Learning Disabilities Research & Practice, *33*(2), 64–74

DOI: 10.1111/ldrp.12160

LEARNING DISABILITIES
RESEARCH & PRACTICE

Review of Single Subject Research Examining the Effectiveness of Interventions for At-Risk English Learners

Doris Luft Baker
Southern Methodist University
Catherine Richards-Tutor
California State University
Anthony Sparks
Southern Methodist University
Rebecca Canges
Metropolitan State University, Denver

We review published single subject design (SSD) studies that examine the effects of interventions for English learners at-risk or with learning disabilities. Results of our literature search yielded 10 studies, five in reading, one in reading and behavior, and four in mathematics that met our inclusion criteria. Seven studies targeted Spanish-speaking English learners, and three studies included students who spoke other languages than English and/or English only students. Two studies in mathematics included native language instruction. Six studies included English learners in second grade and above, and one study included high school students. We were able to calculate effect sizes (Hedges *g*) for eight of the 10 studies. Findings indicated a significant effect of the intervention for 12 of the 18 dependent variables measured. SSD methodology has the potential to help researchers and practitioners better understand what interventions work for English learners, and under what circumstances.

The use of single subject design (SSD) studies to identify evidence-based practices that benefit students who are at-risk for developing a learning disability, or who have been identified as having a disability, can be a powerful and useful design to ensure the validity of interventions (Horner et al., 2005). SSD studies use a rigorous scientific methodology that requires the systematic and detailed analysis of individual behaviors with the purpose of documenting causal relationships between the independent and dependent variables (Cook & Cook, 2016). SSD studies are particularly useful and cost effective for the identification of educational and behavioral interventions that might be appropriate for later large-scale analysis. Moreover, given that high-quality SSD studies also tend to provide in-depth information regarding the intervention, instruction, and student participants, it is a useful methodology to better understand the particular needs of a population provided the following components are included in the study: (a) the practice or intervention is defined operationally, (b) a detailed description of the context is included, (c) fidelity of implementation is reported, (d) results indicate that changes in the dependent measures are related to the independent variables, and (e) the effects of the intervention can be replicated across multiple sites and diverse participants (Horner et al., 2005; Kratochwill et al., 2010).

A benefit of using SSD studies over randomized controlled trials (RCTs) is the ability to analyze one student, or a small group of student responses to an intervention, and then provide a targeted intervention that will benefit that particular student or small group. By focusing on one participant or small group at a time, SSD studies cut down on some threats to internal validity, such as differential selection, selection maturation, and diffusion (Springer, 2010). Another benefit of using SSD studies is the ability to "bypass" certain complications presented by group designs, such as group variability (Neuman, 2011).

SSD studies might be also useful to study heterogeneous populations such as English learners (ELs), particularly English learners who are at-risk for academic failure. "English learners" refers to students who do not speak English as their native language, and who benefit from additional supports until they are proficient enough in English that they can benefit from mainstream classroom instruction (August & Shanahan, 2006). For example, according to the Institute of Education Sciences (2015), the approximate number of Spanish-speaking English learners in schools is 79 percent. However, not all Spanish-speaking English learners share similar demographic characteristics. Approximately, 50 percent of them were born in the United States, but their parents could have been originally from Mexico, Puerto Rico,

Requests for reprints should be sent to Doris Luft Baker, Department of Teaching and Learning, Simmons School of Education, Southern Methodist University. Electronic inquiries should be sent to dluftdebaker@smu.edu.

Cuba, El Salvador, Honduras, Spain, or another Latin American country. Furthermore, a large percentage (32 percent, according to Kena et al., 2015) live in poverty, and their parental level of education could range from 52 percent who have a high school education or below, to 10 percent who have a Bachelor's degree (Goldenberg, Reese, & Rezaei, 2011). Similarly, the 21 percent of English learners who are not Spanish-speaking have many different nationalities, and there are large differences in socioeconomic status, parental level of education, and level of English language proficiency among them.

Although not all English learners are at-risk for academic failure, according to the National Assessment of Educational Progress, English learners appear to score significantly lower in mathematics and reading compared to non-English learners. For example, the most recent data indicate that English learners scored 38 points lower than non-English learners on the fourth-grade reading assessment, and 45 points lower on the eighth-grade reading assessment. The average scaled score in reading in 2013 was 221 in fourth grade, and 266 in eighth grade. Similarly, English learners scored 25 points lower than non-English learners on the fourth-grade mathematics assessment, and 41 points lower on the eighth-grade mathematics assessment. The average scaled score in mathematics in 2013 was 242 in fourth grade, and 285 in eighth grade (Kena et al., 2015).

The academic struggles of English learners have several different causes that might be difficult to tease out, such as low socioeconomic status, low English language proficiency, low vocabulary difficulties in decoding (Baker, Richards-Tutor, Gersten, Baker, & Smith, 2017), and low understanding of procedural mathematics (e.g., understanding addition and subtraction; Orosco, Swanson, O'Connor, & Lussier, 2011). Other potential reasons for poor academic performance can be attributed to poor quality instruction on key elements such as vocabulary (Baker et al., 2017; Tong, Lara-Alecio, Irby, Mathes, & Kwok, 2008), and lack of access to research-based materials in the native language if programs provide native language instruction (Baker, Basaraba, & Polanco, 2016; Vaughn et al., 2006).

For example, Baker et al. (2016) found that vocabulary instruction was the least taught reading component in Spanish or in English in bilingual classrooms, and that students in classrooms that provided better vocabulary instruction also made larger gains on vocabulary measures than students in classrooms with low vocabulary instruction. Regarding the quality of instruction, Tong et al. (2008) conducted an experiment to examine whether enhancing a transitional bilingual program (where students learn to read in their native language first before transitioning to English-only instruction) improved reading outcomes for ELs. Their findings indicated that ELs in the enhanced transitional program outperformed ELs in the typical transitional bilingual program on almost all the measures in Spanish and in English (effect size [*ES*] ranged from 0.12 to 0.71).

Thus, addressing the needs of this struggling population of students using evidence-based interventions should be a priority (Every Student Succeeds Act [ESSA], 2015). According to Gersten et al. (2005), evidence-based interventions refer to practices that have been tested experimentally in at least two studies where participants have been randomly assigned to a treatment or control group, and a weighted effect size that is significantly different from zero has been calculated. Although this rigorous testing of interventions is desirable, a review by Richards-Tutor, D. L. Baker, Gersten, S. K. Baker, and Smith (2015) indicated that few interventions targeting English learners at-risk or with learning disabilities have been tested using these rigorous experimental designs. In fact, the researchers found only 12 studies from 2000 to 2012 that followed Gersten et al.'s (2005) definition of evidence-based practices. Out of these 12 studies, only two were conducted in Spanish, 8 of the 12 were conducted in kindergarten and first grade, only one addressed the needs of upper grade English learners at-risk, only one specifically included English learners with reading disabilities, and all of them focused solely on reading interventions.

Given the dearth of RCTs that include English learners who are struggling academically, the need for examples in the research that provide a detailed description of methods, participants, and interventions for English learners, and the need for a more in-depth examination of individual response to intervention of a heterogeneous English learner population, we decided to search for other studies that used an experimental design such as single-subject experiments. Thus, the purpose of this article is to review published SSD studies that examined the effects of interventions for English learners at-risk or with a disability from 2000 to 2015. Our additional criteria to include studies are specified in detail in the "Methods" section.

TYPES OF SINGLE-SUBJECT RESEARCH AND QUALITY INDICATORS

In general, single-subject research uses three different types of designs with some variations: baseline-treatment, multiple-baseline, and alternating treatments (Springer, 2010). Baseline-treatment design or reversal design (e.g., ABAB; Cook & Cook, 2016) involves two different stages: baseline (usually denoted by the letter A), where the researcher obtains a sense of student performance in typical conditions; and intervention, where the researcher introduces the independent variable of study (usually denoted by the letter B). After the intervention, student performance is measured and compared to the performance in typical conditions. Sometimes, to determine if the effects of the intervention were maintained, the intervention is withdrawn (A_2) to examine whether improved outcomes of the intervention during the first intervention phase (i.e., B_1) were not a coincidence. After collecting multiple data points during the second baseline, the researcher introduces the intervention again (B_2). If the behavior changes during this second intervention are similar to the behavior changes during the first intervention, then it can be reasonably concluded that changes in behavior were caused by the intervention.

Multiple-baseline designs involve the repeated measure of student performance at baseline before the intervention is introduced to other students at different points in time. A variation of multiple-baseline designs is multiple-probe designs (Gast, Lloyd, & Ledford, 2014). The main difference

between a multiple-probe design and the multiple-baseline design is that a multiple-probe design measures student behaviors less frequently than a multiple-baseline design. According to Horner and Baer (1978), if there is a reason to believe that baseline measures will remain steady, then a multiple-probe design might be more appropriate than using a multiple-baseline design.

The last type of SSD study is alternating-treatments design. Unlike switching from baseline to intervention with at least 3 data points in each phase, as in baseline designs, alternating-treatments design allows for more than one treatment (i.e., intervention) to be tested and compared to another intervention from the start (Springer, 2010). An advantage of using alternating-treatments design is the ability to quickly compare different interventions, allowing the researcher to make appropriate instructional decisions immediately (Neuman, 2011). A disadvantage of this design, however, is that it is not possible to create a trend line for each different condition because they are all occurring very close to each other.

Horner et al. (2005) proposed a set of quality indicators for SSD studies that helps researchers and practitioners determine whether single-subject research methods have been applied adequately within a study. These quality indicators included specific criteria for participant descriptions; dependent and independent variable standards; inclusion and description of baseline phase; and discussion of internal, external, and social validity. Although the scope of this study is not to review the included studies to determine evidence-based practices, we did use these quality indicators as a guide for the inclusion criteria in this literature review.

ANALYZING RESULTS FROM SSD STUDIES

Until recently, statistical analyses of SSD studies relied on graphical interpretation to determine effects of the intervention (Kazdin, 2010). However, more recently, two methods have been developed to statistically calculate between-case (i.e., the variability between participants in the study) and within-case (i.e., the variation of one participant over time) effect sizes. For example, if a study included multiple English learners and the researchers wanted to know the variation between English learners across the different phases, then a between-case effect size would be calculated. However, if the study wanted to obtain an idea of how the intervention worked for one English learner, then a within-case effect size would be calculated.

A major issue with SSD studies involves external validity, or the extent to which results can be generalized to a larger population (Springer, 2010). Given that participants of SSD studies are not randomly selected from the population they represent, it is hard to draw conclusions on the extent to which effects would hold for the larger population. Neuman (2011) has argued that generalizability should not be the central focus of SSD studies, given that the goal is to examine the effects of a treatment for an individual or a group of individuals. However, Springer (2010) has suggested that SSD studies can strengthen their external validity by researchers choosing materials for the intervention and assessments carefully, and by providing researchers and practitioners with detailed descriptions of all materials and procedures used during the intervention. Thus, the purpose of this manuscript is to review published SSD studies since 2000 that have included English learners at-risk or with a learning disability to further our knowledge base on effective interventions, and/or features of an intervention, that might improve outcomes for struggling English learners.

METHOD

Criteria for Selecting Studies

We used the following criteria to review SSD studies: (a) the majority of participants were English learners identified as at-risk or with a learning disability (using either standardized tests or valid screening measures); (b) clear and replicable definitions of the dependent and independent variables were provided, including fidelity of implementation; and (c) some type of internal, external and social validity information was provided. These criteria are closely aligned to the Horner et al. (2005) quality indicators, which have been universally accepted.

The study team searched for studies in PsycInfo and ERIC following key words in peer-reviewed journals: English learners, language minority students, second language learners, intervention, response to intervention, at-risk, learning disabilities, reading, reading difficulty, writing, writing difficulty, math, math difficulty, behavior, behavior difficulties. We then specifically searched for studies in the following journals: *International Journal of Learning Disabilities*, *Journal of Educational Psychology, Journal of Learning Disability, Learning Disabilities Research and Practice, Learning Disabilities Quarterly, Reading Research Quarterly, Remedial and Special Education, Scientific Studies of Reading, Journal of Special Education, Exceptional Children, Journal of Literacy Research,* and *Topics in Language Disorders.* We identified 12 potential studies. We then read the abstracts and only selected for further analysis those studies that indicated that they used an SSD and included struggling English learners as participants (i.e., students who were receiving a Tier 2 or Tier 3 intervention). Next, we carefully reviewed the method section of the 12 articles to ensure that the studies met our other criteria listed above. Ten studies out of the 12 met our inclusion criteria.

We then summarized the information from each of the 10 studies by two broad features: (a) features of the research studies such as research design, grade level, participant characteristics, and setting; and (b) features of the intervention such as group size, duration, personnel delivering the intervention, content, and intervention methods. Table 1 presents the features of the research studies, and Table 2 presents the features of the interventions used in these studies. Two raters independently summarized the information from each of the studies using these categories and subcategories. Agreement between raters was 90 percent or above for each of the features. Disagreements were resolved through discussions by the two raters.

TABLE 1
Features of the Single Subject Design Studies

Authors	*Design*	*Grade Level*	*English Learner Sample Size*	*EL Primary Language*	*Risk Determination*
Gyovai et al. (2009)	Multiple baseline	K-1	Eleven ELs from K and one EL from first 10 Somali, 1 Vietnamese, 1 Hispanic	Four languages, majority Somali	At-risk or some risk on DIBELS; Below average grade level performance on Woodcock Johnson-III
Orosco (2014a)	Multiple baseline	3	Six ELs Latino	Spanish	At-risk in Math on District Math Assessment; Below grade level of District Reading Assessment; Scored lower than 25th percentile on Bateria III Woodcock-Munoz; Teacher referral
Orosco (2014b)	Multiple baseline	3	Six ELs Latino	Spanish	At-risk in Math on District Math Assessment; Below grade level of District Reading Assessment; Scored lower than 25th percentile on Bateria III Woodcock-Munoz; Teacher referral
Orosco (2013)	Multiple baseline	2	Six ELs. Latino	Spanish	At-risk in Math on District Math Assessment; Below grade level of District Reading Assessment; Scored lower than 25th percentile on Bateria III Woodcock-Munoz; Teacher referral
Orosco et al. (2011)	Multiple baseline	2	Six ELs Latino	Spanish	At-risk in Math on District Math Assessment; Below grade level on District Reading Assessment; Below basic on Woodcock-Johnson NU Test of Achievement III; Teacher referral
Preciado, Horner, and Baker (2008)	Multiple baseline	2, 3, 4	Two from second, one from third, one from fourth Latino	Spanish	At-risk behaviors during reading based on FBAs; Demonstrated Escape-Maintain behaviors based on Functional Analyses (FAs); Teacher referral
Rahn et al. (2015)	Multiple baseline	K, 1	One from K, one from first, one Chinese and Somali	Hmong and Somali	At-risk for Reading based on Minneapolis Kindergarten Assessment; Teacher referral
Ross and Begeny (2011)	Alternating treatments	2	Five	Spanish	At-risk or some risk on DIBELS DORF; Teacher referral
Santoro et al. (2006)	Multiple baseline	2	Four	Three languages	Teacher referral; Terra Nova Test of Cognitive Skills
Viel-Ruma et al. (2010)	Multiple baseline	Ninth to eleventh	Six, three ELs, Hispanic	Spanish	School Determination

Statistical Analysis

To calculate the between-subject effect sizes, we used an SPSS macro developed by Marso and Shadish (2015). This macro calculates an estimate of an effect size (i.e., Hedges *g*) for SSD studies that include at least three participants, and a baseline or multiple-baseline design (Hedges, Pustejovsky, & Shadish, 2012, 2013). Effect sizes for alternate-treatments design cannot be calculated using this estimator, because in alternate-treatments designs, students tend to receive different interventions on the same day, or between very short periods of time. Therefore, a trend for each of the interventions cannot be calculated. However, the benefit of using Marso and Shadish's estimator over others for baseline or multiple-baseline designs is because of the way it addresses issues related to observations being dependent upon one another as well as accessibility to the software (Shadish, Hedges, Horner, & Odom, 2015).

The main statistical considerations to account for when computing between-effect sizes are trend, dependency of observations, and nonnormality of residuals (Shadish et al., 2015). Like the other available effect-size estimators, Marso and Shadish's estimator accounts for dependency of observations. However, unlike other estimators, Marso and Shadish's makes no assumption about trend. To attenuate concerns of trend, the recommended detrending option was applied to each study and compared to the original effect sizes (Marso & Shadish, 2015). None of the effect sizes were considerably different after detrending, which confirmed the use of Marso and Shadish's estimator. Finally, all the currently available

TABLE 2
Single Subject Intervention Features

Authors	*Group Size*	*Intervention Duration*	*Interventionist*	*Intervention Program*	*Intervention Content*
Gyovai et al. (2009)	4	15, 11, or 7 weeks; 2–4 days per week; 20 min per day; (560–1,200 min)	Researchers	Early Reading Intervention	Letter naming, letter sounds, letter writing, initial sound isolation, and picture matching.
Orosco (2014a)	1-on-1 instruction	5 weeks; 17 sessions (20–25 min per session)	One teacher One researcher	Estrategia Dinámica de Matemáticas (EDM)	The teacher systematically modifying word problem vocabulary via a four-level linguistic modification procedure to match the students' level of vocabulary, and then providing strategy instruction with probes that assessed students' ability to solve problems.
Orosco (2014b)	1-on-1 instruction	5 weeks; 17 sessions (20–25 min per session)	One teacher One researcher	Dynamic Strategic Math (DSM)	Systematically modifying the vocabulary (via a four-level linguistic modification procedure) to the individual student's understanding level of the word problems and then providing strategy instruction with probes that assessed students' ability to solve problems.
Orosco (2013)	1-on-1 instruction	5 weeks; 17 sessions (20–25 min per session)	One teacher One researcher	Estrategia Dinámica de Matemáticas (EDM)	The teacher systematically modifying word problem vocabulary via a four-level linguistic modification procedure to match the students' level of vocabulary, and then providing strategy instruction with probes that assessed students' ability to solve problems.
Orosco et al. (2011)	1-on-1 instruction	5 weeks; 17 sessions (20–25 min per session)	One teacher One researcher	Dynamic Strategic Math (DSM)	Systematically modifying the vocabulary (via a four-level linguistic modification procedure) to the individual student's understanding level of the word problems and then providing strategy instruction with probes that assessed students' ability to solve problems.
Preciado, Horner, and Baker (2008)	1-on-1 instruction	14 weeks; 1 hr of instruction conducted 4 days a week	Six graduate students	Reading Mastery and a language-matched instructional priming (LMIP) Program	Teaching decoding skills, reviewing/previewing the content and vocabulary in the story that would be read in class the next day, reviewing instructions for completing the next day's independent task, and teaching more socially acceptable social skills.
Rahn, et al. (2015)	1-on-1 instruction	17 weeks; 10–15 min per session one time a day for 4 days a week	Six graduate students	Incremental Rehearsal	Letter sound expression and fluency
Ross and Begeny (2011)	1-on-1 instruction and small group	8 weeks; 13 min per session, 180 min	Research Assistants	NA	Listening passage preview, repeated reading, retell, phase drill error correction
Santoro et al. (2006)	1-on-1 instruction	7, 8, 11, or 14 weeks; 4 days per week; 30 min each session	One teacher and two graduate students in special education	Read Well	Phonological awareness, phonics, fluency, vocabulary, and comprehension
Viel-Ruma et al. (2010)	6	5 weeks; 30–45 min sessions	Three special education teachers	Expressive Writing I	Direct Instruction program including mechanics, sentence writing, paragraph writing, editing

TABLE 3
Effect Sizes for 16 Dependent Variables

Dependent Variable by Study	*n*	*g*	$S^{2/g}$	*95 Percent CI*	
				Lower Bound	*Upper Bound*
Gyovai et al. (2009)					
Phonological awareness	12	1.028*	.047	.606	1.453
Phonemic segmentation	12	.808*	.055	.348	1.268
Orosco (2014a)					
Word problem solving	6	2.698*	.231	1.756	2.640
Orosco (2014b)					
Word problem solving	6	2.672*	.247	1.698	3.646
Orosco (2013)					
Word problem solving	6	2.062*	.217	1.149	2.975
Orosco et al. (2011)					
Word problem solving	6	1.982*	.208	1.088	2.876
Preciado, Horner, and Baker (2008)					
Percent intervals with total problem behavior	4	1.479*	.185	.636	2.322
Percent completed tasks	4	1.790*	.107	1.149	2.431
Percent correct tasks	4	1.373*	.082	.812	1.934
Santoro et al. (2006)					
Oral reading fluency	4	.440	.107	−.201	1.081
Phonemic segmentation	4	.737	.393	−.492	1.966
Letter naming fluency	4	−.060	.080	−.614	.494
Letter sound fluency	4	1.154	.482	−.207	2.515
Nonsense word fluency	4	.485	.205	−.402	1.372
Viel-Ruma et al. (2010)					
Percent correct word sequence					
Native English speakers	3	1.721*	.173	.906	2.536
English learners	3	.931*	.188	.081	1.781
Number of words written					
Native English speakers	3	.871*	.111	.218	1.524
English leaners	3	.682	.143	−.059	1.423

Note. g, Hedges *g*; $S^{2/g}$, variance of Hedges *g*.
*$p < .05$.

effect-size estimators assume normally distributed residuals. However, current literature suggests that the between-effect size estimators may be robust to nonnormal residuals due to the nature of the count data commonly used in single-subject research (Shadish et al., 2015). We also conducted a significance test for each of the effect sizes calculated (Shadish, Hedges, & Pustejovsky, 2014), and 95 percent confidence intervals were created with z-scores to determine if the interval with the effect sizes included zero. If the interval did not include zero, the effect size was significant at the $p < .05$ level. Results of the significance test, and the percentage of variance explained is included in Table 3.

RESULTS

Ten studies met our criteria for inclusion in this review. We summarize results based on study characteristics (i.e., design, participants, and setting) and on features of the intervention (i.e., group size, intervention duration, intervention program, and intervention content). We report effect sizes for each outcome measured in the intervention for eight of the 10 studies. The reason to include effect sizes for only eight studies was because one of the studies did not have three participants (Rahn et al., 2015), and the other study used an alternating-treatments design with three conditions (Ross & Begeny, 2011). The three conditions (i.e., 1-on-1, small-group, and no-treatment) were assigned to each student in a predetermined random order that did not allow for estimation of the between-subjects effects using Marso and Shadish's macro because each student received a different order of the intervention at each time point. Consequently, there were not enough consecutive data points to establish a trend for each of the three different conditions. Data from these two studies were inspected visually.

Study Characteristics

All 10 studies took place in different regions of the United States. Nine of the 10 studies used a multiple-baseline design, as indicated in Table 1, and one study used an alternating-treatments design (Ross & Begeny, 2011). The sample sizes of the single-subject studies varied from 2 to 11. Two of the studies included kindergarten and first-grade students (Gyovai, Cartledge, Kourea, Yurick, & Gibson, 2009; Rahn et al., 2015); four included only second-grade students (Orosco, 2013; Orosco et al., 2011; Ross & Begeny, 2011; Santoro,

Jitendra, Starosta, & Sacks, 2006); two included only third-grade students (Orosco,2014a, 2014b); and one study included second-, third-, and fourth-grade students (Preciado, Horner, & Baker, 2008). Finally, one study included ninth-, tenth-, and eleventh-grade students (Viel-Ruma, Houchins, Jolivette, Fredrick, & Gama, 2010). Nine out of the 10 studies included only English learners as participants. Viel-Ruma et al. included both students who were English learners and native English speakers. Ross and Begeny (2011), Orosco et al. (2011), Orosco (2013, 2014a, 2014b), and Preciado et al. (2008) included English learners who spoke Spanish, while student participants in the other three studies had a variety of ethnic backgrounds and spoke multiple first languages (Gyovai et al., 2009; Rahn et al., 2015; Santoro et al., 2006). In each of the 10 studies, English learner status was determined by school designation. Each of the studies included English learners at-risk for reading disabilities/difficulties, and only one of the studies included students with identified learning disabilities (Viel-Ruma et al., 2010). In Santoro et al.'s study, one student was identified with a learning disability during the study. See Table 1 for details on the procedures used to identify students.

Features of the Interventions

Group Size

Unlike the typical small-group interventions conducted in RCT studies (see Richards-Tutor et al., 2015), the single-subject studies used both individual interventions and small-group interventions as indicated in Table 2. For example, seven of the studies delivered the interventions to students individually (Orosco et al., 2011; Orosco, 2013, 2014a, 2014b; Preciado et al., 2008; Rahn et al., 2015; Santoro et al., 2006), and in the other four studies, interventions were delivered in small group, or individually and in small group. In Ross and Begeny's (2011) study, students were provided with a mix of individual and small-group intervention sessions. In Gyovai et al.'s (2009) study, students were provided the intervention in a small group of four, while the Viel-Ruma et al. (2010) studies delivered the intervention in small groups of six.

Duration of the Interventions

The duration of the interventions varied across studies. For example, in Ross and Begeny's (2011) study, all participants received 180 min of instruction for 8 weeks of intervention. In the Santoro et al. (2006) and the Gyovai et al. (2009) studies, intervention duration varied by participant. In Gyovai's study, students received 560–1,200 min of instruction that ranged from 7 weeks to 15 weeks. Similarly, in Santoro et al.'s (2006) study, students received between 840 and 1,680 min of instruction in 30-min sessions, 4 days a week, for 7–14 weeks. In the Orosco studies, students received instruction over the course of 5 weeks in seventeen 20–25 min sessions. Rahn et al.'s (2015) study also lasted 17 weeks. However, students were provided the intervention 4 days a week for 10–15 min a day. Likewise, Preciado et al. (2008) implemented their intervention 4 days a week in 1-hr sessions for 14 weeks. Finally, Viel-Ruma et al.'s (2010) study delivered their small group intervention over the course of 5 weeks, for 30–45 min sessions per day.

Methods of Intervention Delivery

Systematic, explicit instruction such as modeling, scaffolding, and corrective feedback was a common feature of the interventions across all studies. The four Orosco studies were the only ones that included vocabulary instruction to meet the needs of English learners, and two of the studies (i.e., Orosco, 2013, 2014a) used the strategies from Dynamic Strategic Math in the students' native language, Spanish.

Personnel Delivering the Interventions

In all cases, personnel were trained on how to deliver the intervention, and in addition, they were observed and provided with feedback. Interventions were delivered by teachers, researchers, and graduate students across all the studies. In two of the studies, special education teachers delivered the intervention (Santoro et al., 2006; Viel-Ruma et al., 2010), whereas in the four studies by Orosco, a teacher and a researcher delivered the interventions. In the final four studies, research assistants, undergraduates or graduate students delivered the interventions (Gyovai et al., 2009; Preciado et al., 2008; Rahn et al., 2015; Ross & Begeny, 2011).

Fidelity of Implementation

All 10 studies reported adequate levels of fidelity. Interventionists in each study received training before the intervention began, and often coaching was provided if specific elements of the intervention were not implemented as intended. In addition, personnel were trained on measuring fidelity of implementation through the continuous direct measurement of the independent variable as suggested by Horner et al. (2005).

Content of the Interventions

Two of the single case studies reported here (Preciado et al., 2008; Santoro et al., 2006) used a multicomponent intervention that included various components of reading such as decoding, fluency, vocabulary, and comprehension. However, two other studies (Gyovai et al., 2009; Ross & Begeny, 2011) focused on just two components of reading: phonological awareness and alphabetic knowledge in kindergarten (Gyovai et al., 2009), and fluency and vocabulary in second grade (Ross & Begeny, 2011). Two of the studies used previously existing intervention curricula such as Read Well[1] (Santoro et al., 2006), and Early Reading Intervention[2] (Gyovai et al., 2009). Rahn et al. (2015) implemented the Incremental Rehearsal intervention that focused on letter sound expression and fluency. The Orosco et al. (2011) and Orosco (2013, 2014a, 2014b) studies all focused on teaching students

dynamic mathematics strategies for modifying word problems that also included vocabulary explanations and modifications. The 2011 and 2014b studies delivered the intervention in English (i.e., Dynamic Strategic Mathematics, DSM), and the 2013 and 2014a studies delivered the intervention in Spanish (i.e., Estrategia Dinámica de Matemáticas, EDM). Finally, Viel-Ruma et al. (2010) used an existing direct instruction intervention, Expressive Writing, which included mechanics, as well as sentence and paragraph writing and editing.

Literacy

Two of the single-subject studies measured phonological awareness (Gyovai et al., 2009; Santoro et al., 2006). Gyovai et al. included students in kindergarten and first grade, and Santoro et al. included students in second grade. Both studies used DIBELS measures (i.e., Dynamic Indicators of Basic Early Literacy Skills; Good & Kaminski, 2002). In both studies, all students made growth after receiving the intervention in phonemic awareness, as measured by the DIBELS Phonemic Segmentation Fluency subtest. However, only two of four students in Santoro et al.'s study made growth in decoding after receiving the intervention, as measured by the DIBELS Nonsense Word Fluency subtest. Moreover, in Santoro et al.'s (2006) study, only three out of four students made growth in oral reading fluency after the intervention, as measured by the DIBELS Oral Reading Fluency subtest. As indicated in Table 3, effect sizes for early reading variables were not significant. In Ross and Begeny's (2011) study, four of the five second grade students made growth after receiving the individual intervention, and two of the five students made growth after receiving the small-group intervention. We could not calculate effect sizes, however, because of the type of design used in this study.

Viel-Ruma et al.'s (2010) study focused on writing for secondary students. In this study, both English learners and native English speakers made growth on the two measures of writing used: percent correct word sequences and number of words written. The effect sizes ranged from 0.62 to 1.72. Effect sizes for native English speakers were higher, and the effect size for the number of words written was not significant for English learners, but it was significant for native English speakers.

Mathematics

The four mathematics studies measured word problem solving. In the four studies, all students made overall growth with significant effect sizes, as indicated in Table 3.

Behavior

One study focused on behavior gains of students (Preciado et al, 2008). Overall, all students made growth on the three behavior outcomes: percent intervals with total problem behavior, percent completed tasks, and percent correct tasks. Effect sizes were all significant and ranged from 1.37 to 1.79.

DISCUSSION

This article reviewed published single-subject research studies since 2000 designed to improve the academic performance of English learners at-risk or with a learning disability. Our thorough search only yielded 10 SSD studies that met our criteria aligned with Horner et al. (2005) quality standards for SSD studies, and only one was conducted with English learners who had been diagnosed with learning disabilities. This was somewhat surprising, particularly given the fact that SSD appears to be a valid and feasible approach to experimentally examine more closely the needs of a small group of students (Cook & Cook, 2016; Horner et al., 2005). For example, a school district may only have a small percentage of students who are English learners, but these students may have very particular needs. In another case, a school district might have a large number of English learners, but they all might have very different backgrounds and native languages, which might make it difficult to test the effects of interventions for such a heterogeneous population using RCTs.

Out of the 10 studies we found, we were only able to calculate effect sizes for eight, using Marso and Shadish (2015) tools. Based on our independent calculations, eight studies out of the 10 reviewed indicated that the intervention had significantly changed outcomes for students, with large effect sizes (i.e., above 0.75). Studies with the largest effect sizes were the studies by Orosco (i.e., Orosco, 2013, 2014a, 2014b; Orosco et al., 2011) in second and third grades, the study by Preciado et al. (2008) in second to fourth grade, the study by Viel-Ruma et al. (2010) in ninth to eleventh grade, and the study by Gyovai et al. (2009) in kindergarten and first grade.

Although all the studies reviewed used a multiple-baseline design, they all varied in terms of the type of intervention that was provided, the duration, and the group configuration, with the exception of the four Orosco studies (i.e.,Orosco, 2013, 2014a, 2014b; Orosco et al., 2011). Only these four studies used the same type of strategies to teach word problem solving, and all participants in the four studies attended the same elementary school. In two of the studies, participants were in second grade, and in the other two, participants were in third grade. However, the intervention in each grade was conducted in either Spanish or English, with a different group of six Latino students. In all of the other six studies, none of the interventions were used twice, and all students varied in terms of age, grade, and school they attended. Nonetheless, all these studies suggest that SSD studies can potentially be a viable approach to identify effective practices for English learners at-risk or with learning disabilities, as suggested by Horner et al. (2005) and Cook and Cook (2016). We recommend that future research consider this approach, particularly given the heterogeneity of our English-learner population who may be at-risk or may have learning disabilities.

Only one of the SSD studies we found specifically examined interventions for students with identified learning

disabilities who were also English learners. This was somewhat surprising, given that SSD studies seem to be a good fit for studying effective interventions for this population of students. However, most of the intervention research we located focused on students in the early elementary grades, before students are identified for learning disabilities. Often teachers are reluctant to refer English learners to receive special education services before they have had the opportunity to develop their English language proficiency. Thus, research has indicated that English learners tend to be underidentified in the early grades for learning disabilities, but overidentified in the upper elementary grades and beyond (Artiles, Rueda, Salazar, & Higareda, 2005; Samson & Lesaux, 2009). Although we do not want to overidentify students, underidentification can potentially lead to a reduction in services at an earlier age. Also, the underidentification of English learners in the earlier grades can lead to a lack of experimental research that examines effective interventions for this specific population in earlier grades, when it might be easier to provide them with the supports they need to succeed in school.

It is very difficult to provide specific recommendations about best practices for English learners at-risk or with learning disabilities from this review of the literature, because not only did we find a small number of SSD studies, but there was also quite a bit of variability across the studies (i.e., with the exception of the four mathematics studies), in terms of their purpose, type of intervention provided, and participant characteristics. For this reason, we were unable to conduct a meta-analysis, as suggested by Shadish et al. (2015) where we could calculate an effect size for all combined interventions that were similar in scope and that targeted a similar population of students. This type of analysis would be desirable because it would increase the external validity of an intervention, or specific features of an intervention. In this review, only the Orosco mathematics studies might have been suitable to conduct a meta-analysis.

Next we compare the outcomes of this review to the outcomes from the review conducted by Richards-Tutor et al. (2015) of RCTs that included English learners at-risk or with a disability. Similar to our findings, Richards-Tutor et al. (2015) found only 12 published studies between 2000 and 2012 that targeted the English learner student population, and as with this review, only the four mathematics studies by Orosco (i.e., Orosco, 2013, 2014a, 2014b; Orosco et al., 2011) focused on examining the effects of a language or vocabulary intervention as a core component of the supports that English learners at-risk or with a learning disability need. Richards-Tutor et al. (2015) found that the only variables that appeared to be common across studies were small group size and direct, explicit instructional methods. Moderator variables such as minutes of intervention, number of students in the group, and personnel delivering the intervention were not significant predictors of English learner performance. In this study, with the exception of the four mathematics studies, it was difficult to discern the specific characteristics that would significantly improve outcomes for English learners, given that the studies were so different in scope and focus. However, all studies did use small group instruction and explicit, direct instructional methods, just like the studies reviewed by Richards-Tutor et al. (2015). The Orosco studies all used the same Dynamic Strategic Math intervention that included a strong emphasis on vocabulary with different scaffolding procedures that included (a) preteaching of mathematics concepts, (b) comprehensive strategy instruction to solve word problems, and (c) a collaborative approach to check for understanding.

The small number of studies in both reviews suggests the need to conduct more rigorous studies where the design of the interventions includes specific instructional strategies known to be effective for English learners. Examples of these strategies include teaching academic vocabulary breadth and depth throughout the intervention, providing students with many opportunities to develop their oral and written language through engaging conversations, activities to develop morphological awareness, explanations of the grammatical and syntactic structures of word problems, and videos and illustrations (S. K. Baker et al., 2014; Lesaux, Kieffer, Kelley, & Harris, 2014; Mayer, 2009).

Thus, more research studies ought to be conducted to address the needs of this very heterogeneous population of students that is growing worldwide. Single-subject research is a cost-effective experimental approach that can potentially provide important insights into what works for English learners at-risk or with learning disabilities. The studies reviewed here indicate that, with one exception, the strategies provided to students were effective and significantly increased their literacy, mathematics, and behavior outcomes. Thus, single-subject research can be a way to further examine distinctive independent variables that might affect outcomes of interventions for English learners, such as varying language proficiency levels, language of instruction, engaging academic conversations, and including robust vocabulary instruction in all academic areas (i.e., reading, writing, mathematics, science, and social studies). Moreover, the possibility of combining SSD studies to conduct a meta-analysis and thus increase the external validity of interventions for ELs at-risk of or with a learning disability would be highly desirable. The recent focus of expert methodologists in the field on making the SSD approach more available to researchers and practitioners is a first step in better understanding and using SSDs to address the needs of an increasingly diverse population of students.

ACKNOWLEDGEMENT

We would like to acknowledge the work by the late Dr. William R. Shadish on Single Subject Design (SSD) methodologies. His contributions to the field inspired us to take a closer look at the role of SSDs to test the effectiveness of interventions intended to support English learners at risk for learning disabilities.

NOTES

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

1. Sprick, Howard, and Fidanque (1998–2000).
2. Simmons and Kame'enui (2003).

REFERENCES

Artiles, A. J., Rueda, R., Salazar, J., & Higareda, I. (2005). Within-group diversity in minority special education disproportionate representation: The case of English language learners in California's urban school districts. *Exceptional Children*, *71*, 283–300.

August, D., & Shanahan, T. (2006). *Executive summary: Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Retrieved from http://www.cal.org/resource-center/publications/developing-literacy.

Baker, D. L., Basaraba, D., & Polanco, P. S. (2016). Connecting the present to the past: Furthering the research on bilingualism and bilingual education. *Review of Research in Education*, *40*, 821–883. https://doi.org/10.3102/0091732X16660691.

Baker, D. L., Richards-Tutor, C., Gersten, R., Baker, S. K., & Smith, J. (2017). Building literacy for English learners within response to intervention. In E. C. Lopez, S. G. Nahari, and S. L. Proctor (Eds.), *The handbook of multicultural school psychology. An interdisciplinary perspective* (2nd ed.) (pp. 97–117). New York, NY: Routledge.

Baker, S. K., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., et al. (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.

Cook, B., & Cook, L. (2016). Research designs and special education research: Different designs address different questions. *Learning Disabilities Research and Practice*, *31*(4), 190–198. https://doi.org/10.1111/ldrp.12110.

Every Student Succeeds Act (ESSA) (2015). *Every Student Succeeds Act of 2015, Pub. L. No 114–95 § 114 Stat. 1177 (2015-2016)*. Retrieved from the US Department of Education website https://www.ed.gov/essa?src=rn

Gast, D. L., Lloyd, B. P., & Ledford, J. R. (2014). Multiple baseline and multiple probe designs. In D. L. Gast & J. R. Ledford, *Single case research methodology: Applications in special education and behavioral sciences* (pp. 251–296). New York, NY: Routledge.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*(2), 149–164.

Goldenberg, C., Reese, L., & Rezaei, A. R. (2011). Contexts for language and literacy development among dual-language learners. In A. Yücesan Durgunoğlu & C. Goldenberg (Eds.), *Language and literacy development in bilingual settings* (pp. 3–28). New York, NY: The Guilford Press.

Good, R. H., & Kaminski, R. (2002). *Dynamic indicators of basic early literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Education Achievement.

*Gyovai, L. K., Cartledge, G., Kourea, L., Yurick, A., & Gibson, L. (2009). Early reading intervention: Responding to the learning needs of young at-risk English language learners. *Learning Disability Quarterly*, *32*, 143–162.

Hedges, L. V., Pustejovsky, J., & Shadish, W. R. (2012). A standardized mean difference effect size for single-case designs. *Research Synthesis Methods*, *3*, 224–239.

Hedges, L. V., Pustejovsky, J., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods*, *4*, 324–341. https://doi.org/10.1002/jrsm.1086.

Horner, R. D., & Baer, D. M. (1978). Multipleprobe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis*, *11*(1), 189–196.

Horner, R. D., Carr, E., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*(2), 165–179.

Kazdin, A. (2010). *Single-case research designs: Methods for clinical and applied settings*. New York, NY: Oxford University Press.

Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., et al. (2015). *The Condition of Education 2015* (NCES 2015–144). Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., et al. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Lesaux, N. K., Kieffer, M. J., Kelley, J. G., & Harris, J. R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, *51*(6), 1159–1194. https://doi.org/10.3102/0002831214532165.

Marso, D., & Shadish, W. R. (2015). *DHPS Effect Size Macro [Computer Software]*. Merced, CA: University of California, Merced.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge.

Neuman, S. B. (2011). Single-subject experimental design. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methodologies* (pp. 383–402). New York, NY: Guilford Press.

*Orosco, M. J. (2013). The development of a math comprehension strategy in Spanish for Latino English language learners at risk for math disabilities. *International Journal for Research in Learning Disability*, *1*(2), 86–108.

*Orosco, M. J. (2014a). A math intervention for third grade Latino English language learners at risk for math disabilities. *Exceptionality: A Special Education Journal*, *22*(4), 205–225. https://doi.org/10.1080/09362835.2013.86535.

*Orosco, M. J. (2014b). Word problem strategy for Latino English Language Learners at risk for math disabilities. *Learning Disability Quarterly*, *37*(1), 45–53.

*Orosco, M. J., Swanson, H. L., O'Connor, R., & Lussier, C. (2011). The effects of dynamic strategic math on English language learners' word problem solving. *The Journal of Special Education*, *47*(2), 96–107.

*Preciado, J. A., Horner, R. H., & Baker, S. K. (2008). Using a function-based approach to decrease problem behaviors and increase academic engagement for Latino English language learners. *The Journal of Special Education*, *42*, 227–240. https://doi.org/10.1177/0022466907313350.

*Rahn, N. L., Wilson, J., Egan, A., Brandes, D., Kunkel, A., Peterson, M. et al. (2015). Using incremental rehearsal to teach letter sounds to English language learners. *Education and Treatment of Children*, *38*(1), 71–91.

Richards-Tutor, C., Baker, D. L., Gersten, R., Baker, S. K., & Smith, J. M. (2015). The effectiveness of reading interventions for English learners: A research synthesis. *Exceptional Children*, 1–26. https://doi.org/10.1177/0014402915585483.

*Ross, S. G., & Begeny, J. C. (2011). Improving Latino, English language learners' reading fluency: The effects of small group and one-on-one intervention. *Psychology in the Schools*, *48*(6), 604–618.

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*, 148–162.

*Santoro, L. E., Jitendra, A. K., Starosta, K., & Sacks, G. (2006). Reading well with Read Well enhancing the reading performance of English language learners. *Remedial and Special Education*, *27*(2), 105–115.

Simmons, D. C., & Kame'enui, E. J. (2003). *Scott Foresman early reading intervention*. Glenview, IL: Scott Foresman.

Shadish, W. R., Hedges, L. V., Horner, R. H., & Odom, S. L. (2015). *The role of between-case effect size in conducting, interpreting, and summarizing Single-Case Research (NCER 2015-002*). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Retrieved from Institute website http://ies.ed.gov/.

Shadish, W. R., Hedges, L. V., & Pustejovsky, J. E. (2014). Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: A primer and applications. *Journal of School Psychology*, *52*(2), 123–147.

Sprick, M., Howard, L. M., & Fidanque, A. (1998–2000). *Read well: Critical foundations in primary reading*. Longmont, CO: Sopris West.

Springer, K. (2010). *Educational research: A contextual approach*. Hoboken, NJ: Wiley.

Tong, F., Lara-Alecio, R., Irby, B., Mathes, P., & Kwok, O. (2008). Accelerating early academic oral English development in transitional bilingual and structured English immersion programs. *American Educational Research Journal*, *45*, 1011–1044. https://doi.org/10.3102/0002831208320790.

Vaughn, S., Linan-Thompson, S., Mathes, P. G., Cirino, P. T., Carlson, C. D., Pollard-Durodola, S. D., et al. (2006). Effectiveness of Spanish intervention for first-grade English language learners at risk for reading difficulties. *Journal of Learning Disabilities*, *39*, 56–73.

*Viel-Ruma, K., Houchins, D. E., Jolivette, K., Fredrick, L. D., & Gama, R. (2010). Direct instruction in written expression: The effects on English speakers and English language learners with disabilities. *Learning Disabilities Research & Practice*, *25*(2), 97–108.

About the Authors

Doris Luft Baker is an Assistant Professor in the Simmons School of Education and Human Development at Southern Methodist University. She earned her doctorate in Educational Leadership from the University of Oregon. Her current research interests are the development and testing of interventions and formative assessments that help improve the academic outcomes of English learners and native Spanish speakers with and without learning disabilities using evidence-based practices and technology.

Catherine (Cara) Richards-Tutor, PhD, is a professor in the Special Education Program and California State University, Long Beach. She earned her doctorate from the University of California, Santa Barbara. Her research interests include interventions and assessments for English learners at-risk and with learning disabilities and preparing teachers to use these evidence-based interventions and assessments.

Anthony Sparks is a doctoral student at Southern Methodist University. Previously, he received his Master's in Education from the University of North Carolina at Greensboro. His current research interests include the effects of policies on English learner academic outcomes.

Rebecca Canges is an Associate Professor in the School of Education at Metropolitan State University, Denver. She earned her doctorate in Educational Leadership from the University of Southern California. Dr. Canges' research interests are in the areas of effective instruction for English language learners and students with disabilities, teacher preparation, and methods for establishing social acceptance for students with disabilities in schools.