL students. Funding and conducting future studies that rigorously evaluate the effects of reading programs for primary grade ELs would allow researchers to better understand how student and intervention characteristics impact intervention effects and could enable educators to better individualize interventions for K–3 ELs. These studies would do well to also investigate the effects of interventions in Grades 2–3, as only Foorman et al. (2018) included students enrolled in grades beyond Grade 1 in their sample, and no study included Grade 3 ELs. Among other limitations, the measurement of reading comprehension is limited in Grades K–1; it would be valuable to determine the effects of reading interventions on reading comprehension for ELs in Grades 2 and 3.

Furthermore, there is a need (also articulated by Richards-Tutor et al., 2016) for more research on interventions that target vocabulary and language development. This best-evidence synthesis also found a need for more research in this domain. In particular, the findings from Nelson et al.'s (2011) study of *Early Vocabulary Connection*, which spent the majority of instructional time on oral language activities, suggests that interventions targeting academic language development in young EL students (with or without reading difficulties) have the potential to produce large effects on oral language outcomes, and

effects on oral language that are larger than those achieved by studies focusing primarily on foundational skills and/or on multiple components with a lesser focus on building academic language knowledge. Finally, there is also value in measuring intervention follow-up effects in order to determine how intervention impacts are sustained over time. It is even possible that intervention impacts may increase over time as a result of increasing students' access to grade-level texts in the upper elementary grades, when a greater percentage of classroom time is spent *reading to learn* as compared to *learning to read* (Chall & Jacobs, 1983).

This best-evidence synthesis also revealed that studies with high-quality counterfactuals tended to have lower reading outcomes than studies with low-quality counterfactuals. This suggests that counterfactual quality is worth considering when interpreting intervention and synthesis findings. Future research needs to acknowledge that not all counterfactual conditions are equal. When studies compare an intervention to a high-quality counterfactual condition, the intervention's effectiveness may appear to be reduced. So that future meta-analyses and syntheses can control for counterfactual quality, researchers need to better measure and document the nature of business-as-usual instruction received by students; note that four of the ten studies included in this best-evidence synthesis did not include counterfactual descriptions.

Recommendations for Practice

This best-evidence synthesis demonstrates the effectiveness of a number of commercially available programs and/or approaches that are replicable by educators. *Proactive Reading*, delivered in a small group setting to students with or at risk for RD for nearly a school year, produced consistently large effects on foundational skills, fluency, and reading comprehension outcomes, if not on oral language outcomes. Outcomes were similar but smaller for *Sound Partners*, whose studies also included samples of students with or at risk for RD. Outcomes for *Sound Partners* were shown to be effective even when delivered for a shorter duration, when delivered in a 1:1 setting. Even though *Reading Rescue* was associated with only one effect size in this best-evidence synthesis, the fact that the composite foundational skill and reading comprehension standardized measure effect size was $g = 0.50$ makes it worth considering when resources are available to provide 6 months of intervention in a 1:1 setting. Additionally, *Early Vocabulary Connection*, delivered in a small group setting for 20 weeks, had the largest effects on oral language outcomes. Finally, for administrators or educators looking to identify an effective whole-class intervention, *K-PALS* is a feasible program shown to improve foundational skills and fluency. These latter two studies included students with diverse reading abilities rather than requiring included students to be with or at risk for RD.

In summary, this best-evidence synthesis suggests that reading interventions targeting foundational reading skills, fluency, and comprehension have the potential to benefit reading outcomes for K–2 ELs. That said, this best-evidence synthesis clearly identified that more high-quality research is still needed; having a sufficient sample of high-quality, rigorous studies to analyze will provide greater confidence in findings. Additionally, this best-evidence synthesis documented a need for evaluations of reading interventions with ELs in Grades 2, and a better documentation of business-as-usual reading practices. Finally, more research is needed to better understand the impact of interventions targeting oral language on improved reading outcomes for K–2 ELs.

References

Adams, G., & Engelmann, S. (1996). *Research on direct instruction: 25 years beyond DISTAR*. Educational Achievement Systems.

Adlof, S., & Catts, H. (2015). Morphosyntax in poor comprehenders. *Reading and Writing, 28*(7), 1051–1070. https://doi.org/10.1007/11145.2015.95623.

Adlof, S., Catts, H., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities, 43*(4), 332–345. https://doi.org/10.1177/00222.2010.0369067.

Al Otaiba, S., Connor, C. M., Foorman, B., Schatschneider, C., Greulich, L., & Sidler, J. F. (2009). Identifying and intervening with beginning readers who are at-risk for dyslexia: advances in individualized classroom instruction. *Perspectives on Language and Literacy, 35*, 13–19.

Babayiğit, S. (2014). The role of oral language skills in reading and listening comprehension of text: a comparison of monolingual (L1) and bilingual (L2) speakers of English language. *Journal of Research in Reading, 37*(1), S22–S47. https://doi.org/10.1111/j.1467-9817.2012.01538.x.

Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., Gersten, R., Haymond, K., Kieffer, M. J., Linan-Thompson, S., & Newman-Gonchar, R. (2014). Teaching academic content and literacy to English learners in elementary and middle school (NCEE 2014-4012). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: http://ies.ed.gov/ncee/wwc/publications_reviews.aspx.

*Baker, D., Burns, D., Kame'enui, E., Smolkowski, K., & Baker, S. (2016). Does supplemental instruction support the transition from Spanish to English reading instruction for first-grade English learners at risk of reading difficulties? *Learning Disability Quarterly*, *39*, 226-239. https://doi.org/10.1177/073.2016.715616757, 4

Boscardin, C., Muthén, B., Francis, D., & Baker, E. (2008). Early identification of reading difficulties using heterogeneous developmental trajectories. *Journal of Educational Psychology, 100*(1), 192–208. https://doi.org/10.1037/00220663.100.1.192.

Capps, R. (2005). *The new demography of America's schools*. Urban Institute.

Chall, J. S., & Jacobs, V. A. (1983). Writing and reading in the elementary grades: developmental trends among low SES children. *Language Arts, 60*, 617–626.

Cheung, A., & Slavin, R. E. (2005). Effective reading programs for English language learners and other language-minority students. *Bilingual Research Journal, 29*(2), 241–267. https://doi.org/10.1080/15235882.2005.10162835.

Cheung, A., & Slavin, R. (2012). Effective reading programs for Spanish-dominant English language learners (ELLs) in the elementary grades. *Review of Educational Research, 82*(4), 351–395. https://doi.org/10.3102/0034654312465472.

Cho, E., Capin, P., Roberts, G., Roberts, G., & Vaughn, S. (2019). Examining sources and mechanisms of reading comprehension difficulties: comparing English learners and non-English learners within the simple view of reading. *Journal of Educational Psychology, 111*(6), 982–1000. https://doi.org/10.1037/edu0000332.

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities, 39*(6), 507–514. https://doi.org/10.1177/00222194060390060301.

Cosentino de Cohen, Deterding, C. N., & Clewell, B. C. (2005). Who's left behind? Immigrant children in high and low LEP schools. *Urban Institute (NJ3)*.

*Ehri, L. C., Dreyer, L. G., Flugman, B., Gross, A. (2007). Reading rescue: an effective tutoring intervention model for language-minority students who are struggling readers in first grade. *American Education Research Journal*, *44*, 414-448. https://doi.org/10.3102/2F0002831207302175, 2

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2018). *Learning disabilities: from identification to intervention*. Guilford Publications.

*Foorman, B. R., Herrera, S., & Dombek, J. (2018). The relative impact of aligning tier 2 intervention materials with classroom core reading materials in grades K-2. *The Elementary School Journal*, *118*, 477-504. https://doi.org/10.1086/696021, 3

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: a longitudinal, individual growth curves analysis. *Journal of Educational Psychology, 88*(1), 3–17. https://doi.org/10.1037/00220663.88.1.3.

Fry, R., & Gonzales, F. (2008). One-in-five and growing fast: a profile of Hispanic public school students. Pew Research Center.

Fuchs, D., Fuchs, L. S., Al Otaiba, S., Thompson, A., Yen, L., McMaster, K. N., Svenson, E., & Yang, N. J. (2001). K-PALS: Helping kindergartners with reading readiness: teachers and researchers in partnerships. *Teaching Exceptional Children, 33*(4), 76–80. https://doi.org/10.1177/2F004005990103300411.

Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., & Jayanthi, M. (2020). Meta-analysis of the impact of reading interventions for students in the primary grades. *Journal of Research on Educational Effectiveness, 13*(2), 401–427. https://doi.org/10.1080/19345747.2019.1689591.

Gottardo, A., Mirza, A., Koh, P. W., Ferreira, A., & Javier, C. (2018). Unpacking listening comprehension: the role of vocabulary, morphological awareness, and syntactic knowledge in reading comprehension. *Reading and Writing, 31*(8), 1741–1764. https://doi.org/10.1007/s11145-017-9736-2.

Grant, A., Gottardo, A., & Geva, E. (2011). Reading in English as a first or second language: the case of grade 3 Spanish, Portuguese, and English speakers. *Learning Disabilities Research & Practice, 26*(2), 67–83.

Greenberg, E., Dunleavy, E., & Kutner, M. (2007). *Literacy behind bars: results from the 2003 national assessment of adult literacy prison survey.* NCES 2007-473. National Center for Education Statistics.

Hall, C., Roberts, G. J., Cho, E., McCulley, L. M., Carroll, M., & Vaughn, S. (2017). Reading instruction for English language learners in the middle grades: a synthesis of the research. *Educational Psychology Review, 29*(4), 763–794. https://doi.org/10.1007/s10648-016-9372-4.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107–128.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. https://doi.org/10.1002/jrsm.5.

Hernandez, D. J. (2011). *Double jeopardy: how third-grade reading skills and poverty influence high school graduation.* Annie E. Casey Foundation. http://www.aecf.org/m/resourcedoc/AECF-Double-Jeopardy-2012-Full.pdf

Hernandez, D. J., Denton, N. A., & Macartney, S. E. (2008). Children in immigrant families: looking to America's future. *Society for Research in Child Development, 22*, 1–24. https://doi.org/10.1002/j.23793988.2008.tb00056.x.

Hussar, B., Zhang, J., Hein, S., Wang, K., Roberts, A., Cui, J., Smith, M., Bullock Mann, F., Barmer, A., & Dilig, R. (2020). *The condition of education 2020 (NCES 2020-144). U.S. Department of Education.* National Center for Education Statistics Retrieved March 1, 2020, from https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020144.

Hutchinson, J. M., Whiteley, H. E., Smith, C. D., & Connors, L. (2004). The early identification of dyslexia: children with English as an additional language. *Dyslexia, 10*(3), 179–195. https://doi.org/10.1002/dys.275.

Institute of Education Sciences. (2020). What Works Clearinghouse standards handbook version 4.1. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf

Irby, B. J., Lara-Alecio, R., Quiros, A. M., Mathes, P., & Rodriguez, L. (2004). *English Language and Literacy Acquisition evaluation research program (Project ELLA): Second annual evaluation report.* U.S. Department of Education, Institute for Educational Sciences.

Irby, B. J., Tong, F., Lara-Alecio, R., Mathes, P. G., Acosta, S., & Guerrero, C. (2010). Quality of instruction, language of instruction, and Spanish-speaking English language learners' performance on a state reading achievement test. *Texas Association for Bilingual Education, 12*, 1–42.

Kieffer, M. J., & Lesaux, N. K. (2012). Development of morphological awareness and vocabulary knowledge in Spanish-speaking language minority learners: a parallel process latent growth curve model. *Applied PsychoLinguistics, 33*(1), 23–54. https://doi.org/10.1017/S0142716411000099.

Lakeshore Learning Materials. (1997). *Question of the day.* Author.

Leafstedt, J. M., Richards, C. R., & Gerber, M. M. (2004). Effectiveness of explicit phonological-awareness instruction for at-risk English learners. *Learning Disabilities Research & Practice, 19*(4), 252–261. https://doi.org/10.1111/j.1540-5826.2004.00110.x.

Lemons, C. J., Fuchs, D., Gilbert, J. K., & Fuchs, L. S. (2014). Evidence-based practices in a changing world: reconsidering the counterfactual in education research. *Educational Researcher, 43*(5), 242–252. https://doi.org/10.3102/2F0013189X14539189.

Lesaux, N. K., Rupp, A. A., & Siegel, L. S. (2007). Growth in reading skills of children from diverse linguistic backgrounds: findings from a 5-year longitudinal study. *Journal of Educational Psychology, 99*(4), 821–834.

Liberatti, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine, 6*(10), 1–28. https://doi.org/10.1016/j.jclinepi.2009.06.006.

Lipka, O., & Siegel, L. S. (2012). The development of reading comprehension skills in children learning English as a second language. *Reading and Writing, 25*(8), 1873–1898. https://doi.org/10.1007/s11145-011-9309-8.

Lovett, M. W., Frijters, J. C., Wolf, M., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2017). Early intervention for children at risk for reading disabilities: the impact of grade at intervention and individual differences on intervention outcomes. *Journal of Educational Psychology, 109*(7), 889–914. https://doi.org/10.1037/edu0000181.

Ludwig, C., Guo, K., & Georgiou, G. K. (2019). Are reading interventions for English language learners effective? A meta-analysis. *Journal of Learning Disabilities, 52*(3), 220–231. https://doi.org/10.1177/0022219419825855.

Mathes, P. G., & Torgesen, J. K. (2005*). Early interventions in reading*, *level 1*. SRA/McGraw-Hill.

*McMaster, K. L., Kung, H., Han, I., & Cao, M. (2008). Peer-assisted learning strategies: a "tier 1" approach to promoting English learners' response to intervention. *Exceptional Children*, *74*, 194-214. https://doi.org/10.1177/001440290807400204, 2

McNamara, J. K., Scissons, M., & Gutknecht, N. (2011). A longitudinal study of kindergarten children at risk for reading disabilities: the poor really are getting poorer. *Journal of Learning Disabilities, 44*(5), 421–430. https://doi.org/10.1177/0022219411410040.

Montgomery, J. (2007). *Bridge of vocabulary.* Pearson.

National Reading Panel. (2000). *Report of National Reading Panel: teaching children to read*. National Institute of Child Health and Human Development.

Nelson, J. R., & Vadasy, P. F. (2007). Early vocabulary connections: first words to know and decode. Sopris West.

*Nelson, J. R., Vadasy, P. F., & Sanders, E. A. (2011). Efficacy of a tier 2 supplemental root word vocabulary and decoding intervention with kindergarten Spanish-speaking English learners. *Journal of Literacy Research*, *43*, 184-211. https://doi.org/10.1177/1086296X11403088, 2

O'Connor, R. E., Bocian, K. M., Beach, K. D., Sanchez, V., & Flynn, L. J. (2013). Special education in a 4-year Response to Intervention (RtI) environment: characteristics of students with learning disability and grade of identification. *Learning Disabilities Research & Practice, 28*(3), 98–112. https://doi.org/10.1111/ldrp.12013.

Phillips, B. (2014). Promotion of syntactical development and oral comprehension: development and initial evaluation of a small-group intervention. *Child Language Teaching and Therapy, 30*(1), 63–77. https://doi.org/10.1177/0265659013487742.

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: toward a model of comprehension. *Journal of Educational Psychology, 97*(2), 246–256. https://doi.org/10.1037/0022-0663.97.2.246.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2016). The effectiveness of reading interventions for English learners: a research synthesis. *Exceptional Children, 82*, 144–169. https://doi.org/10.1177/0014402915585483.

Risher, Tipton and Zhipeng (2017). robumeta: robust variance meta regression. R package version 2.0. https://CRAN.R-project.org/package=robumeta

Samson, J. F., & Lesaux, N. K. (2009). Language-minority learners in special education: rates and predictors of identification for services. *Journal of Learning Disabilities, 42*(2), 148–162. https://doi.org/10.1177/0022219408326221.

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (2015). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities, 48*(4), 369–390. https://doi.org/10.1177/0022219413504995.

Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., & Carroll, M. (2016). A century of progress: reading interventions for students in grades 4–12, 1914–2014. *Review of Educational Research, 86*(3), 756–800. https://doi.org/10.3102/0034654316652942.

Slavin, R. E. (1986). Best-evidence synthesis: an alternative to meta- analytic and traditional reviews. *Educational Researcher, 19*, 141–162 jstor.org/stable/42900104.

Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: a best-evidence synthesis. *Review of Educational Research, 78*(3), 427–515. https://doi.org/10.3102/0034654308317473.

Slavin, R. E., & Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic review in education. *Educational Evaluation and Policy Analysis, 31*(4), 500–506. https://doi.org/10.3102/2F0162373709352369.

Slavin, R. E., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: a best-evidence synthesis. *Reading Research Quarterly, 43*(3), 290–322. https://doi.org/10.1598/RRQ.43.3.4.

Slavin, R. E., Madden, N. A., Chambers, B., & Haxby, B. (2009). *2 million children: Success for All* (2nd ed.). Corwin Press.

Sullivan, A. L. (2011). Disproportionality in special education identification and placement of English language learners. *Exceptional Children, 77*(3), 317–334. https://doi.org/10.1177/2F001440291107700304.

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20*(3), 375–393. https://doi.org/10.1037/met0000011.

Tong, F., Irby, B. J., Lara-Alecio, R., & Mathes, P. G. (2008a). English and Spanish acquisition by Hispanic second graders in developmental bilingual programs: a 3-year longitudinal randomized study. *Hispanic Journal of Behavioral Sciences, 30*(4), 500–529. https://doi.org/10.1177/0739986308324980.

Tong, F., Lara-Alecio, R., Irby, B., Mathes, P., & Kwok, O. (2008b). Accelerating early academic oral English development in transitional bilingual and structured English immersion programs. *American Educational Research Journal, 45*(4), 1011–1044. https://doi.org/10.3102/2F0002831208320790.

Tong, X., Deacon, S. H., & Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: another piece of the puzzle. *Journal of Learning Disabilities, 47*(1), 22–33. https://doi.org/10.1177/2F0022219413509971.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology, 40*(1), 7–26. https://doi.org/10.1016/S0022-4405(01)00092-9.

*Vadasy, P. F. & Sanders, E. A. (2010). Efficacy of supplemental phonics-based instruction for low-skilled kindergarteners in the context of language minority status and classroom phonics instruction. *Journal of Educational Psychology, 102*, 786-803. https://doi.org/10.1037/a0019639, 4

*Vadasy, P. F. & Sanders, E. A. (2011). Efficacy of supplemental phonics-based instruction for low-skilled first graders: how language minority status and pretest characteristics moderate treatment response. *Scientific Studies of Reading, 15*, 471-497. https://doi.org/10.1080/10888438.2010.501091, 6

Vadasy, P., & Sanders, E. (2012). Two-year follow-up of a kindergarten phonics intervention for English learners and native English speakers: contextualizing treatment impacts by classroom literacy instruction. *Journal of Educational Psychology, 104*(4), 987–1005. https://doi.org/10.1037/a0028163.

Vadasy, P., Sanders, E., & Peyton, J. (2006). Code-oriented instruction for kindergarten students at risk for reading difficulties: a randomized field trial with paraeducator implementers. *Journal of Educational Psychology, 98*(3), 508–528. https://doi.org/10.1037/0022-0663.98.3.508.

Vadasy, P., Sanders, E., & Abbott, R. (2008). Effects of supplemental early reading intervention at 2-year follow up: reading skill growth patterns and predictors. *Scientific Studies of Reading, 12*(1), 51–89. https://doi.org/10.1080/10888430701746906.

*Vaughn, S., Cirino, P. T., Linan-Thompson, S., Mathes, P. G., Carlson, C. D., Cardenas-Hagen, E., Pollard-Durodola, S. D., Fletcher, J. M., & Francis, D. J. (2006a). Effectiveness of a Spanish intervention and an English intervention for English-language learners at risk for reading problems. *American Educational Research Journal, 43*, 449-487. https://doi.org/10.3102/2F00028312043003449, 3

*Vaughn, S., Mathes, P., Linan-Thompson, S., Cirino, P., Carlson, C., Pollard-Durodola, S., Cardenas-Hagen, E., & Francis, D. (2006b). Effectiveness of an English intervention for first-grade English language learners at risk for reading problems. *The Elementary School Journal, 107*, 153-180. https://doi.org/10.1086/510653, 2

Vaughn, S., Elbaum, B. E., Wanzek, J., Scammacca, N., & Walker, M. A. (2014). *Code sheet and guide for education-related intervention study syntheses.* The Meadows Center for Preventing Educational Risk.

Ventriglia, L., & González, L. (2000). *Santillana Intensive English.* Santillana USA.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48 http://www.jstatsoft.org/v36/i03/.

Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-analyses of the effects of tier 2 type reading interventions in grades K-3. *Educational Psychology Review, 28*(3), 551–576.

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current evidence on the effects of intensive early reading interventions. *Journal of Learning Disabilities, 51*(6), 612–624. https://doi.org/10.1177/2F0022219418775110.

 Springer

Affiliations

Garrett J. Roberts[1] **· Colby Hall**[2] **· Eunsoo Cho**[3] **·**
Brooke Coté[1] **· Jihyun Lee**[4] **· Bingxin Qi**[5] **· Jacklyn Van Ooyik**[1]

Colby Hall
csh4t@virginia.edu

Eunsoo Cho
escho@msu.edu

Brooke Coté
brooke.cote@du.edu

Jihyun Lee
jihyunl@utexas.edu

Bingxin Qi
bingxin.qi@du.edu

Jacklyn Van Ooyik
jacklyn.vanooyik@du.edu

[1] Department of Teaching and Learning Sciences, Morgridge College of Education, The University of Denver, 1999 E. Evans Ave, Denver, CO 80210, USA

[2] School of Education and Human Development, The University of Virginia, 417 Emmet Street South, P.O. Box 400273, Charlottesville, VA 22904, USA

[3] Department of Counseling, Educational Psychology, and Special Education, College of Education, Michigan State University, 620 Farm Lane, East Lansing, MI 48824, USA

[4] Department of Educational Psychology, The University of Texas at Austin, 1912 Speedway STE 504, Austin, TX 78712, USA

[5] Department of Research Methods and Information Science, Morgridge College of Education, The University of Denver, 1999 E. Evans Ave, Denver, CO 80210, USA