

# Social Validity in Early Language Interventions for Dual Language Learners: A Systematic Review of the Literature

Topics in Early Childhood Special Education  
2020, Vol. 40(1) 39–51  
© Hammill Institute on Disabilities 2020  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/0271121419901289  
tecse.sagepub.com



Anne L. Larson, PhD<sup>1</sup> , Zhe Gigi An, PhD<sup>2</sup>, Carla Wood, PhD<sup>3</sup>,  
Yuuko Uchikoshi, PhD<sup>4</sup>, Lauren M. Cyclic, PhD<sup>5</sup>, Carol Scheffner Hammer, PhD<sup>6</sup>,  
Kelly Escobar, PhD<sup>6</sup>, and Kate Roberts, MA<sup>2</sup>

## Abstract

The social validity of intervention research has been emphasized in special education and related fields for decades. There is relatively little focus on social validity that considers culturally and linguistically diverse populations. Eleven articles met the inclusionary criteria for this systematic review and were evaluated to describe social validity in early language intervention research—specifically with young (birth through 5 years) dual language learners (DLLs). We examined how social validity was considered and addressed, how social validity was measured, and the results of social validity assessments. Definitions of social validity were generally vague, and researchers in most studies only considered social validity post intervention rather than prior to starting or during the interventions. Participants generally viewed interventions positively, but incomplete data were often reported.

## Keywords

language/communication, development, cultural diversity, families, intervention strategies, focus on measurement/instrument/test design, research methodologies

The number of children in the United States who are considered dual language learners (DLLs) is substantial and growing. DLLs currently comprise approximately one-third of children below the age of five, indicating 24% growth in this population since 2000 (Park et al., 2017). DLLs are defined in this study as children who are learning two or more languages from an early age as a result of their home exposure to a language other than (or in addition to) English. Although DLLs represent many racial, ethnic, and linguistic backgrounds, most DLLs in the United States are of Hispanic or Latino origin and exposed to Spanish (Park et al., 2017). DLLs also experience varied socioeconomic circumstances, but a majority (58%) of them live in families with low incomes below 200% of the federal poverty level (Park et al., 2017). Limited data exist to specify the number of DLLs who qualify for early intervention (EI) or early childhood special education (ECSE); however, Head Start reported in 2006 that approximately 8% of enrolled DLLs had an identified disability—most of whom qualified for intervention services due to speech and/or language delay (U.S. Department of Health and Human Services, 2013).

Given these demographic trends, a focus on language development is needed to support DLLs with and without

disabilities to prepare for school (e.g., Priest et al., 2001). DLLs face unique complexities related to preparing for schooling in a language other than, or in addition to, the home language(s). Often related to limited experience with English and socioeconomic factors, some DLLs enter school with less robust English language skills than their peers from monolingual English-speaking backgrounds (Oller & Eilers, 2002; Place & Hoff, 2011). Unfortunately, this performance discrepancy tends to widen for DLLs over time (e.g., Moiduddin et al., 2012; Reardon & Galindo, 2009) and further challenges DLLs' future academic and social development due to the predictive relations between early language skills and other areas of development (Morgan et al., 2015).

<sup>1</sup>Utah State University, Logan, USA

<sup>2</sup>University of Wisconsin–Madison, USA

<sup>3</sup>Florida State University, Tallahassee, USA

<sup>4</sup>University of California, Davis, USA

<sup>5</sup>University of Oregon, Eugene, USA

<sup>6</sup>Teachers College, Columbia University, New York, NY, USA

## Corresponding Author:

Anne L. Larson, Department of Special Education and Rehabilitation,  
Utah State University, 2865 Old Main Hill, Logan, UT 84322-2865, USA.  
Email: anne.larson@usu.edu

Effective language interventions aimed at addressing the unique needs of DLLs and their caregivers may help address concerns for DLLs with or at risk for delays. The likelihood that an intervention is successfully incorporated into practice often depends on whether the intervention addresses goals that are important to consumers, contains procedures that are practical and acceptable, and leads to outcomes that are considered meaningful (Leko, 2014; Lloyd & Heubusch, 1996; Strain et al., 2012). Consumers who feel that interventions align with their experiences, perspectives, and values are more inclined to participate in the intervention and, consequently, to experience positive outcomes (Dunst et al., 2016; Strain et al., 2012; Wolf, 1978). As such, ensuring the social validity of early language interventions for DLLs (alongside their effectiveness) may support the language development of this group.

### Early Language Interventions for DLLs

Two recent systematic reviews have provided important information about the effectiveness of early language interventions for DLLs with and without disabilities; however, neither review provided an in-depth exploration of social validity. Durán et al. (2016) reviewed 26 language intervention studies for young (below age six) DLLs with, or at risk for, language difficulties. Although the authors noted variability in the quality (i.e., rigor) of the interventions reviewed, they concluded that bilingual preschool instruction, supplemental language and literacy instruction, individualized vocabulary instruction, and parent book reading interventions have demonstrated positive effects on children's English and home language abilities. Larson and colleagues (2020) reached similar conclusions in a review of 41 language intervention studies that addressed children who were considered culturally and linguistically diverse, including DLLs. Larson and colleagues (2020) found that interventions for this population are more likely to be adapted on the basis of home language than culture, and that interventions that are both linguistically and culturally adapted appear to best support language outcomes for children from diverse backgrounds. Moreover, fewer than a third of the reviewed studies measured social validity; however, the discussion on social validity within the review was limited.

### Social Validity and Its Application for DLLs

Social validity has been discussed in the field of applied behavior analysis to elevate the experiences, perspectives, and values of intervention consumers. As defined by Wolf (1978), social validity includes three main components: acceptance of the intervention goals ("Are the specific behavioral goals really what society wants?"), procedures

("Do the participants, caregivers and other consumers consider the treatment procedures acceptable?"), and outcomes/effects ("Are consumers satisfied with the results?") (p. 207). Since this definition, social validity has been conceptualized in a variety of ways across the social sciences, including but not limited to *social importance*, *clinical significance*, *ecological validity*, *cultural validity*, *cultural adaptation*, and *cultural significance* (Carter & Wheeler, 2019). Despite various terms for, and conceptualizations of, social validity, the operationalization of the concept (particularly for interventions with DLLs) remains unclear and vague, and its inclusion and consideration are limited (Hurley, 2012; Snodgrass et al., 2018).

A growing body of evidence suggests the importance of incorporating social validity when developing interventions for populations from diverse cultural and linguistic backgrounds (e.g., Castro et al., 2010; Parra Cardona et al., 2012). More specifically, some researchers suggest that social validity is achieved through "cultural adaptation"—adjusting previously developed interventions to the cultural and linguistic characteristics of the target population (Mejia et al., 2017). Intervention adaptation and design is especially critical to consider for early language interventions for DLLs where much of the current literature is based on White, English-speaking, middle-class caregivers and their children (Roberts & Kaiser, 2015).

The three components of social validity outlined by Wolf (1978) provide a starting point to guide researchers and practitioners in developing and delivering language interventions that best meet the needs of DLLs; however, few researchers report using social validity to adapt interventions prior to or during implementation, and, instead, most researchers only apply social validity evaluations at the end of an intervention. For example, in a review of special education articles that reported on social validity, Snodgrass and colleagues (2018) found that nearly all (96%) assessments of social validity occurred only after the intervention. Solely measuring social validity at the conclusion of an intervention limits the potential impact of results to future studies and intervention development. In contrast, measuring social validity before or during the intervention may allow for cultural adaptations that increase intervention acceptability. Thus, there is a need to determine when social validity is being considered in interventions with DLLs as a means to design and adapt socially accepted interventions.

### Previous Reviews of Social Validity

Several recent reviews have evaluated the use of social validity in intervention studies for children with, or at risk for, disabilities, but none have specifically focused on DLLs. Snodgrass et al. (2018), for example, reviewed 429 articles addressing social validity within single-case research studies in special education journals between 2005 and 2016, and

Ledford et al. (2016) examined evidence of social validity more specifically in social skill interventions for young children with autism spectrum disorders (from 1994 to 2013). Neither of these reviews commented on the cultural or linguistic characteristics of participants included in the studies, nor how social validity was considered based on participants' background. Spear and colleagues (2013) summarized participant characteristics (participant racial and ethnic background, but not linguistic background) in a review of 22 single-case design studies (from 2008 to 2011) focused on children with or at risk for emotional behavioral disorders; however, they did not relate these characteristics to social validity specifically.

The proportion of studies that have included social validity has varied among reported reviews. Spear and colleagues (2013) reported 68% of studies including some assessment of social validity, whereas Snodgrass et al. (2018) found that only 26.8% of studies reported measurement of social validity data. Ledford and colleagues' (2016) review demonstrated that fewer than half of their included studies measured social validity. When researchers addressed social validity, results of recent reviews noted that very few (e.g., 6.8% of the studies reviewed by Snodgrass and colleagues, 2018) addressed all three aspects of social validity as originally conceptualized by Wolf (1978). In addition, no researchers reported on intervention consumers' evaluation of intervention procedures or components (Ledford et al., 2016). Results from these reviews, along with other research (e.g., Hurley, 2012; Odom & Strain, 2002), reveal the underutilization of social validity, as well as a lack of consideration for the content of an intervention within research for children with and/or at risk for disabilities.

Moreover, within studies that include social validity, there is great variation in the ways in which social validity is assessed. Some researchers (e.g., Leko, 2014) have emphasized the value of qualitative methods in social validity assessments, including more subjective social validity assessment methods such as interviews, focus groups, surveys, and rating scales. Yet, other researchers (e.g., Ledford et al., 2016) have argued for the need for more objective and psychometrically sound social validity assessments such as normative comparisons and measures of maintenance or sustained use of an intervention (see Barton, Meadan-Kaplansky, & Ledford, 2018, for a brief review). Considerations for the measurement of social validity in interventions with culturally and linguistically diverse populations may be especially critical as these groups may benefit from different procedures to assess social validity based on language (e.g., use of interpreters) and cultural preference (e.g., involving a peer assessor rather than a researcher).

This study expands on earlier reviews by focusing specifically on social validity within early language interventions for DLLs. Additional research is needed to understand

how social validity is conceptualized and measured in intervention studies involving DLLs to advance the field's consideration of this critical intervention component for this growing population. Similarly, it is important to analyze and consider the results of social validity assessment in DLLs to support future development of high-quality, culturally and linguistically responsive interventions. An improved understanding of the components associated with socially valid interventions for DLLs could inform design and implementation of interventions that promote participant engagement and child language outcomes.

We applied a conceptual framework that highlighted attention to cultural and linguistic diversity on the three aspects of social validity that Wolf (1978) defined. That is, we examine the extent to which researchers address social validity of the intervention goals (e.g., whether caregivers of DLLs believe that developing their child's English skills and/or skills in languages other than English is important), the intervention procedures (e.g., whether a play-based approach to facilitating language matches the cultural perspectives of the family), and the intervention effects/outcomes (e.g., whether caregivers of DLLs noted any meaningful changes in their child's dual language development as a result of the intervention). We also propose a fourth component of social validity that expands on acceptance of intervention procedures and is focused on intervention *content* (e.g., themes represented in intervention materials such as books). In addition to looking at intervention procedures by gathering stakeholder opinions on *how* the intervention is delivered and whether it is feasible and ethical for a particular population (Barton et al., 2018), a focus on content helps ensure that the subject matter of an intervention is responsive to culturally and linguistically diverse populations.

We frame the social validity of an intervention's content as the extent to which an intervention incorporates the cultural and linguistic background of the stakeholders in its adaptation or design. Our definition of content considers *what* the intervention includes. Acceptability of an intervention's content might ask about the cultural and/or linguistic sensitivity of the content or materials and the specific strategies used by intervention agents prior to or after asking participants to use the strategy or materials within the intervention procedures.

To understand how social validity has been addressed in interventions for young DLLs, we focused on three questions: (a) To what extent is social validity considered in early language interventions with DLLs? (b) How is social validity being measured within early language interventions with DLLs? (c) What are the results of social validity assessment in early language interventions with DLLs? With the results to these questions, we discuss implications and further directions for research and practice.

## Method

### Search Strategy

We followed guidelines from Preferred Reporting Items for Systematic Reviews (PRISMA; Liberati et al., 2009). We conducted a database search from 1975 to 2015 as reported in Larson et al. (2020), then used the following terms to update the search through 2018 across five databases (Web of Science, PsycINFO, ERIC, MEDLINE, and Academic Search Ultimate): (language OR vocabulary OR communication OR talk OR interact\* OR word\* OR gesture\*) AND (interven\* OR strateg\* OR training OR program OR development) AND (infant\* OR toddler\* OR preschool\* OR child\*) AND (parent\* OR mother\* OR father\* OR caregiver\* OR teach\* OR adult\* OR “child care” OR childcare) AND (bilingual OR DLL OR “dual language learner\*” OR “English language learner\*” OR ELL OR “English learner”).

### Inclusionary Criteria

Studies were included in this synthesis if they were written in English in a peer-reviewed journal, were conducted in the United States, and used either a single-case or group design. Single-case research design studies had at least three opportunities to demonstrate the experimental effect (i.e., three participants in a multiple baseline design across participants), and group design studies included a control or comparison. All included studies also had at least some participants who spoke a language other than English (LOTE) and involved children birth through kindergarten with an average age of no more than 5 years 11 months (or adult participants who focused on this age group). In addition, researchers in included studies reported on interventions designed to increase language skills (e.g., receptive or expressive vocabulary, syntax, etc.) for children who are DLLs. These criteria included studies directed at influencing caregiver behavior associated with child language outcomes, such as caregivers’ use of responsive communication. Researchers of included studies also considered participant culture or linguistic diversity in (a) intervention design (e.g., materials linked to children’s personal cultural experiences, translation of an existing intervention, etc.), (b) measurement and outcomes (e.g., assessment in both languages), or (c) analysis (e.g., factoring initial child dual language ability or caregiver cultural preferences into interpretation of study data). Finally, to be included in the review, each study addressed or measured at least one component of social validity (i.e., acceptance of the intervention goals, procedures, content, or outcomes and effects as defined previously).

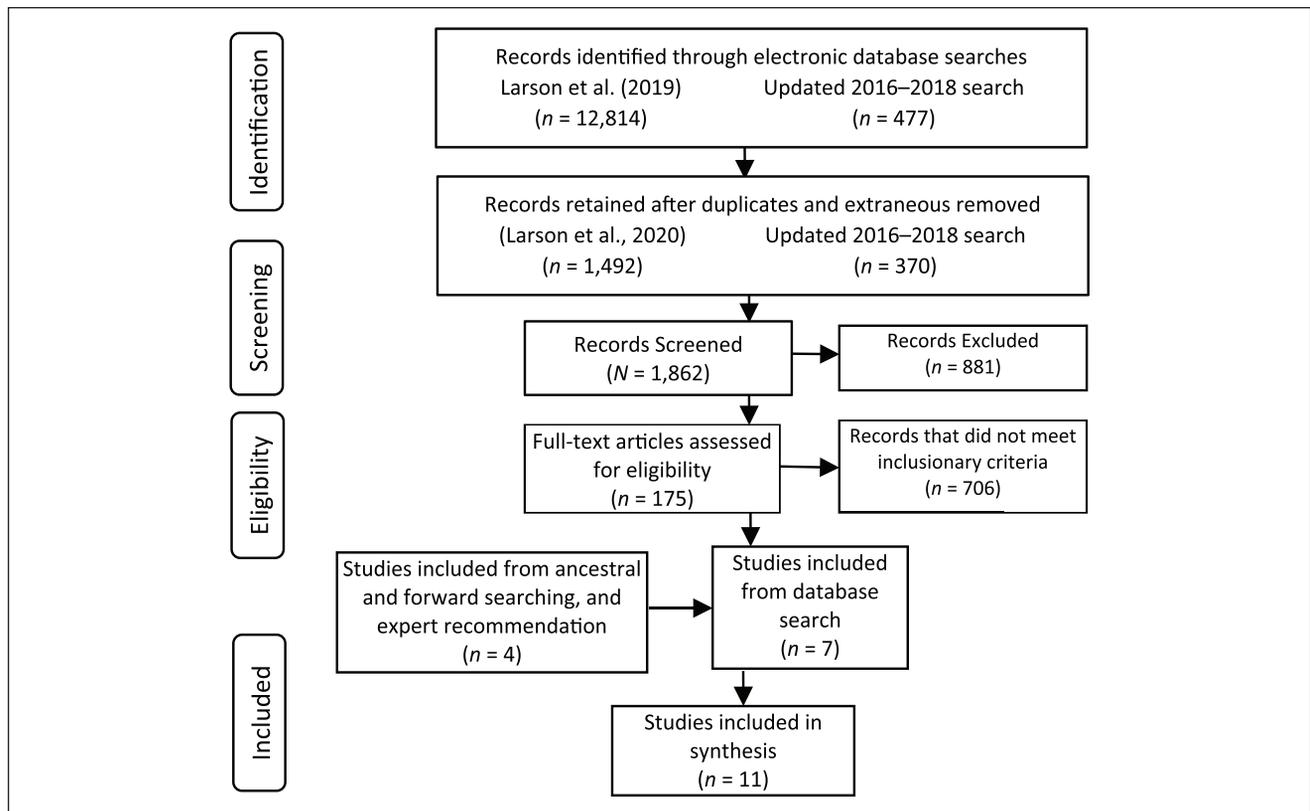
### Study Selection Process

To determine study inclusion in this synthesis, database-identified articles were title and abstract screened by individual reviewers and then full-text screened (when

inclusionary criteria were unclear based on information from title and abstract alone). The first and second authors independently considered all studies ( $n = 65$ ) during the full-text screening. Interrater agreement on study inclusion decisions between the first and second authors was 97% (63 out of 65 studies). Consensus on disagreements was reached with all coauthors. We also used snowball searching to conduct a hand search of reference lists and forward search (using the Google Scholar “cited by” feature) from database-selected studies meeting inclusion criteria. These methods resulted in two additional studies that met inclusionary criteria. All authors also came to consensus on the inclusion of two additional studies suggested by the authors of this study.

### Coding and Analyses

To address the first research question, articles were coded to determine whether study authors listed or operationalized social validity (or a related term) and where in the article the study authors mentioned social validity (e.g., research questions, discussion, etc.). During this process, coders identified instances where authors documented addressing cultural and linguistic relevance, as was done in Peredo et al. (2017), where participants were asked how to change or improve the intervention content to “be more congruent with Hispanic cultural and parenting values” (p. 3). Coders also identified the specific aspects of social validity that were addressed and/or measured within each study with a particular focus on the cultural and linguistic relevance of an intervention for the target population, such as *goals and objectives* (related to the perceived importance of the intervention targets), *procedures* (perceptions of the appropriateness of steps involved in the intervention and its feasibility), *content* (acceptance of the materials used in the intervention), and *effects* (perceived importance of the intervention outcomes). Coders also noted when social validity was taken into consideration (i.e., before, during, and/or after the intervention period). To answer the second research question, coders noted when components of social validity were measured, documented information about the format and process used for measurement (e.g., questionnaires, focus groups, etc.), the tool used to gather social validity data (i.e., researcher-developed measures), the administrator of social validity measures, and the specific questions or prompts that were used to gather social validity data. Coders also noted whether any cultural or linguistic adaptations were used to gather social validity data, such as translating questionnaires, using an interpreter, and using nonresearchers as data collectors to avoid any perceived power relations that may have been exacerbated by differences in cultural background between researchers and participants. To answer the third research question, coders noted detailed information on reported social validity outcomes (e.g., whether qualitative and/or quantitative reporting was used, if social validity findings



**Figure 1.** Study selection process.

were positive, and whether findings were supported by data).

All articles were double-coded by two authors, and coding reliability was calculated by summing the number of agreements across studies ( $n = 533$ ), then dividing the total number of possible agreements across studies ( $n = 660$ ), and multiplying by 100. Overall reliability was 81% (range = 72%–90%). Coders had strong agreement across the vast majority of codes; however, disagreement occurred on some individual items such as whether authors reported social validity on intervention effects and the specific category for the reported level of social validity (very high/moderate/low and positive/negative/mixed). To account for disagreements, all final codes were determined by consensus with the coders and the first author. All final codes were entered in REDCap (Research Electronic Data Capture)—a web-based program designed to support data collection and automatic data exports (Harris et al., 2009, 2019). Quantitative data (e.g., number of studies that addressed social validity before implementing the intervention) were summarized with descriptive statistics such as frequencies and proportions. Descriptive information (e.g., author-provided definitions of social validity) was independently analyzed by two authors who reviewed all coder comments and identified common themes (Patton, 2015).

## Results

Figure 1 shows a flow diagram of the search process. The search for this review resulted in 1,862 articles after accounting for duplicates across databases. Of these, 175 articles remained after title and abstract screening. Seven articles remained after full-text screening. With ancestral, forward searching and author-recommended studies, 11 studies met inclusionary criteria for this review. Table 1 includes detailed information about each study. Parents were the intervention agents in all but three studies in this review (Bernhard et al., 2006, 2008; Simon-Cerejido & Gutiérrez-Clellen, 2014). Other intervention agents ( $n = 3$ ) were early educators.

### Social Validity Considerations

To address the first research question, we examined how social validity was defined, framed, and addressed within each study. Across studies, no definition of social validity was explicitly given, and researchers in only three studies specifically identified the construct using the term *social validity* (Cooke et al., 2009; Hammer & Sawyer, 2016; Peredo et al., 2017). Researchers addressed and evaluated the following constructs relevant to Wolf's (1978) definition

**Table 1.** Characteristics of Identified Studies on Social Validity That Included Culturally and Linguistically Diverse Populations.

Study	Intervention	Child participants ( <i>n</i> ; <i>M</i> or range)	Setting	Language(s) represented
Bernhard et al. (2006)	<i>Early Authors Program (EAP)</i>	367; 37.3 months	Early learning program	English, Haitian Creole
Bernhard et al. (2008)	EAP	367; 37.3 months	Early learning program	English, Haitian Creole
Boyce et al. (2010)	<i>Storytelling for the Home Enrichment of Language and Literacy Skills intervention (SHELLS)</i>	68; 41 months	Home	Spanish, Other
Caesar and Nelson (2014)	<i>Supporting the Acquisition of Language and Literacy through School-Home Activities</i>	19; 48.6 months (experimental), 46.4 months (control)	Home and early learning program	Spanish, Mixteco
Cooke et al. (2009)	Parent tutoring with the assistance of audio prompting	3; 4–5 years	Home and early learning program	Spanish <sup>a</sup>
Hammer and Sawyer (2016)	Culturally responsive interactive book-reading	73; 56.3 months	Home	Spanish
Ijalba (2015)	Parent-implemented language and literacy intervention	24; 43 months	Home	Spanish
Lim and Cole (2002)	Parent-implemented picture book reading	21; 3.3 years	Home	Korean
Peredo et al. (2017)	<i>Enhanced Milieu Teaching (EMT) en Español</i>	3; 35.7 months	Home	Spanish
Roberts (2008)	Storybook reading in primary language or English	33; 52.1 months	Home and early learning program	English, Hmong, Spanish
Simon-Cerejido and Gutiérrez-Clellen (2014)	<i>The Vocabulary, Oral Language and Academic Readiness (VOLAR) program</i>	107; 53 months	Early learning program	Spanish

<sup>a</sup>Parents reported to be from “Spanish-speaking countries”; language of participants not specified.

of social validity, including *acceptance, enjoyment, satisfaction, perceived effectiveness, feasibility, and sustainability*. Several researchers also used terms that closely aligned with our conceptualization of social validity with an emphasis on cultural and linguistic diversity. These terms included *cultural appropriateness and cultural acceptance* (Peredo et al., 2017), *cultural validity* (Hammer & Sawyer, 2016), *cultural relevance* (Boyce et al., 2010; Caesar & Nelson, 2014), *ecological validity* (Lim & Cole, 2002), and *cultural validation* (Bernhard et al., 2008).

We also determined where the concept of social validity appeared in each article (see Table 2) as a means to evaluate the extent to which study authors addressed social validity within a scientific approach (Creswell, 2008; Snodgrass et al., 2018). Researchers in three studies discussed social validity (or a related term) in the introduction section; researchers in five studies included a research question related to social validity; and researchers in all but one study (Simon-Cerejido & Gutiérrez-Clellen, 2014) included some mention of social validity in the methods section. Researchers in most studies ( $n = 9$ ; 82%) addressed the concept within the results and discussion sections. Hammer and Sawyer (2016) were the only ones to address social validity across all major components of the research paper.

Table 2 provides a summary of the aspects of social validity (i.e., goals, procedures, content, and outcomes/

effects) that were addressed and/or measured in each included study. Researchers from three studies targeted the *goals and objectives* aspect of social validity. For example, Roberts (2008) asked the caregivers two questions about goals and objectives: (a) “How important is it for <child name> to keep his/her ability to speak Spanish/Hmong?” and (b) “How important is it for <child name> to learn English?” (p. 128). Researchers in five of the studies examined social validity in terms of intervention *content*. For instance, Peredo et al. (2017) adapted a language intervention by having Spanish–English bilingual providers with extensive experience working with Spanish-speaking families view videos of the intervention strategies and provide feedback on the cultural appropriateness of strategies to determine how to better align the intervention with Latino cultural and parenting values. Researchers across a majority of studies addressed the *procedure* aspect of social validity ( $n = 10$ ; 91%)—the practicality of receiving or implementing the intervention. Lim and Cole (2002) measured the social validity of the intervention procedure at the end of the study; the authors asked the participants whether any of the techniques were difficult to use due to cultural differences between Korean and American cultures. Researchers in eight studies examined the social validity of intervention *outcomes* or *effects*. Bernhard et al. (2006, 2008) used an exit interview to understand the

**Table 2.** Social Validity Characteristics of Identified Studies.

Study	Where did social validity appear in the article?					What aspect was addressed?			
	Intro	RQs	Method	Result	Discussion	Goal	Content	Procedure	Effects
Bernhard et al. (2006)			✓	✓					✓
Bernhard et al. (2008)		✓	✓	✓	✓		✓	✓	✓
Boyce et al. (2010)		✓	✓	✓	✓		✓	✓	✓
Caesar and Nelson (2014)	✓		✓	✓	✓			✓	
Cooke et al. (2009)		✓	✓	✓	✓	✓		✓	✓
Hammer and Sawyer (2016)	✓	✓	✓	✓	✓	✓	✓	✓	
Ijalba (2015)			✓	✓			✓	✓	✓
Lim and Cole (2002)			✓	✓	✓		✓	✓	
Peredo et al. (2017)	✓	✓	✓		✓		✓	✓	✓
Roberts (2008)			✓	✓	✓	✓	✓	✓	✓
Simon-Cerejido and Gutiérrez-Clellen (2014)					✓			✓	✓

Note. RQs = research questions.

participants' perceived effectiveness of the intervention on child outcomes.

### Social Validity Measurement

The most common time to gather data on social validity within language interventions for DLLs was at the end of the study ( $n = 10$ ; 91%). Ijalba (2015) and Peredo et al. (2017) gathered social validity data both before and after the intervention. For example, to build interactive picture books, Ijalba (2015) interviewed mothers about objects, activities, and themes influencing their communication with their children prior to the intervention. Post intervention, mothers were also asked about their satisfaction with the intervention. Roberts (2008) gathered social validity survey data during and after the intervention to understand family caregivers' views regarding the perceived importance of their home language and English, English acquisition, and primary language maintenance.

Simon-Cerejido and Gutiérrez-Clellen (2014) assessed social validity only during the intervention by encouraging participating teachers to provide feedback about any aspect of the intervention. Hammer and Sawyer (2016) considered social validity throughout the intervention at multiple time points—during intervention design (i.e., prior to implementing the study), throughout the implementation of the intervention, and at the end of the intervention. In their study, Latina mothers (separate from participants in the study) worked through an iterative process with a book author to determine appropriate cultural content of books to be used during the book reading intervention. Throughout the intervention, mothers who read the books to their children also provided data on whether they read the books, used the interactive strategies, or had suggestions to improve the intervention. At the end of the intervention,

participants were asked an open-ended question about their participation in the intervention.

Researchers in nine studies (81%) collected information about social validity from the adult recipients of the intervention. In seven of these studies, researchers elicited inputs from caregivers, specifically parents. Mothers were the only caregiver included in the assessment of social validity in five of the seven studies. Teachers (who were both the intervention agents and recipients) reported on their satisfaction with and perceived sustainability of the intervention in Bernhard et al. (2008). In two other studies, literacy specialists who provided support to the teachers implementing the Early Authors Program (Bernhard et al., 2006) and preschool teachers who implemented the Vocabulary, Oral Language, and Academic Readiness curriculum (Simon-Cerejido & Gutiérrez-Clellen, 2014) provided information on social validity. Not all recipients or agents of the intervention in each study provided information on social validity. Researchers in only five studies reported that social validity information was collected from all adult recipients (Cooke et al., 2009; Hammer & Sawyer, 2016; Ijalba, 2015; Lim & Cole, 2002; Peredo et al. 2017). In the remaining studies where researchers collected social validity data from adults (Bernhard et al., 2008; Boyce et al. 2010; Caesar & Nelson, 2014; Roberts, 2008), it was unclear whether all adult participants or only a portion of the adult participants provided information for social validity measurements.

Researchers used researcher-developed subjective measures of social validity in all identified studies. Researchers in the majority of the studies ( $n = 10$ ; 91%) measured social validity through interviews, surveys, or a combination. For instance, Boyce et al. (2010) assessed parent satisfaction with the intervention by asking families how the intervention supported their family values and customs. Caesar and

Nelson (2014) conducted a focus group 2 weeks after the conclusion of the intervention where parents provided feedback on the feasibility of the intervention procedures.

Researchers in six studies (55%) explicitly described the questions and/or items that were used to gather social validity information. Researchers in three studies inquired whether there were any cultural issues with the intervention (Hammer & Sawyer, 2016; Lim & Cole, 2002; Peredo et al. 2017). For example, Lim and Cole (2002) asked about cultural differences between Korean and American cultures, and whether there were any difficulties while reading picture books to children using the specific techniques from the study. Roberts (2008) asked about the parents' language preference in conducting the intervention. Bernhard and colleagues (2008) and Ijalba (2005) collected information about the participants' overall engagement and satisfaction with the intervention.

Researchers in eight studies made cultural and/or linguistic adaptations to social validity assessments. Researchers in three studies used interpreters, three used translated materials, and two included both interpreters and cultural liaisons (i.e., bilingual staff from the community; Hammer & Sawyer, 2016; Roberts, 2008). Social validity assessments were typically administered by researchers and their assistants ( $n = 9$ ; 82%). Roberts (2008), however, asked parents of the children in the preschool to collect social validity data (after being trained in survey administration) based on the belief that having a cultural and linguistic match between survey administrators and participants would increase the caregivers' comfort and thus increase participant elaborations of their responses. In one other case, the role of the social validity assessor could not be determined.

### Social Validity Results

Researchers in all studies except Bernhard et al. (2008) ( $n = 10$ ; 91%) reported social validity results using summary statements such as,

Because of the integration of the Latino culture throughout the intervention, we hypothesized that the mothers would find the intervention socially and culturally valid . . . mothers expressed pleasure with the books, indicating the books allowed them to share their culture with their children. (Hammer & Sawyer, 2016, p. 74)

Researchers in five studies reported summary statements (46%) with quantitative descriptions. For example, Boyce et al. (2010) reported the range, mean, and standard deviation of parent ratings of the intervention. Researchers in nine studies included qualitative data such as caregivers' responses to the question, "What did you like about the books coming home from school for reading with your child?" (Roberts, 2008, p. 130). Author-reported summary

statements related to the overall level of social validity among participants who completed social validity assessments were highly positive across studies; however, these findings were supported by limited evidence. For example, not all recipients or agents of the intervention in each study participated in social validity assessments, and only Ijalba (2015) and Peredo et al. (2017) reported on all data relevant to social validity measurements and results. Other researchers reported only partial results. For instance, Bernhard et al. (2008) indicated that social validity was understood by interviewing both teachers and literacy specialists, yet study authors only reported on results from literacy specialist interviews.

### Discussion

Social validity is widely recognized as an important component of intervention research (e.g., Cook et al., 2013; Horner et al., 2005) and is especially important in interventions with DLLs and others from culturally and linguistically diverse backgrounds to ensure that interventions are appropriate for the target population. The percentage of language intervention studies for DLLs that include social validity (approximately 24%) reflects underutilization of social validity practices and is similar to findings in other reviews on this topic (e.g., Snodgrass et al., 2018; 26.8%). We analyzed 11 language intervention studies for young DLLs that included a discussion of social validity. Despite our suggested rationale for the importance of socially valid interventions for DLLs, we found variation in the extent to which social validity was considered and how it was measured. Researchers across all identified studies reported similar (positive) social validity outcomes despite the variability in measurement procedures.

### Consideration of Social Validity

No explicit definitions of social validity were provided by any of the reviewed studies. Consistent with previous reviews (Ledford et al., 2016; Snodgrass et al., 2018) and other research (Hurley, 2012; Odom & Strain, 2002; Spear et al., 2013), we identified a wide range of terms used to refer to social validity and researchers in only one study (Hammer & Sawyer, 2016) consistently addressed social validity across sections of reported research (i.e., introduction and rationale, research questions, methods, results, and discussion). Researchers across the majority of identified studies focused on social validity within the results and discussion. Acceptance of the intervention outcomes or effects at the end of a study was the most widely considered aspect of social validity, whereas the social validity of the goals, content, and procedures might be more likely to be addressed in earlier sections of a formal research paper.

## Measurement of Social Validity

**Timing.** Similar to findings by Snodgrass et al. (2018), social validity data in language intervention studies for DLLs were primarily gathered after intervention implementation. The Division for Early Childhood's (2014) Recommended Practices call for EIs that are culturally and linguistically responsive; however, there is little evidence of culturally informed interventions in the available literature. We found that researchers in three studies (Hammer & Sawyer, 2016; Ijalba, 2015; Peredo et al., 2017) reported gathering social validity data from caregivers or other stakeholders to inform the intervention content and design prior to implementing the intervention. It is important to collect social validity data before and throughout an intervention to assess participants' acceptance of the intervention. Hammer and Sawyer (2016) and Peredo et al. (2017) provided good examples of gathering information about the intervention strategies and materials with members from the target population prior to implementing the intervention. Hammer and Sawyer (2016) then followed up using home visitors to collect input on participants' perspectives of the intervention after each home visit.

**Respondents.** Social validity data were most often gathered from adult participants (primarily mothers), and not from child participants; however, a few researchers inquired about perspectives from other stakeholders. For example, in Hammer and Sawyer (2016) in, interviews with mothers included information about other family members' (e.g., fathers, grandparents) engagement with the intervention. Cultural differences in family ecosystems may influence which stakeholders are essential to include in the assessment of social validity (e.g., children, parents, extended family members, teachers, administrators). Generally, families from Latinx backgrounds (the largest group of participants from the identified studies) value the participation of all caregivers, including extended family and older siblings of the child, in supporting young children's development (Czyk & Hammer, 2018). It is also common for Latinx families to encourage a focus on family well-being (Landale et al., 2006); thus, perceptions of multiple individuals may be relevant to social validity for a given intervention with this group. Recognizing that family structure and family members' roles may vary across cultures, it was surprising that perspectives of diverse stakeholders (e.g., grandparents, siblings, other family members) were not elicited in the majority of the identified studies. The lack of assessment of social validity across a wide variety of stakeholders reflects a narrow breadth of consideration in the assessment of social validity within this body of research.

**Measurement.** Researchers across the identified studies exclusively used subjective measures of social validity, and,

on average, measures were brief with a limited array of questions or items. We recognize the value of both subjective (e.g., surveys and focus groups) and objective (e.g., observational assessment of participants' sustained use of an intervention and participants' choice in selecting a preferred intervention) measures of social validity. Subjective and objective measures may bring different perspectives and voices into social validity results (Barton et al., 2018). Unlike previous reviews of social validity; however, no researchers from the studies included in this review used validated measures of social validity, and few authors reported on the development process for the social validity assessment tools that were used. For example, there were no descriptions of systematic processes including defensibility of the type of social validity data measured, how the assessment tools were developed, or reports of interobserver agreement or triangulation. This limits our knowledge of the potential influencing factors that contribute to perceptions of socially valid interventions. Furthermore, limited use of validated measures makes it difficult to examine the quality of the social validity assessments that were used. As noted in Snodgrass et al. (2018), this finding suggests that studies generally reflect low quantity and quality of the social validity method.

Researchers across most identified studies used surveys or questionnaires to measure social validity. Snodgrass et al. (2018) and Spear et al. (2013) also noted the frequency of using these measures in their reviews of social validity assessment. Researcher-made questionnaires might not fully capture all relevant or important dimensions of social validity, including the cultural and linguistic appropriateness of the intervention content. Interactive dialogues or semi-structured interviews may serve as a viable alternative to static surveys and questionnaires, particularly when members of the target population(s) are included in reviewing and interpreting the findings (i.e., member-checking; Cho & Trent, 2006; Miles et al., 2013). Several researchers used more open-ended approaches to collect social validity data that may be particularly important in understanding stakeholders' acceptability of an intervention as related to their culture. For example, Hammer and Sawyer (2016) gathered information about mothers' "beliefs about parenting and aspects of their culture about which they wanted their children to learn" (p. 63) as researchers developed the intervention. Researchers across several identified studies acknowledge the complexities and multiple dimensions of *culture* (e.g., Banks, 2015; Phillips & Shonkoff, 2000), and the different interpretations of culture that may be considered by individuals (e.g., surface aspects of culture such as language, food, and customs as compared with deep aspects of culture like beliefs, values, and norms; Resnicow et al., 2000). Interviews and similar open-ended approaches may facilitate discussion with participants and other stakeholders that leads to rich input on the acceptability of an intervention as it relates to an individual's culture.

**Response reporting.** It is possible that journal space restrictions influenced the amount of social validity data included in the identified studies. Many researchers did not fully summarize the social validity data or individual responses; therefore, uncertainty remains about the distribution of response ratings and author summaries of social validity results. Researchers in some studies did not report using cultural or linguistic adaptations when assessing social validity (e.g., using an interpreter); this makes it difficult to determine whether respondents fully understood the questions.

### Social Validity Outcomes

Although participants' ratings of social validity were highly positive or affirmative in the included studies, it is possible that ratings may have been inflated by several potentially influencing factors, including cultural norms or expectations for relationships (particularly those with perceived experts) and question types.

**Cultural norms.** Researchers in this review were directly involved in the collection of social validity data, and all researchers reported positive social validity results. Stakeholders from various cultural groups may have viewed these professionals as *experts* and, thus, may have been less likely to speak up or initiate negative feedback about an intervention—particularly if researchers and participants were from differing cultural or linguistic backgrounds. It is also possible that individuals from certain cultures may be more reluctant to share aspects of the intervention that were unfavorable. For example, the tendency for positive responses may have been influenced by Latinx families' adherence to the traditional value of *simpatía*, prioritizing harmonious relationships and focusing on the positive aspects of experiences (Triandis et al., 1984). As such, respondents may have considered it inappropriate or disrespectful to convey negative perceptions or dissatisfactions. Training others (who may be viewed as peers) to gather social validity data may be preferable as was done in Hammer and Sawyer (2016) where home visitors were community members who had similar cultural, linguistic, and education backgrounds as the intervention participants.

**Question types.** Furthermore, the confirmatory nature of some of the questions used in the included studies may have implied an expectation of affirmative or positive responses. For example, a confirmatory question type (e.g., *Did you like the intervention?*) would have a high likelihood of generating a large proportion of affirmative responses. Such questions have been reported in the literature to generate high incidences of positive responses, particularly in EI studies where parents historically confirm high satisfaction with services and outcomes (McWilliam et al., 1998). As

suggested above, it may be preferable to use an open-ended question structure as was done in Peredo et al. (2017) where researchers asked, "How can we make this intervention better for other Hispanic families?" (p. 7). The use of open-ended questions to elicit constructive feedback may be more likely to generate varied and detailed information about stakeholders' acceptance of an intervention.

### Limitations

It should be noted that the findings in this review may not reflect the use of social validity in studies of DLLs outside the scope of language interventions. This review focused on language interventions specifically which represents a relatively narrow, but critical piece of literature due to the large percentage of DLLs with or at risk for disabilities who have difficulty with language and communication development. In addition, it is possible that social validity assessment and results may show different trends for interventions that focused on other aspects of developmental outcomes for DLLs. Similarly, all included studies were written in English, conducted in the United States, and published in peer-reviewed journals. As such, it cannot be assumed that a review of unpublished work (e.g., dissertations) or studies conducted outside the United States would yield similar findings.

The social validity codes employed in this study reflect a fairly narrow definition of social validity. Although our effort to conceptualize social validity informed this review, it should be noted that this review did not lend to a comprehensive study of other potential dimensions or manifestations of social validity, such as *engagement*. It may be beneficial for researchers in future studies to conduct a more exhaustive review of engagement and consider other potential dimensions of social validity that were beyond the scope of this study.

### Implications and Future Directions

There are several implications and suggestions for future research based on our findings. First, there is a need for researchers to have an explicit and inclusive definition of social validity (i.e., considering the cultural and linguistic appropriateness of intervention content) so that further research tools and methods may be developed that are appropriate for stakeholders from a variety of diverse populations. Similarly, there is a need for validated measurement strategies and reporting standards related to social validity. The current findings suggest that more rigorous subjective and objective (i.e., observable, less subject to bias) methods of social validity measurement might capture child and family preferences. We also acknowledge that qualitative methods might yield valuable information and have a particularly critical place in determining social validity for

DLLs because of the rich data available to examine stakeholders' perceptions of an intervention as related to their cultural and linguistic backgrounds.

Given the varying ways that social validity is conceptualized in the literature, additional research is needed to inform what (and who) is important to include in the process of social validity consideration in interventions for young DLLs. The inclusion of multiple stakeholders (i.e., parents, teachers, and extended family members) may extend the breadth of perspectives considered and allow for triangulation across groups. Furthermore, the design and implementation of culturally adapted interventions may be enhanced by consideration and measurement of social validity at multiple time points in the study (i.e., before, during, and after the intervention) and including stakeholders in the development and design of language interventions. Similarly, objective measures of participants' sustained use of an intervention after the study's conclusion is an indicator of procedural acceptability (Barton et al., 2018).

Additional research is needed to identify factors that contribute to family's perceptions of socially, culturally, and linguistically valid practices. We recommend researchers explore factors that may promote systems change in prioritizing the inclusion of social validity in the design and development of interventions and the inclusion of detailed descriptions of social validity in research reports. We recognize that measuring social validity data may place an additional burden on participants and other stakeholders. Furthermore, reporting detailed social validity data in page-restricted research journals is not always feasible. Given the importance of social validity (especially in interventions with DLLs); however, we suggest continued innovations in the measurement and reporting of social validity.

Additional research may also be beneficial to identify best practices for eliciting authentic and in-depth responses about participants' acceptance of an intervention. In studies with DLLs, it may also be important for researchers to assess whether materials, instructions, and intervention strategies were easy for caregivers to understand. Future research may explore the impact of varying cultural norms (e.g., where it may be impolite/inappropriate to share directly) on social validity responses. Given that the majority of DLLs in the United States are living in poverty (Park et al., 2017), future research should also consider how participant incentives may affect social validity ratings in this group. The receipt of books, monetary reward, or other tokens of appreciation may increase the overall perceived value of participation. Furthermore, it is possible that an incentive decreases the likelihood of negative feedback if the respondent feels indebted to the researcher(s). Researchers in future studies may consider these effects. Additional research is needed to examine the effects of including multiple sources within the same study to

triangulate results and compare and contrast views on social validity across groups. For example, information provided by stakeholders from similar cultural and linguistic backgrounds as study participants should be compared with data gathered from the participants themselves. In studies where two or more cultural and linguistic groups are represented, social validity data may be assessed differentially to determine whether interventions are appropriately designed, carried out, and similarly effective for different groups.

## Conclusion

Our findings support the need for additional efforts to ensure that language interventions for young DLLs are appropriate, meaningful, and responsive to the values and preferences of a culturally and linguistically diverse population. Given that social validity may impact an intervention's likelihood of adoption (Wolf, 1978), social validity remains a critical component of intervention research. Innovative practices are needed to identify social, cultural, and linguistically valid interventions to facilitate better outcomes for young DLLs.

## Acknowledgments

The authors acknowledge Drs. Judith Carta and Melissa Baralt for their contributions to related work that supported this project.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported, in part, by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) under grant UA6MC 27762, Bridging the Word Gap Research Network. This information or content and conclusions are those of the authors and should not be construed as the official position or policy of, nor should any endorsements be inferred by HRSA, HHS, or the U.S. Government.

## ORCID iD

Anne L. Larson  <https://orcid.org/0000-0002-4974-3687>

## References

- References marked with an asterisk indicate studies in the meta-analysis.
- Banks, J. A. (2015). *Cultural diversity and education: Foundations, curriculum, and teaching*. Routledge. <https://doi.org/10.4324/9781315622255>

- Barton, E. E., Meadan-Kaplansky, H., & Ledford, J. R. (2018). Independent variables, fidelity, and social validity. In *Single case research methodology* (pp. 133-156). Routledge.
- \*Bernhard, J. K., Cummins, J., Campoy, F. I., Ada, A. F., Winsler, A., & Bleiker, C. (2006). Identity texts and literacy development among preschool English language learners: Enhancing learning opportunities for children at risk for learning disabilities. *Teachers College Record, 108*(11), 2380–2405. <https://doi.org/10.1111/j.1467-9620.2006.00786.x>
- \*Bernhard, J. K., Winsler, A., Bleiker, C., Ginieniewicz, J., & Madigan, A. L. (2008). "Read my story!" Using the early authors program to promote early literacy among diverse, urban preschool children in poverty. *Journal of Education for Students Placed at Risk, 13*(1), 76–105. <https://doi.org/10.1080/10824660701860458>
- \*Boyce, L. K., Innocenti, M. S., Roggman, L. A., Jump Norman, V. K., & Ortiz, E. (2010). Telling stories and making books: Evidence for an intervention to help parents in migrant head start families support their children's language and literacy. *Early Education and Development, 21*, 343–371. <https://doi.org/10.1080/10409281003631142>
- \*Caesar, L., & Nelson, N. W. (2014). Parental involvement in language and literacy acquisition: A bilingual journaling approach. *Child Language Teaching and Therapy, 30*(3), 317–336. <https://doi.org/10.1177/0265659013513028>
- Carter, S. L., & Wheeler, J. J. (2019). Background of social validity. In S. L. Carter (Ed.), *The social validity manual* (pp. 1–19). Elsevier. <https://doi.org/10.1016/C2009-0-01897-2>
- Castro, F. G., Barrera, M., Jr., & Holleran Steiker, L. K. (2010). Issues and challenges in the design of culturally adapted evidence-based interventions. *Annual Review of Clinical Psychology, 6*, 213–239. <https://doi.org/10.1146/annurev-clinpsy-033109-132032>
- Cho, J., & Trent, A. (2006). Validity in qualitative research revisited. *Qualitative Research, 6*, 319–340. <https://doi.org/10.1177/1468794106065006>
- Cook, B. G., Cook, L., & Landrum, T. J. (2013). Moving research into practice: Can we make dissemination stick? *Exceptional Children, 79*, 163–180. <https://doi.org/10.1177/001440291307900203>
- \*Cooke, N. L., Machiewicz, S. M., Wood, C. L., & Helf, S. (2009). The use of audio prompting to assist mothers with limited English proficiency in tutoring their pre-kindergarten children on English vocabulary. *Education and Treatment of Children, 32*, 213–229. <https://doi.org/10.1353/etc.0.0057>
- Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Pearson.
- Cycyk, L. M. & Hammer, C. S. (2018). Beliefs, values, and practices of Mexican immigrant families towards language and learning in toddlerhood: Setting the foundation for early childhood education. *Early Childhood Research Quarterly*. Advance online publication. <http://doi.org/10.1016/j.ecresq.2018.09.009>
- Division for Early Childhood. (2014). DEC recommended practices in early intervention/earlychildhood special education 2014. <http://www.dec-sped.org/recommendedpractices>
- Dunst, C. J., Raab, M., & Hamby, D. W. (2016). Interest-based everyday child language learning. *Revista de Logopedia, Foniatria y Audiología, 36*(4), 153–161. <https://doi.org/10.1016/j.rlfa.2016.07.003>
- Durán, L. K., Hartzheim, D., Lund, E. M., Simonsmeier, V., & Kohlmeier, T. L. (2016). Bilingual and home language interventions with young dual language learners: A research synthesis. *Language Speech and Hearing Services in Schools, 47*(4), 347–371. [https://doi.org/10.1044/2016\\_LSHSS-15-0030](https://doi.org/10.1044/2016_LSHSS-15-0030)
- \*Hammer, C. S., & Sawyer, B. (2016). Effects of a culturally responsive interactive book-reading intervention on the language abilities of preschool dual language learners: A pilot study. *NHSA Dialog, 18*(4), 59–79.
- Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., . . . Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Information, 95*, Article 103208.
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research Electronic Data Capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Information, 42*(2), 377–381.
- Horner, R. H., Car, E. G., 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*, 165–179. <https://doi.org/10.1177/001440290507100203>
- Hurley, J. J. (2012). Social validity assessment in social competence interventions for preschool children: A review. *Topics in Early Childhood Special Education, 32*, 164–174. <https://doi.org/10.1177/0271121412440186>
- \*Ijalba, E. (2015). Effectiveness of a parent-implemented language and literacy intervention in the home language. *Child Language Teaching and Therapy, 31*(2), 207–220. <https://doi.org/10.1177/0265659014548519>
- Landale, N. S., Oropesa, R. S., & Bradatan, C. (2006). Hispanic families in the United States: Family structure and process in an era of family change. In National Research Council (US) Panel on Hispanics in the United States (Ed.), *Hispanics and the future of America* (pp. 138–178). National Academies Press.
- Larson, A. L., Cycyk, L. M., Carta, J. J., Hammer, C. S., Baralt, M., Uchikoshi, Y., An, Z. G., Wood, C. (2020). A systematic review of language-focused interventions for young children from culturally and linguistically diverse backgrounds. *Early Childhood Research Quarterly, 50*(1), 157–178. <https://doi.org/10.1016/j.ecresq.2019.06.001>
- Ledford, J. R., Hall, E., Conder, E., & Lane, J. D. (2016). Research for young children with autism spectrum disorders: Evidence of social and ecological validity. *Topics in Early Childhood Special Education, 35*, 223–233. <https://doi.org/10.1177/0271121415585956>
- Leko, M. M. (2014). The value of qualitative methods in social validity research. *Remedial and Special Education, 35*, 275–286. <https://doi.org/10.1177/074193514544002>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., . . . Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses

- of studies that evaluate health care interventions: Explanation and elaboration. *PLOS Medicine*, 6(7), Article e1000100. <https://doi.org/10.1371/journal.pmed.1000100>
- \*Lim, Y. S., & Cole, K. N. (2002). Facilitating first language development in young Korean children through parent training in picture book interactions. *Bilingual Research Journal*, 26(2), 367–381. <https://doi.org/10.1080/01443410.2012.744159>
- Lloyd, J. W., & Heubusch, J. D. (1996). Issues of social validation in research on serving individuals with emotional and behavior disorders. *Behavioral Disorders*, 22, 8–14. <https://doi.org/10.1177/019874299602200105>
- McWilliam, R. A., Tocci, L., & Harbin, G. L. (1998). Family-centered services: Service providers' discourse and behavior. *Topics in Early Childhood Special Education*, 18, 206–221. <https://doi.org/10.1177/0271112149801800404>
- Mejia, A., Leijten, P., Lachman, J. M., & Parra-Cardona, J. R. (2017). Different strokes for different folks? Contrasting approaches to cultural adaptation of parenting interventions. *Prevention Science*, 18(6), 630–639. <http://doi.org/10.1007/s11121-016-0671-2>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative data analysis*. SAGE.
- Moiduddin, E., Aikens, N., Tarullo, L., West, J., & Xue, Y. (2012, September). *Child outcomes and classroom quality in FACES 2009* (OPRE Report 2012-37a). Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, 86(5), 1351–1370. <https://doi.org/10.1111/cdev.12398>
- Odom, S. L., & Strain, P. S. (2002). Evidence-based practice in early intervention/early childhood special education: Single-subject design research. *Journal of Early Intervention*, 25, 151–160. <https://doi.org/10.1177/105381510202500212>
- Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children* (Vol. 2). Multilingual Matters.
- Park, M., O'Toole, A., & Katsiaticas, C. (2017). *Dual language learners: A national demographic and policy profile*. Migration Policy Institute.
- Parra Cardona, J. R., Domenech-Rodriguez, M., Forgatch, M., Sullivan, C., Bybee, D., Holtrop, K., . . . Bernal, G. (2012). Culturally adapting an evidence-based parenting intervention for Latino immigrants: The need to integrate fidelity and cultural relevance. *Family Process*, 51(1), 56–72. <https://doi.org/10.1111/j.1545-5300.2012.01386.x>
- Patton, M. Q. (2015). *Qualitative research and methods: Integrating theory and practice* (4th ed.). SAGE.
- \*Peredo, T. N., Zelaya, M. I., & Kaiser, A. P. (2017). Teaching low-income Spanish-speaking caregivers to implement EMT en Español with their young children with language impairment: A pilot study. *American Journal of Speech-Language Pathology*, 27, 136–153. [https://doi.org/10.1044/2017\\_AJSLP-16-0228](https://doi.org/10.1044/2017_AJSLP-16-0228)
- Phillips, D. A., & Shonkoff, J. P. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. National Academies Press.
- Place, S., & Hoff, E. (2011). Properties of dual language exposure that influence 2-year-olds' bilingual proficiency. *Child Development*, 82(6), 1834–1849. <https://doi.org/10.1111/j.1467-8624.2011.01660.x>
- Priest, J. S., McConnell, S. R., Walker, D., Carta, J. J., Kaminski, R. A., McEvoy, M. A., . . . Shinn, M. R. (2001). General growth outcomes for young children: Developing a foundation for continuous progress measurement. *Journal of Early Intervention*, 24(3), 163–180. <https://doi.org/10.1177/10538151010240030101>
- Reardon, S. F., & Galindo, C. (2009). The Hispanic-White achievement gap in math and reading in the elementary grades. *American Educational Research Journal*, 46(3), 853–891. <https://doi.org/10.3102/0002831209333184>
- Resnicow, K., Soler, R., Braithwaite, R. L., Ahluwalia, J. S., & Butler, J. (2000). Cultural sensitivity in substance use prevention. *Journal of Community Psychology*, 28, 271–290. [https://doi.org/10.1002/\(SICI\)1520-6629\(200005\)28:3<271::AID-JCOP4>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1520-6629(200005)28:3<271::AID-JCOP4>3.0.CO;2-I)
- Roberts, M. Y., & Kaiser, A. P. (2015). Early intervention for toddlers with language delays: A randomized controlled trial. *Pediatrics*, 135(4), 686–693. <https://doi.org/10.1542/peds.2014-2134>
- \*Roberts, T. A. (2008). Home storybook reading in primary or second language with preschool children: Evidence of equal effectiveness for second-language vocabulary acquisition. *Reading Research Quarterly*, 43(2), 103–130. <https://doi.org/10.1598/RRQ.43.2.1>
- \*Simon-Cerejido, G., & Gutiérrez-Clellen, V. F. (2014). Bilingual education for all: Latino dual language learners with language disabilities. *International Journal of Bilingual Education and Bilingualism*, 17(2), 235–254. <https://doi.org/10.1080/13670050.2013.866630>
- Snodgrass, M. R., Chung, M. Y., Meadan, H., & Halle, J. W. (2018). Social validity in single case research: A systematic literature review of prevalence and application. *Research in Developmental Disabilities*, 74, 160–173. <https://doi.org/10.1016/j.ridd.2018.01.007>
- Spear, C. F., Strickland-Cohen, M. K., Romer, N., & Albin, R. W. (2013). An examination of social validity within single-case research with students with emotional and behavioral disorders. *Remedial and Special Education*, 34(6), 357–370. <https://doi.org/10.1177/0741932513490809>
- Strain, P. S., Barton, E. E., & Dunlap, G. (2012). Lessons learned about the utility of social validity. *Education and Treatment of Children*, 35, 183–200. <https://doi.org/10.1353/etc.2012.0007>
- Triandis, H. C., Marin, G., Lisansky, J., & Betancourt, H. (1984). Simpatia as a cultural script of Hispanics. *Journal of Personality and Social Psychology*, 47(6), 1363–1375. <https://doi.org/10.1037/0022-3514.47.6.1363>
- U.S. Department of Health and Human Services. (2013). *Report to Congress on dual language learners in Head Start and Early Head Start*. <https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/report-congress-dual-language-learners.pdf>
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11, 203–214. <https://doi.org/10.1901/jaba.1978.11-203>