English learners remain the fastest-growing group of students in American schools, with large increases occurring in most regions of the country (U.S. Department of Education, 2011). For a variety of reasons, including the fact that these students have to learn a second or even third language while also mastering grade-level content, they form a significant portion of students who struggle academically. Only 7% of fourth-grade and 3% of eighth-grade English learners score at or above proficiency on reading assessments as compared to 38% and 37% of native English speakers (National Center for Educational Statistics, 2014). These data highlight the importance of instructional interventions to support their academic progress as well as their English language proficiency abilities. However, little is known about the additional support English learners receive in schools as part of a response-to-intervention model.

This article reviews published experimental studies from 2000 to 2012 that evaluated the effects of providing reading interventions to English learners who were at risk for experiencing academic difficulties, including students with learning disabilities. The interventions included explicit instruction, and 10 used published intervention programs. Moderator variables, such as group size, minutes of intervention, and type of personnel delivering the intervention, were not significant predictors of outcomes.

Abstract

Scott K. Baker, Doris L. Baker, Russell Gersten, Catherine Richards-Tutor, and Jeanie Mercier Smith

A Research Synthesis: Interventions for English Learners: The Effectiveness of Reading Interventions
Exceptional Children

studies investigating the instructional effec-
tiveness of interventions, results of these
studies have not been synthesized in a cohe-
sive manner for scientific purposes or to
guide practitioners in their implementation
of effective interventions for this population
of students. In this article, we describe and
synthesize this recent research. Our review
includes summarizing findings from previ-
ous reviews, manifestations of poor educa-
tion and intervention support for English
learners, and the disproportionate overrepre-
sentation and underrepresentation of English
learners in special education (August &
Hakuta, 1997; August & Shanahan, 2006;
Klingner et al., 2006). We also discuss the
implications for both future research and
practice.

Past Syntheses on
Interventions for English
Learners

In 2000, Gersten and Baker reviewed research
on effective instructional practices for English
learners. The authors found few experimental
studies and consequently expanded their synthe-
sis to include descriptive and qualitative studies
and also incorporated the professional opinions
of practitioners with specific expertise in work-
ing effectively with English learners (i.e., Noblit
& Hare, 1988; Ogawa & Malen, 1991). Studies
reviewed included case study research, qualita-
tive research, descriptive studies, and a small
number of experiments and quasi-experiments.
This analysis was not limited to interventions for
struggling English learners in reading but
included any research on K–8 instruction.

Results of the synthesis suggested several prom-
ising approaches for improving instruction for
English learners: (a) using vocabulary as a cur-
riculum anchor across multiple subject areas,
(b) using graphic organizers and other physical
artifacts to reinforce concept acquisition and
growth in academic vocabulary, (c) using coop-
erative and peer-tutoring strategies to enhance
engagement and nonthreatening articulation and
discussion of newly acquired content, (d) strate-
gically using the native language when neces-
sary, and (e) modulating cognitive and language
demands depending on the lesson objectives.

In a more recent review, Klingner et al.
(2006) located eight studies conducted since
1997 that met their criteria for credible
approaches toward reading instruction for Eng-
lish learners. The evidence base of this review
centered on qualitative studies rather than exper-
imental research. Only two of the eight studies
reviewed were experimental, and one of the two
was conducted in India. The authors articulated
what they viewed as promising practices based
on the studies reviewed. These promising prac-
tices included (a) combining phonological
awareness (PA) with other English language
development activities, (b) teaching and encour-
gaging the use of reading comprehension strate-
gies in the first and second language, (c) helping
students develop a strong foundation in reading
in both their native language and in English, and
(d) heavy emphasis on rich vocabulary instruc-
tion and native language support. In their review,
the authors noted that these practices provide
an important foundation for promoting success
in English language learners. However, in both
their native language and in English, and
in the process of developing a strong foundation in
reading

At the same time, guidance to practitioners
regarding how to effectively teach reading to
English learners remains a pressing national
priority. In response to this demand, the Insti-
tute of Education Sciences (IES) has published
two practice guides (S. Baker et al., 2014; Ger-
sten et al., 2007) intended to provide practice
guidance to those teaching English learners.
The guides are intended to serve as means for
informing practice. Both guides were con-
ducted using best practices from the litera-
ture. The IES practice guide contains a number of
evidence-based recommendations and instructional
strategies. Each guide includes a section on
language and literacy support for English
learners.

In summary, findings from both syntheses
converge in the importance of providing strong
vocabulary instruction and native language sup-
port. However, in neither the Gersten et al.
(2000) nor the Klingner et al. (2006) syntheses
was the evidence base sufficient to draw clear
conclusions about best practices. Both were
intended to serve as means for delineating prom-
ineering practices that could subsequently be eval-
uated with rigorous research.

At the same time, guidance to practitioners
regarding how to effectively teach reading to
English learners remains a pressing national
priority. In response to this demand, the Insti-
tute of Education Sciences (IES) has published
two practice guides (S. Baker et al., 2014; Ger-
sten et al., 2007) intended to provide practice
guidance to those teaching English learners.
The guides are intended to serve as means for
informing practice. Both guides were con-
ducted using best practices from the litera-
ture. The IES practice guide contains a number of
evidence-based recommendations and instructional
strategies. Each guide includes a section on
language and literacy support for English
learners.

In summary, findings from both syntheses
converge in the importance of providing strong
vocabulary instruction and native language sup-
port. However, in neither the Gersten et al.
(2000) nor the Klingner et al. (2006) syntheses
was the evidence base sufficient to draw clear
conclusions about best practices. Both were
intended to serve as means for delineating prom-
ineering practices that could subsequently be eval-
uated with rigorous research.

At the same time, guidance to practitioners
regarding how to effectively teach reading to
English learners remains a pressing national
priority. In response to this demand, the Insti-
tute of Education Sciences (IES) has published
two practice guides (S. Baker et al., 2014; Ger-
sten et al., 2007) intended to provide practice
guidance to those teaching English learners.
The guides are intended to serve as means for
informing practice. Both guides were con-
ducted using best practices from the litera-
ture. The IES practice guide contains a number of
evidence-based recommendations and instructional
strategies. Each guide includes a section on
language and literacy support for English
learners.

In summary, findings from both syntheses
converge in the importance of providing strong
vocabulary instruction and native language sup-
port. However, in neither the Gersten et al.
(2000) nor the Klingner et al. (2006) syntheses
was the evidence base sufficient to draw clear
conclusions about best practices. Both were
intended to serve as means for delineating prom-
ineering practices that could subsequently be eval-
uated with rigorous research.

At the same time, guidance to practitioners
regarding how to effectively teach reading to
English learners remains a pressing national
priority. In response to this demand, the Insti-
tute of Education Sciences (IES) has published
two practice guides (S. Baker et al., 2014; Ger-
sten et al., 2007) intended to provide practice
guidance to those teaching English learners.
The guides are intended to serve as means for
informing practice. Both guides were con-
ducted using best practices from the litera-
ture. The IES practice guide contains a number of
evidence-based recommendations and instructional
strategies. Each guide includes a section on
language and literacy support for English
learners.

In summary, findings from both syntheses
converge in the importance of providing strong
vocabulary instruction and native language sup-
port. However, in neither the Gersten et al.
(2000) nor the Klingner et al. (2006) syntheses
was the evidence base sufficient to draw clear
conclusions about best practices. Both were
intended to serve as means for delineating prom-
ineering practices that could subsequently be eval-
uated with rigorous research.

At the same time, guidance to practitioners
regarding how to effectively teach reading to
English learners remains a pressing national
priority. In response to this demand, the Insti-
tute of Education Sciences (IES) has published
two practice guides (S. Baker et al., 2014; Ger-
sten et al., 2007) intended to provide practice
guidance to those teaching English learners.
The guides are intended to serve as means for
informing practice. Both guides were con-
ducted using best practices from the litera-
ture. The IES practice guide contains a number of
evidence-based recommendations and instructional
strategies. Each guide includes a section on
language and literacy support for English
learners.

In summary, findings from both syntheses
converge in the importance of providing strong
vocabulary instruction and native language sup-
port. However, in neither the Gersten et al.
(2000) nor the Klingner et al. (2006) syntheses
was the evidence base sufficient to draw clear
conclusions about best practices. Both were
intended to serve as means for delineating prom-
ineering practices that could subsequently be eval-
uated with rigorous research.

At the same time, guidance to practitioners
regarding how to effectively teach reading to
English learners remains a pressing national
priority. In response to this demand, the Insti-
tute of Education Sciences (IES) has published
two practice guides (S. Baker et al., 2014; Ger-
sten et al., 2007) intended to provide practice
guidance to those teaching English learners.
The guides are intended to serve as means for
informing practice. Both guides were con-
ducted using best practices from the litera-
ture. The IES practice guide contains a number of
evidence-based recommendations and instructional
strategies. Each guide includes a section on
language and literacy support for English
learners.

In summary, findings from both syntheses
converge in the importance of providing strong
vocabulary instruction and native language sup-
port. However, in neither the Gersten et al.
(2000) nor the Klingner et al. (2006) syntheses
was the evidence base sufficient to draw clear
conclusions about best practices. Both were
intended to serve as means for delineating prom-
ineering practices that could subsequently be eval-
uated with rigorous research.

At the same time, guidance to practitioners
regarding how to effectively teach reading to
English learners remains a pressing national
priority. In response to this demand, the Insti-
tute of Education Sciences (IES) has published
two practice guides (S. Baker et al., 2014; Ger-
sten et al., 2007) intended to provide practice
guidance to those teaching English learners.
The guides are intended to serve as means for
informing practice. Both guides were con-
ducted using best practices from the litera-
ture. The IES practice guide contains a number of
evidence-based recommendations and instructional
strategies. Each guide includes a section on
language and literacy support for English
learners.
necessary scaffolds students need to understand the concepts taught. Small-group instruction provides English learners with extended opportunities to use English and multiple opportunities to interact closely with the teacher. Moreover, skilled teachers can take advantage of small-group instructional opportunities to not only provide targeted and modulated instruction to meet the precise needs of individual children but also provide additional opportunities for English learners to speak, hear, and read English. These additional opportunities may help English learners develop their English language proficiency, an important component of comprehension (D. Baker, Park, & Baker, 2013; Farnia & Geva, 2011; Gottardo & Mueller, 2009).

In this article, we extend the findings of previous reviews and the two IES practice guides by calculating the measurable impacts of interventions from published experimental studies that have been conducted since 2000 with English learners identified as at risk or with learning disabilities. We included only experimental studies (i.e., randomized control trials [RCTs]) because we were interested in determining if there were interventions that had been subjected to the peer-review process and were thus more likely to be effective. In addition, we included only studies that (a) reported on English learners identified as at risk or with learning disabilities, and (b) included English learners in kindergarten through 12th grade.

Method

We used the following criteria to identify RCT studies in peer-reviewed journals: (a) the study sample comprised English learners in kindergarten through 12th grade who were identified as at risk or with a learning disability (using either standardized tests or valid screening measures), (b) data were disaggregated by English learner status if not all participants were English learners, and (c) information about fidelity of implementation was reported. These criteria were used because they allowed for precise impact estimates to be calculated for English learners specifically and helped ensure that interventions were delivered as intended.

During January through March 2013, the search focused on studies from 2000 to 2012 in PsycINFO and ERIC, using the following keywords individually in peer-reviewed journals: English learners, language minority students, second language learners, intervention, response to intervention, at-risk, learning disabilities, reading difficulty, writing difficulty, math difficulty. We then specifically searched for studies in the following journals: Journal of Educational Psychology, Journal of Learning Disability, Learning Disabilities Research & Practice, Learning Disabilities Quarterly, Reading Research Quarterly, Remedial and Special Education, Scientific Studies of Reading, The Journal of Special Education, Exceptional Children, Journal of Literacy Research, and Topics in Language Disorders.

Once we identified potential studies, we read the abstracts and selected for further analysis only those that met the following criteria: (a) English learners were the focus of the study, (b) the study was conducted in a school setting, (c) the study was conducted with a group of English learners, and (d) the study included a control group. We also excluded studies that were conducted outside of the United States, studies that were not published in English, and studies that were not peer-reviewed.

Finally, we calculated the effect sizes and summarized these effects in terms of meaningful categorizations.
Exceptional Children

those studies that indicated they used an RCT design and included struggling K–12 English learners as participants (i.e., English learners who were receiving a Tier 2 or Tier 3 intervention). Next, we carefully reviewed the Method section of each article to ensure that the studies met our other criteria. We located three studies that met these criteria but used single-case design methods. We do not reference these studies in this review, but we published a separate technical report on these studies (Richards-Tutor, Baker, Gersten, Baker, & Smith, 2014).

Once we identified eligible studies, we used a coding form to summarize the information from each study by two broad categories: (a) features of the research study (i.e., research design, grade level, participant characteristics, setting) and (b) characteristics of the interventions (i.e., group size, duration, personnel delivering the intervention, intervention content, intervention methods). Table 1 presents the features of the research studies, and Table 2 presents the characteristics of the interventions. Two raters independently coded each study; all addressed reading or pre-reading. The comprehensive literature search yielded 12 studies; all addressed reading or pre-reading.

The comprehensive literature search yielded 12 studies; all addressed reading or pre-reading.

Once we identified eligible studies, we used a coding form to summarize the information from each study by two broad categories: (a) features of the research study (i.e., research design, grade level, participant characteristics, setting) and (b) characteristics of the interventions (i.e., group size, duration, personnel delivering the intervention, intervention content, intervention methods). Table 1 presents the features of the research studies, and Table 2 presents the characteristics of the interventions. Two raters independently coded each study; all addressed reading or pre-reading.

Calculation of Effect Size

To calculate the effect size, we used an adjusted mean difference in the numerator and the pooled unadjusted standard deviation in the denominator. If adjusted means were not reported, we calculated effect sizes using a difference-in-differences approach (i.e., we computed a gain score for both experimental and control groups and then calculated the difference in gain scores). For the calculation of effect sizes, we used an adjusted mean difference in the numerator and the pooled unadjusted standard deviation in the denominator. If adjusted means were not reported, we calculated effect sizes using a difference-in-differences approach (i.e., we computed a gain score for both experimental and control groups and then calculated the difference in gain scores).

Results

The comprehensive literature search yielded 12 studies; all addressed reading or pre-reading. The comprehensive literature search yielded 12 studies; all addressed reading or pre-reading.

Calculation of Effect Size

To calculate the effect size, we used an adjusted mean difference in the numerator and the pooled unadjusted standard deviation in the denominator. If adjusted means were not reported, we calculated effect sizes using a difference-in-differences approach (i.e., we computed a gain score for both experimental and control groups and then calculated the difference in gain scores).
<table>
<thead>
<tr>
<th>Authors</th>
<th>Grade level</th>
<th>Sample size</th>
<th>EL primary language</th>
<th>EL determination</th>
<th>Reading risk determination</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Connor, Bocian, Beebe-Frankenberger, &amp; Linklater (2010)</td>
<td>K</td>
<td>35</td>
<td>Spanish</td>
<td>CELDT</td>
<td>DIBELS (LNF &lt; 9 and ISF &lt; 7) PPVT (SS &lt; 85)</td>
</tr>
<tr>
<td>Vadasy &amp; Sanders (2010)</td>
<td>K</td>
<td>84</td>
<td>28 different languages, including Spanish = 49%, Vietnamese = 15%, Somali = 6%, Chinese = 6%, Tagalog = 3%</td>
<td>Parent survey</td>
<td>In bottom half of class on LN, LS (Fuchs et al., 2001), and CTOPP Sound Matching (composite-score z score)</td>
</tr>
<tr>
<td>Gunn, Biglan, Smolkowski, &amp; Ary (2000)</td>
<td>K–3</td>
<td>122</td>
<td>Spanish</td>
<td>Not reported</td>
<td>DIBELS composite: both some risk and high risk categories</td>
</tr>
<tr>
<td>Vaughn, Linan-Thompson, et al. (2006)</td>
<td>I</td>
<td>69</td>
<td>Spanish</td>
<td></td>
<td>School determination WLPB LWID (below 25th percentile) Experimenter word reading list (raw score &lt; 2)</td>
</tr>
<tr>
<td>Vaughn, Mathes, et al. (2006)</td>
<td>I</td>
<td>48</td>
<td>Spanish</td>
<td></td>
<td>School determination WLPB LWID (below 25th percentile) Experimenter word reading list (raw score &lt; 2)</td>
</tr>
<tr>
<td>Vaughn, Cirino, et al. (2006)</td>
<td>I</td>
<td>190 total</td>
<td>Spanish</td>
<td>School determination WLPB LWID (below 25th percentile) Researcher developed word reading list (raw score &lt; 2)</td>
<td></td>
</tr>
<tr>
<td>Begeny, Ross, Greene, Mitchell, &amp; Whitehouse (2012)</td>
<td>2</td>
<td>21</td>
<td>Spanish</td>
<td>School determination GORT fluency or comprehension &lt; 10 (SS mean)</td>
<td></td>
</tr>
<tr>
<td>Denton, Anthony, Parker, &amp; Hasrouck (2004)</td>
<td>2–5</td>
<td>93 total (n = 22, Grade 2; n = 37, Grade 3; n = 28, Grade 4; n = 6, Grade 5)</td>
<td>Spanish</td>
<td>LAS</td>
<td>Teacher recommendation WRMT: LWID and WA subtests &lt; Grade 1 equivalent = emerging decoding group WRMT: LWID and WA subtests &gt; Grade 1 equivalent = established decoding group</td>
</tr>
<tr>
<td>Authors</td>
<td>Grade level</td>
<td>Sample size</td>
<td>EL primary language</td>
<td>EL determination</td>
<td>Reading risk determination</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------------------------------------</td>
<td>--------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Lovett et al. (2008)</td>
<td>2–8</td>
<td>76</td>
<td>Multiple languages, including</td>
<td>Parent survey</td>
<td>Teacher referral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Portuguese = 49% Spanish = 21%</td>
<td></td>
<td>WRMT: LWID and WA (SS &lt; 85)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WRAT-3 (SS &lt; 85)</td>
</tr>
<tr>
<td>Wanzek &amp; Roberts (2012)</td>
<td>4</td>
<td>74</td>
<td>Spanish</td>
<td>School determination</td>
<td>Teacher referral</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GMRT (SS &lt; 25th percentile)</td>
</tr>
<tr>
<td>Vaughn et al. (2011)</td>
<td>7–8</td>
<td>42</td>
<td>Spanish</td>
<td>Not reported</td>
<td>TAKS (SS &lt; 30th percentile)</td>
</tr>
</tbody>
</table>

Note. EL = English learner; CELDT = California English Language Development Test; DIBELS = Dynamic Indicators of Basic Early Literacy Skills, 6th ed. (Good & Kaminski, 2002); LNF = Letter Naming Fluency; ISF = Initial Sound Fluency; PPVT = Peabody Picture Vocabulary Test (Dunn & Dunn, 1981); PA = phonological awareness; LN = letter naming; LS = letter sounds; CTOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999); WLPB = Woodcock Language Proficiency Battery (Woodcock, 1991); LWID = Letter Word Identification; WA = Word Attack; GORT = Gray Oral Reading Test (Bryant, Shih, & Bryant, 2009); WRMT = Woodcock Reading Mastery Test (Woodcock, 1987); LAS = Language Assessment Scales (De Avila & Duncan, 1990); WRAT = Wide Range Achievement Test (Wilkinson, 1993); GMRT = Gates-MacGinitie Reading Test (MacGinitie, MacGinitie, Dreyer, & Hughes, 2006); TAKS = Texas Assessment of Knowledge and Skills; SS = standard score.
conducted in kindergarten or first grade. One study involved second-grade students (Begeny, Ross, Greene, Mitchell, & Whitehouse, 2012) and one study involved fourth graders (Wanzek & Roberts, 2012). Two studies involved students in the upper elementary grades (Denton, Anthony, Parker, & Hasrouck, 2004; Gunn, Biglan, Smolkowski, & Ary, 2000), and two studies involved students at the middle school level (Lovett et al., 2008; Vaughn et al., 2011).

Half the studies included English learners only (Begeny et al., 2012; Denton et al., 2004; Solari & Gerber, 2008; Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006; Wanzek & Roberts, 2012). The remaining six included both English learners and native English speakers, but we were able to separately analyze the English learner subsample. In all but two of the studies, participants were from homes where Spanish was the primary home language. In the two other studies (Vahey & Sanders, 2010; Lovett et al., 2008) the English learner participant sample included numerous languages.

Methods used to determine English learner status. The majority of studies (n = 7) used as their criteria the school designation of English learners (Begeny et al., 2012; Denton et al., 2004; O'Connor, Bocian, Beebe-Frankenberger, & Linklater, 2010; Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Wanzek, Roberts, et al., 2006). Three studies used parent questionnaires or interviews to determine the primary language spoken at home (Lovett et al., 2008; Solari & Gerber, 2008; Vadasy & Sanders, 2010). Two studies did not report specifically how English learner status was determined (Vadasy & Sanders, 2010). Three studies used screening measures of PA, alphabet knowledge, or word and pseudoword reading. In contrast, both Solari and Gerber (2008) and O'Connor et al. (2010) used the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981) to determine risk status. The three first-grade studies by Vaughn et al. (Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006) and the Solari and Gerber study assessed students both in their native language (i.e., Spanish) and in English as part of the risk determination process. Half the studies included both English learners and native English speakers, but we were able to separately analyze the English learner subsample in all but two of the studies. Participants were from homes where Spanish was the primary home language. In the two other studies (Vahey & Sanders, 2010; Lovett et al., 2008) the English learner participant sample included numerous languages.

Methods used to identify risk status and learning disability. Three studies included English learners with identified learning disabilities (Lovett et al., 2008; Vaughn et al., 2011; Wanzek & Roberts, 2012) in their sample. The remaining studies included students deemed at risk for learning disabilities, but definitions of how risk was determined varied. The three kindergarten and first-grade studies included students with identified learning disabilities (Lovett et al., 2008; Vaughn et al., 2011; Wanzek & Roberts, 2012) in their sample. The remaining studies included students deemed at risk for learning disabilities, but definitions of how risk was determined varied. The remaining studies included students deemed at risk for learning disabilities, but definitions of how risk was determined varied.

Duration of the intervention. Intervention sessions ranged from 20 to 60 min. The shortest intervention sessions were 20 min, and the longest were 60 min. The majority of studies reported intervention sessions of 30 to 45 min.

Intervention Characteristics. The intensity of an intervention can be determined by many characteristics, but three are common: group size, duration of the intervention, and quality of the personnel delivering the intervention and the associated amount of training they receive. The intervention sessions ranged from 20 to 60 min. The shortest sessions were 20 min, and the longest were 60 min. The majority of studies reported intervention sessions of 30 to 45 min.

Group size. Two studies (Begeny et al., 2012; Vadasy & Sanders, 2010) used one-on-one tutoring. The remaining studies included groups of three to five students. Half the studies included groups of three to five students, and one study included groups as large as eight students.

Duration of the intervention. Intervention sessions ranged from 20 to 60 min. The shortest sessions were 20 min, and the longest were 60 min. The majority of studies reported intervention sessions of 30 to 45 min.

Intervention Characteristics. The intensity of an intervention can be determined by many characteristics, but three are common: group size, duration of the intervention, and quality of the personnel delivering the intervention and the associated amount of training they receive. The intervention sessions ranged from 20 to 60 min. The shortest sessions were 20 min, and the longest were 60 min. The majority of studies reported intervention sessions of 30 to 45 min.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Group size and composition</th>
<th>Intervention duration</th>
<th>Interventionist</th>
<th>Intervention program</th>
<th>Intervention content</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Connor, Bocian, Beebe-Frankenberger, &amp; Linklater (2010)</td>
<td>2–3; homogeneous groups</td>
<td>36 weeks, 3 days per week, 15 min per day (270–1,430 min)</td>
<td>Paraeducators</td>
<td>Ladders to Literacy</td>
<td>Alphabetics, phonological awareness, oral language</td>
</tr>
<tr>
<td>Solari &amp; Gerber (2008)</td>
<td>4–5; homogeneous groups</td>
<td>8 weeks, 3 days per week, 20 min per day (480 min)</td>
<td>Research assistants</td>
<td>NA</td>
<td>Phonological awareness, listening comprehension</td>
</tr>
<tr>
<td>Vadasy &amp; Sanders (2010)</td>
<td>One-on-one instruction</td>
<td>18 weeks, 4 days per week, 30 min per day (2,160 min)</td>
<td>Paraeducators</td>
<td>NA</td>
<td>Alphabetics, phonological awareness, word reading, spelling, oral language</td>
</tr>
<tr>
<td>Gunn, Biglan, Smolkowski, &amp; Ary (2000)</td>
<td>2–3; homogeneous groups</td>
<td>60 weeks, 5 days per week, 25–30 min per day (7,500–9,000 min)</td>
<td>Paraeducators</td>
<td>Reading Mastery and Corrective Reading</td>
<td>Phonological awareness, letter knowledge, word recognition, fluency, comprehension, oral language skill, vocabulary</td>
</tr>
<tr>
<td>Vaughn, Linan-Thompson, et al. (2006)</td>
<td>3–5; homogeneous groups</td>
<td>32 weeks, 5 days per week, 50 min per day: (4,560–6,900 min)</td>
<td>Bilingual certified teachers hired by research team</td>
<td>Lectura Proactiva</td>
<td>Letter knowledge, word recognition, fluency, comprehension, oracy, vocabulary</td>
</tr>
<tr>
<td>Vaughn, Mathes, et al. (2006)</td>
<td>3–5; homogeneous groups</td>
<td>32 weeks, 5 days per week, 40 min per day (6,400 min)</td>
<td>Bilingual certified teachers hired by research team</td>
<td>Proactive Reading</td>
<td>Letter knowledge, word recognition, fluency, comprehension, oracy, vocabulary</td>
</tr>
<tr>
<td>Vaughn, Cirino, et al. (2006)</td>
<td>3–5; homogeneous groups</td>
<td>32 weeks, 5 days per week, 50 min per day (4,476–6,402 min)</td>
<td>Bilingual certified teachers hired by research team</td>
<td>Lectura Proactiva or Proactive Reading</td>
<td>Letter knowledge, word recognition, fluency, comprehension, oracy, vocabulary</td>
</tr>
<tr>
<td>Begeny, Ross, Greene, Mitchell, &amp; Whitehouse (2012)</td>
<td>One-on-one instruction</td>
<td>20–28 weeks, 2–3 times per week, 10 min per day (600–840 min)</td>
<td>Lead researcher, graduate and undergraduate students</td>
<td>HELPS Fluency Program</td>
<td>Reading fluency, comprehension</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Authors</th>
<th>Group size and composition</th>
<th>Intervention duration</th>
<th>Interventionist</th>
<th>Intervention program</th>
<th>Intervention content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denton, Anthony, Parker, &amp; Hasrouck (2004)</td>
<td>1–4; homogeneous based on decoding</td>
<td>10 weeks, 3 days per week, 40 min per day (1,200 min)</td>
<td>Undergraduate students</td>
<td>Read Well or Read Naturally</td>
<td>Alphabetsics, reading fluency, vocabulary, comprehension</td>
</tr>
<tr>
<td>Lovett et al. (2008)</td>
<td>4–8; homogeneous groups based on decoding</td>
<td>21 weeks*, 4–5 days per week, 60 min per day (6,300 min)</td>
<td>Certified special education teachers</td>
<td>Reading Mastery or Corrective Reading</td>
<td>Alphabetsics, word reading, phonological awareness</td>
</tr>
<tr>
<td>Wanzek &amp; Roberts (2012)</td>
<td>2–4; within schools</td>
<td>28 weeks, 5 days per week, 30 min per day (2,550–3,420 min)</td>
<td>Certified teachers hired by research team</td>
<td>Wilson Reading</td>
<td>Word reading, comprehension</td>
</tr>
<tr>
<td>Vaughn et al. (2011)</td>
<td>4–5 (not reported)</td>
<td>32 weeks, 5 days per week, 50 min per day (8,000 min)</td>
<td>Certified teachers hired by research team</td>
<td>REWARDS and Wilson Reading</td>
<td>Reading fluency, vocabulary, comprehension</td>
</tr>
</tbody>
</table>

Note. The instructional program and content in the control condition was whatever the district and school typically provided to students. Homogenous refers to academic levels and not English proficiency level or English learner status.

*Not reported, calculated by authors.
Exceptional Children

sessions were the kindergarten studies with 10- to 20-min sessions, whereas most other sessions ranged from 30 to 60 min. Six studies included daily intervention sessions; the remainder varied from twice a week to four times a week. Length of intervention also varied substantially. These factors yielded a large range in the total number of minutes of intervention provided, with the range being from 270 min to 9,000 min. The average intervention was approximately 3,600 min in duration, which equals to about 120 thirty-minute lessons.

Personnel delivering instruction. Teachers delivered the instruction in six of the interventions. In three studies, teachers with specific training in bilingual education delivered the interventions (Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; and Vaughn, Mathes, et al., 2006), and in another, special education teachers delivered the intervention (Lovett et al., 2008). Two studies hired outside teachers to deliver the interventions (Vaughn et al., 2011; Wanzek & Roberts, 2012). Three studies employed paraprofessionals (Gunn et al., 2000; O'Connor et al., 2010; Vadasy & Sanders, 2010), and in three other studies, research assistants, undergraduates, or graduate students delivered the intervention (Begeny et al., 2012; Denton et al., 2004; Solari & Gerber, 2008). In all cases, teachers, paraprofessionals, or research personnel were trained on how to deliver the intervention; they were observed regularly; and they were provided with feedback.

Content of the interventions. Half the studies (Denton et al., 2004; Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006; Vaughn et al., 2011; Wanzek & Roberts, 2012) used a comprehensive intervention that covered at least four of the five areas of literacy outlined in the National Reading Panel (Ehri et al., 2001) and the National Literacy Panel for Language Minority Students (August & Shanahan, 2006): phonemic awareness, phonics, fluency, vocabulary, and comprehension. The other half (Begeny et al., 2012; Gunn et al., 2000; Lovett et al., 2008; O'Connor et al., 2010; Solari & Gerber, 2008; Vadasy & Sanders, 2010) focused on just one or two components of reading. In general, studies that targeted the kindergarten level (e.g., O'Connor et al., 2010) focused on PA and alphabetic knowledge, and studies that targeted the intermediate grades (e.g., Begeny et al., 2012) focused on fluency and comprehension.

Five of the studies included vocabulary as one of the proficiencies (Denton et al., 2004; Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006; Vaughn et al., 2011), and four studies focused on oral language development as a key skill targeted in the intervention. None of the studies included building of academic vocabulary or academic language as an explicit goal, although we suspect that those with an oral language or vocabulary component probably did address these topics to some extent.

Ten studies used existing curricula in the treatment condition, including Reading Mastery (Engelmann & Bruner, 1995), Corrective Reading (Engelmann, 1988), Read Well (Sprick, Howard, & Fidanque, 1998), Read Well (Vaughn, Thompson, & Tomato, 2002), Wilson Reading System (Wyatt-Floyd, 2002), ORF (Begeny, 2009), T.L.A.P. (Lamberson & Loomis, 2003), LEAP (Orth & Scruggs, 2001), Cycles (Brown, 1991), Ready Write (Vaughn, Thompson, & Tomato, 2002), and several other locally developed instructional materials. Two studies combined existing and researcher-developed materials, and the decision about which materials to use was based on the skills taught (Vaughn et al., 2011; Wanzek & Roberts, 2012). Two studies developed and tested novel interventions in the treatment condition (Solari & Gerber, 2008; Vadasy & Sanders, 2010). Note that most of these curricula were developed for use with the general population rather than as specialized curricula for English learners.

In five of the studies, multiple interventions were tested in different treatment groups.
Richards-Tutor et al. (Denton et al., 2004; O'Connor et al., 2010; Solari & Gerber, 2008; Vaughn et al., 2011; Wanzek & Roberts, 2012). Only two studies, as indicated in Table 2, included interventions provided in the students’ primary languages in both cases. (Vaughn, Cirino, et al., 2006; Vaughn, Mathes, et al., 2006).

Methods of intervention delivery. The use of systematic, explicit instruction is the best way to describe the treatment intervention across all studies. Vaughn et al. (2011) described the treatment intervention across all studies. Common instructional procedures included modeling, scaffolding, and corrective feedback. Three studies described features of intervention delivery that were designed specifically to meet the needs of English learners learning to read in a relatively new language. The counterfactual: Nature of comparison group intervention and instruction. Of the 12 studies, 1 found the control condition as “business as usual” or “typical practice.” The majority of studies (n = 10) provided some information about the control condition; two studies did not (Gunn et al., 2000; Vaughn et al., 2011). Three studies reported that the control group received the school’s core reading program; Tier 1 instruction (Begeny et al., 2012; O’Connor et al., 2010; Vadasy & Sanders, 2010). Four studies reported that control students received supplemental intervention from their school in addition to Tier 1 (Denton et al., 2004; Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006). In the Denton study, seven of the students in the control group received on average 2,472 min of intervention, and 28 students in the English intervention received 5,256 min of intervention.

The counterfactual: Nature of comparison group intervention and instruction. Of the 12 studies, 11 described the control condition; two studies did not (Gunn et al., 2006; Vaughn, Mathes, et al., 2006). These studies described features of intervention delivery that were designed specifically to meet the needs of English learners learning to read in a relatively new language. Vaughn et al. (2004), Vaughn, Cirino, et al. (2006), Vaughn, Mathes, et al. (2006), and Roberts (2012) reported that eight of the control students received one supplemental intervention per week. These interventions typically took place in groups of two to three students and were delivered in the same way as the two intervention conditions.

Intervention Outcomes

Table 3 presents the outcomes for English learners in each study, summarized across seven domains. Although the National Reading Panel (Ehri et al., 2001) suggested only five domains in reading, we decided to create three separate domains for comprehension because of research suggesting that comprehension effects depend on how this component is measured (Cutting & Scarborough, 2006) and because for English learners, in particular, we thought the additional precision could be helpful in understanding intervention impact. We divided comprehension outcomes in the 12 studies into three categories based on prior knowledge, clarifying meanings of words, and showing differences between English and the students’ primary language.
Table 3. Outcome Effect Sizes by Measurement Domain.

<table>
<thead>
<tr>
<th>Study</th>
<th>PA</th>
<th>Phonics/word reading</th>
<th>Passage reading fluency</th>
<th>Vocabulary/oral language</th>
<th>Reading cloze</th>
<th>Reading comprehension</th>
<th>LC</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Connor, Bocian, Beebe-Frankenberger, &amp; Linklater (2010)</td>
<td>DIBELS PSF ($g = .91^{**}$)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Solari &amp; Gerber (2008)</td>
<td>LC concentration: Early PA ($g = .63$)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>LC concentration: Late PA ($g = -.14$)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>PA concentration: Early PA ($g = -.57$)</td>
<td>LC concentration: WJ III LWID ($g = -.19$)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>PA concentration: Late PA ($g = -.74$)</td>
<td>WJ III WA ($g = .43$)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Vadasy &amp; Sanders (2010)b</td>
<td>CTOPP ($g = .93^{**}$)</td>
<td>WRMT LWID and WA ($g = .61^{**}$)</td>
<td>Passage reading fluency ($g = .90^{**}$)</td>
<td>—</td>
<td>—</td>
<td>WRMT Passage Comprehension ($g = .47^{**}$)</td>
<td>—</td>
</tr>
<tr>
<td>Gunn, Biglan, Smolkowski, &amp; Ary (2000)</td>
<td>—</td>
<td>WJ III LWID ($g = .24$)</td>
<td>DIBELS ORF ($g = .24$)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Vaughn, Linan-Thompson, et al. (2006) Spanish intervention</td>
<td>TOPPS ($g = .58^{*}$)</td>
<td>WLPB-Spanish WA ($g = .91^{**}$)</td>
<td>IDEL ORF ($g = .78^{*}$)</td>
<td>WLPB-Spanish Picture Vocabulary ($g = .28$)</td>
<td>WLPB-Spanish Passage Comprehension ($g = .88^{**}$)</td>
<td>—</td>
<td>WLPB-Spanish LC ($g = .50^{*}$)</td>
</tr>
<tr>
<td>Study</td>
<td>PA</td>
<td>Phonics/word reading</td>
<td>Passage reading fluency</td>
<td>Vocabulary/oral language</td>
<td>Reading cloze</td>
<td>Reading comprehension</td>
<td>LC</td>
</tr>
<tr>
<td>-------</td>
<td>----</td>
<td>----------------------</td>
<td>------------------------</td>
<td>-------------------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>----</td>
</tr>
<tr>
<td>Vaughn, Mathes, et al., (2006)(^b) English intervention</td>
<td>CTOPP ((g = 1.24^{**}))</td>
<td>WLPB-English WA ((g = .69))</td>
<td>DIBELS ORF ((g = .18))</td>
<td>WLPB-English Picture Vocabulary ((g = .09))</td>
<td>—</td>
<td>WLPB-English Listening Comprehension ((g = .26))</td>
<td></td>
</tr>
<tr>
<td>Vaughn, Cirino, et al. (2006) Spanish and English intervention</td>
<td>Spanish intervention: TOPPS ((g = .82^{**})) English intervention: CTOPP ((g = .38))</td>
<td>Spanish intervention: WLPB Spanish LWID ((g = .60^{**}))</td>
<td>Spanish intervention: IDEL ORF 1 ((g = .41))</td>
<td>Spanish intervention: WLPB-Spanish Picture Vocabulary ((g = .42))</td>
<td>—</td>
<td>WLPB-Spanish Listening Comprehension ((g = .23))</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>WLPB Spanish WA ((g = .45))</td>
<td>WLPB Spanish LWID ((g = .48^{*}))</td>
<td>WLPB-Spanish Picture Vocabulary ((g = .28))</td>
<td>WLPB-Spanish Picture Vocabulary ((g = -.14))</td>
<td>English intervention: WLPB-Spanish Listening Comprehension ((g = -.22))</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spanish word reading fluency ((g = .48^{*}))</td>
<td>Spanish intervention: DIBELS ORF 1 ((g = -.39))</td>
<td>WLPB-Spanish Verbal Analogies ((g = .33))</td>
<td>WLPB-Spanish Verbal Analogies ((g = .06))</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Domain average ((g = .51^{*}))</td>
<td>Domain average ((g = .27))</td>
<td>Domain average ((g = .33))</td>
<td>Domain average ((g = .06))</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>English intervention: WLPB-English LWID ((g = .13))</td>
<td>English intervention: WLPB-English WA ((g = .15))</td>
<td>English intervention: WLPB-English Listening Comprehension ((g = -.17))</td>
<td>English intervention: WLPB-English Listening Comprehension ((g = -.11))</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TOWRE ((g = .41))</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study PA</td>
<td>Phonics/word reading</td>
<td>Passage reading</td>
<td>fluency</td>
<td>Vocabulary/oral language</td>
<td>Reading comprehension</td>
<td>LC</td>
<td>Phonics/word study</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------</td>
<td>----------------</td>
<td>---------</td>
<td>--------------------------</td>
<td>----------------------</td>
<td>----</td>
<td>-------------------</td>
</tr>
<tr>
<td>Begeny, Ross, Greene, Mitchell, &amp; Whitehouse (2012)</td>
<td>— — GORT fluency ($g = .95$)</td>
<td>— — GORT comprehension ($g = 1.00^{**}$)</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denton, Anthony, Parker, &amp; Hasrouck (2004)</td>
<td>— — Read Well: WRMT Passage comprehension ($g = .18$)</td>
<td>— — Modified Read Naturally: WRMT Passage comprehension ($g = .15$)</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lovett et al. (2008)</td>
<td>— — Blending ($g = .59^*$)</td>
<td>— — WRMT Reading comprehension ($g = .33$)</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wanzek &amp; Roberts (2012)</td>
<td>— — Word study intervention: WJIII Listening comprehension ($g = .41$)</td>
<td>— — Comprehension intervention: WJIII Listening comprehension ($g = –.42$)</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. (continued)
Table 3. (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>PA</th>
<th>Phonics/word reading</th>
<th>Passage reading fluency</th>
<th>Vocabulary/oral language</th>
<th>Reading cloze</th>
<th>Reading comprehension</th>
<th>LC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>WJII LWID (g = .18)</td>
<td>—</td>
<td>WJ III Passage Comprehension (g = .26)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WJIII WA (g = -.02)</td>
<td>—</td>
<td>WJ III Passage Comprehension (g = -.05)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Standardized intervention:</td>
<td>—</td>
<td>Standardized intervention:</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WJII LWID (g = .23)</td>
<td>—</td>
<td>WJ III Passage Comprehension (g = .26)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WJIII WA (g = -.01)</td>
<td>—</td>
<td>WJ III Passage Comprehension (g = -.05)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. DIBELS = Dynamic Indicators of Basic Early Literacy Skills; PSF = Phoneme Segmentation Fluency; ORF = Oral Reading Fluency; RD = researcher developed; PA = phonological awareness; LC = listening comprehension; WJ III = Woodcock Johnson (3rd ed.; Woodcock, McGrew, & Mather, 2001); LWID = Letter Word Identification; WA = Word Attack; CTOPP = Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999); WRMT = Woodcock Reading Mastery Test (Woodcock, 1987); TOPPS = Test of Phonological Processing in Spanish (Francis et al., 2001); WLPB = Woodcock Language Proficiency Battery (Woodcock, 1991); IDEL = Indicadores Dinámicos del Éxito en la Lectura (Good, Bank, & Watson, 2003); GMRT = Gates-MacGinitie Reading Test (MacGinitie, McGinitie, Dreyer, & Hughes, 2006); GORT = Gray Oral Reading Test (Bryant, Shih, & Bryant, 2009); TOWRE = Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999).

*Research-developed measure.

bStudies had high attrition; either overall attrition or differential attrition effect sizes should be interpreted with caution.

cThe effect size is a mean composite of the two phonics decoding measures.

*p ≤ .05, **p ≤ .01.
the measures used: (a) reading cloze passage performance (e.g., Woodcock Reading Mastery Test, 1987), (b) reading passages with multiple-choice questions, and (c) listening comprehension. Outcomes for measures in English and, if included in the study, primary language measures are reported. For each domain, we report both standardized and researcher-developed measures that were administered in the studies. Researcher-developed measures that included multiple components of PA (e.g., word recognition, fluency, and word attack) were denoted in Table 3.

For each domain, we report the effect size range as well as the median effect size. We purposely do not report the mean effect size because of the variation in both the features of the studies and the characteristics of the interventions. As a representative effect size for the domain, we believe the median better preserves these variations than the mean, which by definition integrates this variation in the single score estimation process.

O'Connor et al. (2010) found significant effects for PA ($g = 0.91$), which was the main skill targeted in the intervention for this study. Vadasy and Sanders (2010) also found significant effects for PA ($g = 0.93$) for an intervention that included multiple components of PA measures. However, the treatment condition in this study focused on listening comprehension and the control group received a PA-only intervention, so the lack of effect on PA for the experimental group is not surprising. In the Vaughn studies (Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006), effect sizes were significant for students who received the Spanish intervention, but results were mixed for the English intervention as shown in Table 3.

PA. Seven studies measured PA (Lovett et al., 2008; O'Connor et al., 2010; Solari & Gerber, 2008; Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006; Vaughn, & Sanders, 2010). Not surprisingly, all of these studies had primary language measures in kindergarten or first grade. Effect sizes ranged from 0.74 to 1.24 with a median of 0.59. Significant effect sizes ranged from 0.86 to 1.24 with a median of 0.98. Significant effect sizes ranged from 0.74 to 1.24 with a median of 0.59. Significant effect sizes ranged from 0.86 to 1.24 with a median of 0.98. Significant effect sizes ranged from 0.74 to 1.24 with a median of 0.59. Significant effect sizes ranged from 0.86 to 1.24 with a median of 0.98. Significant effect sizes ranged from 0.74 to 1.24 with a median of 0.59. Significant effect sizes ranged from 0.86 to 1.24 with a median of 0.98.
Richards-Tutor et al.

Sanders, 2010; Vaughn et al., 2011; Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006). In each of these studies, fluency was measured using a 1-min timed passage and was scored as number of words read correctly, except the study by Begeny et al. (2012), which used the Gray Oral Reading Test (GORT; Bryant, Shih, & Bryant, 2009) fluency subtest to measure fluency. On the GORT fluency subtest, the amount of time it takes to read a passage is recorded and a rate computed. On the fluency measures, effect sizes ranged from –0.39 to 0.95 with a median of 0.28. Significant effect sizes were found in only two of the studies. In the Vadasy and Sanders (2010) study, the effect size was 0.90 for kindergarten students, and in the Vaughn, Linan-Thompson, et al. (2006) study, the effect size was 0.78 on the Spanish measure of reading fluency. In general, we did not find significant effects in passage reading fluency for English learners at risk or with learning disabilities who were taught in English.

Vocabulary and oral language.

Vocabulary and oral language were measured as an outcome in only four studies (Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006; Wanzek & Roberts, 2012). Effect sizes ranged from –0.59 to 0.78 with a median of –0.05. In only one study (Vaughn, Mathes, et al., 2006) was the effect size statistically significant. In three Vaughn et al. studies (Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; and Vaughn, Mathes, et al., 2006), vocabulary and oral language were measured using the Picture Vocabulary and Verbal Analogies subtests of the Woodcock Language Proficiency Battery needs citation. Across the three studies and these two measures, the only significant effect was for English Verbal Analogies for first-grade English learners who received the intervention in English (g = 0.78); however, the domain average for vocabulary in this study was not significant. For students in fourth grade, across three types of interventions (word study focused, comprehension focused, and responsive based on individual needs), no significant differences were found on vocabulary (Wanzek & Roberts, 2012).

Reading comprehension.

Reading comprehension measures were used in eight of the studies (Denton et al., 2004; Lovett et al., 2008; Vadasy & Sanders, 2010; Vaughn et al., 2011; Vaughn, Cirino, et al., 2006; Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006; Wanzek & Roberts, 2012). Typically, this skill was measured using the Passage Comprehension subtest from one of the Woodcock batteries. Effect sizes ranged from –0.21 to 0.88 with a median of 0.22. Significant effect sizes ranged from 0.47 to 0.88 with a median of 0.83. In two of the Vaughn studies (Vaughn, Linan-Thompson, et al., 2006; Vaughn, Mathes, et al., 2006) significant effects were found for the reading comprehension measure for first-grade students. For the Spanish intervention, Spanish passage comprehension was significant at 0.88, and for the English intervention, English passage comprehension was significant at 0.22. Significant effect sizes were found for the Woodcock Passage Comprehension: Spanish form (g = 0.78) on the Spanish measure of reading comprehension, Spanish form (g = 0.22) on the English measure of reading comprehension, Spanish form (g = 0.78) on the Spanish measure of reading comprehension, and English form (g = 0.22) on the English measure of reading comprehension.

Listening comprehension.

Listening comprehension was measured in two studies (Wanzek & Roberts, 2012). In the Vadasy and Sanders (2010) study, effect sizes were significant on listening comprehension for kindergarten students. In the Wanzek and Roberts (2012) study, effect sizes were significant on listening comprehension for fourth-grade students, with a median effect size of 0.50. The pattern of findings is interesting. For upper elementary students, Wanzek and Roberts (2012) found significant positive effects of interventions focused on listening (g = 0.93) when the intervention was tailored to the student’s skill profile but no significant impact when a one-size-fits-all intervention was used. This result reflects a promising area for future research. For the English learners at risk or with learning disabilities who were taught in English, across the set of studies, there were no significant differences found on vocabulary (Wanzek & Roberts, 2012), and for English learners with learning disabilities, across the set of studies, there was no significant difference found on passageway fluency. In each of these studies, listeners were asked to read a passage and examined the number of words read correctly and measured listening comprehension according to the strategy utilized by the student. The results were varied, with some studies showing significant positive effects and others showing no significant effects.
Exceptional Children

Kinder, Solari, and Gerber (2008) found significant positive effects for both their own measure of listening comprehension and the Woodcock Story Retell measure. Only one other effect was significant, for the Spanish reading intervention by Vaughn, Linan-Thompson, et al. (2006).

Moderating Variables

To examine the specific features of the intervention that may have moderated intervention outcomes, we ran regression analyses using the average unweighted effect size as the dependent variable and group size (individual or small group), minutes of intervention (as a continuous variable), and personnel delivering the intervention (research personnel or school-based personnel) as the independent variables. In studies with more than one intervention, we included in the analysis the intervention that we determined the authors hypothesized would have the strongest effect. This was done to ensure that all contrasts were independent, an assumption for ordinary least squares regression. Results indicated that for each regression analysis, no significant relationship was found between the potential moderator variable (group size, minutes of intervention, or personnel delivering the intervention) and the intervention outcomes.

Discussion

In this review, we examined the characteristics and outcomes of intervention studies that included data on English learners who were at risk for reading difficulties or had been identified as having a reading disability. We located 12 studies conducted since 2000 that used an RCT and met our criteria. The number of studies is dramatically smaller than the number of high-quality reading intervention studies that have been conducted with native English speakers over the same time period (see Edmonds et al., 2009; Solis, 2012; Wexler, Vaughn, Roberts, & Denton, 2010, for syntheses on reading interventions for non-English learners). But it compares favorably to periods prior to 2000 that addressed interventions with English learners found by Gersten and Baker (2000).

Unfortunately, given the limited sample of studies, the substantial amount of variation in the ages of the participants and types of interventions conducted, and the variations in measuring outcomes, it is difficult to determine patterns across the studies that would help identify potentially relevant trends. In this discussion, we summarize our findings illustrating patterns where possible, discuss the implications of our findings, and provide directions for future research.

Regarding the reading interventions in Spanish, it is important to take into account that these interventions are, in a way, different from the English interventions, because in the former studies used reading in the native language as an approach to providing reading support to struggling English learners, whereas in the latter studies explored the impact of an intervention in a second language on student reading performance. In a second language, the instruction is provided in the language of the literature and children’s vocabulary is enhanced. The presence of the native language is not necessarily beneficial, however, because in the former studies the instruction is in a way different from the English instruction, which is important in deciding to what extent the intervention is important. In studies with more than one intervention, we included in the analysis the intervention that we determined the authors hypothesized would have the strongest effect. This was done to ensure that all contrasts were independent, an assumption for ordinary least squares regression. Results indicated that for each regression analysis, no significant relationship was found between the potential moderator variable (group size, minutes of intervention, or personnel delivering the intervention) and the intervention outcomes.

Features of the Intervention Studies and Their Relationship to Impacts

Our review revealed a large variability in how English learners were identified and defined across studies. This trend has been an issue for many years, with frequent requests for more consistency in how English learners are defined. For example, in some of the studies, the school designation of English learners was used, whereas in other English learners were identified and defined in research studies. For example, in one study, English learners were identified and defined consistently across schools. This trend has been an issue for our review of English learners, but it is important to recognize that intervention studies are based on small groups, numbers of participants, and teachers. To examine the specific features of the intervention that may have moderated intervention outcomes, we examined the characteristics of the intervention, the reading performance of students in second language, and the intervention's impact.
et al., 2006). The practice of providing greater specification of the student samples is particularly critical for English learners, given the importance of language factors on achievement outcomes.

Across the 12 studies in our sample, only two reported an analysis of the differential effect of English language proficiency on outcomes. In the O'Connor et al. (2010) study, the language proficiency level of participants as measured by the California English Language Development Test did not have an impact on how well kindergarten students responded to the intervention. This is consistent with previous research with English learners on the weak association between language proficiency and reading growth in the early grades, particularly on foundational measures of reading, such as phonemic awareness and decoding (Chiappe, Siegel, & Wade-Woolley, 2002; Gersten et al., 2007). However, it differs from more recent research that found that language proficiency appears to affect early reading skills (D. Baker et al., 2013; Kieffer, 2008).

On the other hand, in the Lovett et al. (2008) study, which targeted English learners in Grades 2 through 8, the finding was that students who began the intervention with higher levels of language proficiency responded more positively to the interventions based on measures of phonemic blending and passage comprehension than students who began the intervention at lower levels of language proficiency. These findings are consistent with evidence that English language proficiency has an impact on student outcomes particularly in the upper elementary grades (see Geva & Farnia, 2012; Kieffer, 2010).

Further research is needed to determine how varying levels of English language proficiency affect the impact of an intervention. In particular, it may be that growth in basic reading skills (decoding and literal comprehension) is not related to higher levels of English language skill but that growth on higher-level skills (e.g., comprehension) is. Moreover, it may be that students who are technically exited from English learner status (often called former English learners; e.g., Parrish et al., 2006), but may not have developed the necessary academic English to be successful in school, are noticeably absent from the intervention studies in this review. That is, the English learner sample in these studies may be lower in English language proficiency than the population of English learners currently in American schools. Given this, the impact of interventions on these students may be lower than the impact of interventions on English language learners who are still in need of additional support. It is important to note that the counterfactual varied dramatically across studies, ranging from providing no intervention at all to providing the school's typical reading intervention. Thus, although the moderator analyses showed no significant role in predicting effect size, it does not mean that these intensity factors are not relevant.

One hallmark of Tier 2 interventions is the focus on small group instruction, particularly for students at risk of falling behind. Current research suggests that reasonably homogeneous small groups are often effective for delivering instruction, particularly for students at risk or with learning disabilities (Ehri et al., 2001; Elbaum, Vaughn, Hughes, & Moody, 1999). Eleven of the interventions reviewed used small-group instruction; two applied the intervention with students individually (Begeny et al., 2012; Vadasy & Sanders, 2010). When we analyzed group size as a moderating variable, we found there was no significant difference between small-group and individual instruction.

Intensity Factors of Group Size, Duration, Personnel, and Quality

Our review indicated there was large variation across interventions in terms of group size, minutes of instruction, and personnel delivering the instruction. As in most meta-analyses, it is hard to disentangle the length of intervention from numerous other factors. As discussed earlier, the nature of the counterfactual varied dramatically across studies, ranging from providing no intervention at all to providing the school's typical reading intervention. Thus, although the moderator analyses showed no significant role in predicting effect size, it does not mean that these intensity factors are not relevant. One hallmark of Tier 2 interventions is the focus on small group instruction, particularly for students at risk of falling behind. Current research suggests that reasonably homogeneous small groups are often effective for delivering instruction, particularly for students at risk or with learning disabilities (Ehri et al., 2001; Elbaum, Vaughn, Hughes, & Moody, 1999). Eleven of the interventions reviewed used small-group instruction; two applied the intervention with students individually (Begeny et al., 2012; Vadasy & Sanders, 2010). When we analyzed group size as a moderating variable, we found there was no significant difference between small-group and individual instruction.
tions or interventions delivered to six or more students, because English learners have more opportunities to practice the skill they are working on as well as their English language proficiency with their peers and the teacher, and the small-group settings provide more opportunities for this than they would get if they were in a large group (D. Baker & Kosty, 2012; Gersten & Jiménez, 1998). However, the results of this research synthesis do not demonstrate consistent, significant positive impacts or even consistently positive effects.

Interventions varied substantially in terms of the amount of instructional time provided. However, length of intervention did not predict magnitude of effect as the moderator analysis indicated. One reason that minutes of intervention may not have influenced effect sizes is that although intervention treatments are longer, so is the instruction provided to the students in the control group. In addition, shorter interventions tended to focus on just one or two reading outcomes and often measured only these specific outcomes, which may have accounted for larger effect sizes for these studies. For example, studies that focused on kindergarten students targeted only foundational reading or prereading skills (i.e., O'Connor et al., 2010; Solari & Gerber, 2008). In contrast, studies that focused on multiple components of reading—as many believe is most appropriate for Grades 1 and up—demonstrated quite mixed results. For the intermediate grades, the one study that tailored interventions to students' skill profiles (Wanzek & Roberts, 2012) tended to be much more effective than those with a “one-size-fits-all” approach. Consistent information about the nature of the counterfactual may play a role in the outcomes of the interventions. However, consistent information about the nature of the counterfactual may not have influenced effect sizes for English learners as much as it has for English-only students (Biancarosa & Snow, 2004; Torgesen et al., 2001).

For older students in middle school, minutes of instruction did not appear to have an impact on the results. For example, in the Vaughn et al. (2011) study, English learners in middle school received a full year of a Tier 3 reading intervention for 50 min a day, approximately 8,000 min of instruction. This intervention did not yield significant effects, suggesting that students in middle school may benefit more from interventions delivered to six or more students.
Although the number of studies is relatively small, we
found that many of the interventions focused on English
language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.

Outcomes of the Interventions

Although the number of studies is relatively small,
we found that many of the interventions focused on
English language skills as part of the intervention.

In summary, the interventions described in this review
showed significant effects for English learners.
Exceptional Children

wonder if more innovative intervention curricula that have a heavy language component might produce stronger effects than the current set of studies.

Implications for Future Research

In our search for studies for this review, we found 12 studies that used an RCT and met our inclusion criteria. More experimental studies ought to be conducted to determine what mal- leable factors have a significant effect on English learners’ academic performance. However, the pace of studies is improving, and the number of studies conducted since 2000 is much greater than similar time periods prior to 2000.

In terms of implications for future research, we recommend that researchers (a) focus on the individual differences in English learners, (b) consider development of interventions that focus on language and vocabulary and measures that capture language comprehension, and (c) include the calculation of an “effort variable” to be able to compare interventions. There are scarce studies that focus on English learners at risk for reading disabilities and even fewer studies that disaggregate the data by student language proficiency levels. English learners are a very heterogeneous group of students. They vary in terms of language proficiency, academic achievement, and the myriad predictor variables that may have an influence on growth and performance, such as poverty status and proficiency in their primary language. Future research should investigate interventions for English learners at varying language proficiency levels, including students who are technically exited from English learner status.

In addition, there is clearly a need to examine the effect of interventions that focus on language development and vocabulary as a core component of English learners at risk and those who have learning disabilities. We found very few studies that included a vocabulary and language development component and even fewer that included an RCT and met our inclusion criteria. More experimental studies ought to be conducted to determine what mal- leable factors have a significant effect on English learners’ academic performance. However, the pace of studies is improving, and the number of studies conducted since 2000 is much greater than similar time periods prior to 2000. In terms of implications for future research, we recommend that researchers (a) focus on the individual differences in English learners, (b) consider development of interventions that focus on language and vocabulary and measures that capture language comprehension, and (c) include the calculation of an “effort variable” to be able to compare interventions.

There are scarce studies that focus on English learners at risk for reading disabilities and even fewer studies that disaggregate the data by student language proficiency levels. English learners are a very heterogeneous group of students. They vary in terms of language proficiency, academic achievement, and the myriad predictor variables that may have an influence on growth and performance, such as poverty status and proficiency in their primary language. Future research should investigate interventions for English learners at varying language proficiency levels, including students who are technically exited from English learner status. Evaluating how individual differences in language proficiency influence intervention outcomes will allow the field to refine interventions to better meet the needs of these students.

In addition, there is clearly a need to examine the effect of interventions that focus on language development and vocabulary as a core component of English learners at risk and those who have learning disabilities. We found very few studies that included a vocabulary and language development component and even fewer that included an RCT and met our inclusion criteria. More experimental studies ought to be conducted to determine what mal- leable factors have a significant effect on English learners’ academic performance. However, the pace of studies is improving, and the number of studies conducted since 2000 is much greater than similar time periods prior to 2000. In terms of implications for future research, we recommend that researchers (a) focus on the individual differences in English learners, (b) consider development of interventions that focus on language and vocabulary and measures that capture language comprehension, and (c) include the calculation of an “effort variable” to be able to compare interventions. There are scarce studies that focus on English learners at risk for reading disabilities and even fewer studies that disaggregate the data by student language proficiency levels. English learners are a very heterogeneous group of students. They vary in terms of language proficiency, academic achievement, and the myriad predictor variables that may have an influence on growth and performance, such as poverty status and proficiency in their primary language. Future research should investigate interventions for English learners at varying language proficiency levels, including students who are technically exited from English learner status. Evaluating how individual differences in language proficiency influence intervention outcomes will allow the field to refine interventions to better meet the needs of these students.

References

References marked with an asterisk indicate studies included in the meta-analysis.


