Article

HAMMILL INSTITUTE ON DISABILITIES

Topics in Early Childhood Special Education
2016, Vol. 35(4) 223–233

DOI: 10.1177/0271121415585956
tecse.sagepub.com
SAGE

Research for Young Children With Autism Spectrum Disorders: Evidence of Social and Ecological Validity

Jennifer R. Ledford, PhD[1], Emilie Hall, BA[1], Emily Conder, BA[1], and Justin D. Lane, PhD[2]

Abstract

The social and ecological validity of a body of research may impact the degree to which interventions will be used outside of research contexts. The purpose of this review was to determine the extent to which social and ecological validity were demonstrated for interventions designed to increase social skills for young children with autism spectrum disorders (ASD). Results indicated that although the percentage of studies including social validity assessment has remained stable over the 20-year review period, subjective assessments of social validity have increased and objective assessments have decreased. Acceptability was measured more often than feasibility or importance. Approximately half of the studies included indigenous implementers, typical social partners, or typical settings. Suggestions include additional research on the validity of measures, explicit reporting by researchers, and the use of multiple, objective, and psychometrically sound social validity assessments.

Keywords

social validity, ecological validity, consumer satisfaction

Single case research has contributed to the development of a number of interventions for young children with disabilities, many of which are behavioral in nature (Odom & Strain, 2002). Beginning nearly 40 years ago, interventionists and researchers using these interventions began arguing that evidence of effectiveness was not sufficient—evidence of relevance and practical significance was also needed (Brooks & Baumeister, 1977; Wolf, 1978). These characteristics can be considered as components of social and ecological validity. Wolf conceptualized three separate components of social validity: (a) significance of goals, (b) appropriateness of procedures, and (c) importance of effects. Ecological validity is a closely related construct: It refers to the likelihood that outcomes from a given study are meaningful outside the research context and might be referred to as the *feasibility of implementation* of an intervention (Brooks & Baumeister, 1977; Gast, 2014). Some researchers consider feasibility to be a component of social validity (cf. Machalicek, O'Reilly, Beretvas, Sigafoos, & Lancioni, 2007). Others have referred to studies having ecological validity when they are implemented in typical contexts (Clarke & Dunlap, 2008) or have suggested that the use of typical implementers and contexts result in improved social validity (Horner et al., 2005). Thus, ecological and social validity are associated constructs that are often highly valued in the fields of early childhood special education and early intervention (ECSE/EI).

The perceived importance of social and ecological validity is due in part to the idea that acceptability, feasibility, and significance of a study may impact the "scaling up" of the related intervention for use in typical contexts. Contemporary researchers have argued that attempts to control environmental variability in educational research (e.g., minimizing the effects of implementer skill by including only highly trained implementers) may result in less variable outcomes and increased chance for experimental control; it might also considerably constrain the extent to which findings are generalizable (cf. Phillips, 2014). Thus, researchers in ECSE/EI may opt for a greater likelihood of demonstrating experimental control, but at the cost of decreased ecological or social validity. This loss of social and ecological validity is critical because there may be a positive correlation between social validity and the extent to which interventions are used (cf. Carter & Pesko, 2008). Thus, when determining whether a practice is evidence based in ECSE/EI, it may be critical that the evidence for social and ecological validity be considered

[1]Vanderbilt University, Nashville, TN, USA
[2]University of Kentucky, Lexington, USA

Corresponding Author:
Jennifer R. Ledford, Vanderbilt University, Peabody Box 228, Nashville, TN 37203, USA.
Email: jennifer.ledford@vanderbilt.edu

Table 1. Descriptions of Potential Evidence for Social and Ecological Validity.

Component	Description
Feasibility	
Typical settings	Study is conducted in the student's home, school, or community-based settings
Typical activities	Study is conducted during activities that naturally occur
Indigenous implementers	Study is conducted by adults or peers who are customarily present
Feasible financial supports	Materials used in the study are not cost-prohibitive for use in typical settings
Practical personnel supports	Supports provided to personnel are not prohibitive for use in typical settings
Maintenance of intervention	Indigenous agents continue using intervention strategies when research support is removed
Maintenance of behavior change	Behavior change is maintained even when research support is removed
Interviews or questionnaires	Indirect consumers report the intervention can be implemented in typical settings
Acceptance/satisfaction	
Participant choice (goals)	Research participants choose target goals
Participant choice (intervention)	Research participants choose intervention components given those likely to be successful
Consumer selection (goals)	Non-participant consumers select the target goal as important
Consumer selection (intervention)	Non-participant consumers select the intervention as acceptable
Interviews or questionnaires	Indirect consumers report the intervention and/or target behaviors are acceptable
Importance/significance	
Interviews or questionnaires	Indirect consumers report the behavior change was important
Normative comparisons	Behavior change results in behavior consistent with or closer to that exhibited by peers
Blind ratings	Non-participant consumers who cannot ascertain condition (e.g., do not know whether they are rating pre- or post-intervention videos) rate target behavior change as important (e.g., rate behavior exhibited during or after intervention conditions as more desirable or less atypical than baseline behavior)

(e.g., Can parents and practitioners reliably implement the intervention? Will they do so?).

In a study using single case design, a consistent and replicated intervention effect (existence of a functional relation) does not necessarily mean conclusions drawn are meaningful to either current participants (a question of social validity) or to potential future consumers (a question of ecological validity). We argue that at least three critical questions exist regarding ecological and social validity: (a) Is the intervention feasible? (b) Are the target behaviors, intervention, and outcomes acceptable and are consumers satisfied with them? and (c) Does the intervention result in significant or important change? Social and ecological validity components related to these three areas (*feasibility, acceptability*, and *significance)* are shown in Table 1. There is some overlap among these terms—for example, a parent might find intervention procedures *acceptable* only if the resulting child behavior change is *significant.* Moreover, social validity assessment may overlap with the assessment of effects—that is, that direct and indirect consumers may be unlikely to rate a change as *significant* if a functional relation is not demonstrated for primary measures (although there is some evidence that this is not necessarily true; Strain, Barton, & Dunlap, 2012).

In previous reviews, researchers have noted insufficient measurement of social validity, and have called for improved measurement and reporting of the extent to which studies are ecologically and socially valid (Hurley, 2012; Kennedy, 1992; McDonald & Machalicek, 2013; Odom & Strain, 2002; Spear, Strickland-Cohen, Romer, & Albin, 2013). Odom and Strain reported that only 15% of intervention studies for young children with disabilities assessed acceptability and only 27% assessed the importance of effects. Similarly, Hurley (2012) reported that only 27% of studies designed to improve social competence for preschoolers with disabilities included measures of social validity. Clarke and Dunlap (2008) reported variable inclusion of social validity measures in intervention studies for children and young adults with disabilities across three journals (3%–31%). They reported higher rates of ecological validity, described in part as the use of typical physical contexts (19%–63%), activity contexts (24%–69%), or social contexts (22%–69%). However, Clarke and Dunlap reported relatively uncommon use of family (2%–10%) or teacher (4%–17%) implementation within their sample of studies. It may be that use of indigenous implementers and typical contexts is higher in research with young children with autism spectrum disorders (ASD), because intervening in natural contexts is valued in ECSE/EI, but this assumption has not been evaluated.

Objective and Subjective Data Collection

Social validity data are commonly collected using subjective post-intervention measures (Kong & Carta, 2011; Machalicek

et al., 2007); these measures are often used to gather opinions regarding acceptability and feasibility from consumers following an intervention study. It is also possible to collect objective data regarding the feasibility, acceptability, satisfaction, and significance of target behaviors, interventions, and outcomes. One objective method for assessing the social validity of an intervention is to have blind raters (e.g., those who are unaware of the condition being implemented) report the extent to which observed behavior is typical, acceptable, or positive; the extent to which the procedures used are enjoyable, beneficial, or feasible; and the extent to which behavior changes between two different (but unlabeled) session types are positive, noticeable, or important. A second procedure for objectively assessing social validity is to compare outcome data with data from individuals who display of age-appropriate or socially acceptable behaviors (Ennis, Jolivette, Fredrick, & Alberto, 2013). In this case, authors might demonstrate that behavior of target participants during baseline conditions is considerably different from peer comparison data, but that this difference is minimal, non-existent, or lessened during or after intervention. Although the purpose of an intervention is not always to completely ameliorate a problem to levels expected of comparison peers, these comparisons may inform general expectations for researchers when assessing meaningful change. Thus, objective measures of social validity exist, although previous studies have suggested that they are not often used (cf. Kennedy, 1992; Spear et al., 2013).

Another example of the objective measurement of *acceptability* is the use of participant choice. Hanley (2010) argued that it is possible to objectively assess the degree to which an intervention is acceptable for direct consumers (e.g., children, even those with limited communication skills), by providing treatment options and assessing choices (e.g., in a simultaneous treatments design; see Ledford, Wolery, & Gast, 2014). This could occur instead of or in addition to the usual subjective measures of acceptability typically assessed via adults. Measures derived from objective sources are difficult to analyze because of inadequate evidence that they are psychometrically sound. An additional concern is that consumers may feel an obligation to report positive results because of perceived researcher preferences. This may be particularly important in single case research because the relatively small number of participants decreases the likelihood of anonymity in reporting. Results from these types of social validity assessments have been described as "indiscriminately positive" (Machalicek et al., 2007).

Although social and ecological validity are often measured using subjective checklists and interviews, several research teams (cf. Ennis et al., 2013; Kennedy, 1992, 2003; Ledford et al., 2014) have suggested objective measurement, including (a) normative comparisons, (b) blind ratings, (c) evidence maintenance of use by indigenous implementers, (d) evidence of maintenance of behavior change, (e) use of participant choice, and (f) consideration of consumer preferences. The analysis of these data, along with assessment of the use of typical settings and indigenous implementers, may allow consumers to determine to what extent a study shows evidence of social and ecological validity.

Social Skills Intervention Research for Young Children With ASD

One body of literature often controversial in regard to acceptability, feasibility, and significance is the literature assessing the effectiveness of social skills interventions for young children with ASD. Thus, meaningful measurement of social validity may be especially important in this area. Although recommended practices for young children include provision of intervention by indigenous implementers in typical social contexts (Division for Early Childhood of the Council for Exceptional Children, 2014), some of the most widely used and researched interventions may not meet these standards (e.g., discrete trial training conducted in a clinic setting by an advanced researcher). There is a longstanding debate about the usefulness of these interventions that (a) may be unacceptable to parents and caregivers, (b) may not be feasible for widespread implementation, and (c) may be used to teach irrelevant skills to children who are in need of practical skills for social communication in typical contexts. Thus, despite the increased number of recommended practices for individuals with ASD (cf. Wong et al., 2014), the social validity of some interventions with research support may be questionable.

Social skills impairments are one of the defining characteristics of ASD, and interventions designed to increase pro-social behaviors for children with ASD have the potential to have important and long-lasting impacts on social relationships and other outcomes, such as educational placement. A variety of interventions exist for increasing pro-social behaviors; reports of the effectiveness of these interventions, as a whole, have been positive (e.g., Reichow & Volkmar, 2010; Wong et al., 2014). However, even when a practice has been deemed evidence based, important questions remain. Among these important questions are ones of feasibility, acceptability, and significance; social skills interventions may only be effective if indigenous implementers can and will use them and if the behavior change they produce is meaningful. This may be an especially important consideration in ECSE/EI settings, where children spend the majority of their time with parents and early childhood practitioners who may not have specific training in evidence-based practice.

The purpose of this review was to analyze studies designed to increase social competence for young children

with ASD to determine to what extent evidence for social and ecological validity was present. Research questions were as follows:

Research Question 1: To what extent do authors report evidence of social validity?
Research Question 2: What types of social validity data are reported?
Research Question 3: To what extent do studies include characteristics that show evidence of ecological validity?
Research Question 4: Which type of evidence of ecological validity is most often reported?

Method

The studies included in this review were also included in a larger review of social skills interventions for individuals with ASD (Ledford, King, Harbin, & Zimmerman, 2015). Inclusion criteria for that review included (a) inclusion of an individual with ASD, (b) use of a single case design with at least three potential demonstrations of effect and visual data presentation, (c) assessment of a social skills intervention with a dependent variable related to human-to-human interactions, and (d) publication in a peer-reviewed journal between 1994 and 2013. An additional criterion for inclusion in the current review was that all participants included were 8 years of age or below and at least half were aged 5 years or younger. Coding was conducted separately for each *study*, which was defined by the use of a stand-alone single case design; some articles included multiple studies. For example, an article with three A-B-A-B designs included three *studies* with three participants, whereas an article with a single multiple baseline across three participants design included a single *study* with three participants.

General coding was done and reported as part of the previously completed review of social skills interventions; these variables included year of publication, journal, intervention components, dependent variables, participant age, design type, and whether a functional relation existed. A doctoral level single case researcher completed general coding, with independent reliability conducted for 21% of all studies in the larger review by a second doctoral level researcher. Both coders were also certified behavior analysts. Average agreement was 95.4% across codes. The presence of a functional relation was separately determined for each study, using a dichotomous yes/no decision. Determinations regarding functional relations were made using visual analysis, with consideration for changes in level, trend, and variability across conditions. Additional information about coding, including specific independent and dependent variable definitions, is available in the article describing the larger review (Ledford, King, et al., 2015).

Variables specific to the questions of social and ecological validity were coded to answer the research questions for this review. These codes included whether authors self-reported the collection of social validity data and what type of data were reported (interview or questionnaire results, blind raters, normative comparisons, and other). In addition, we coded whether direct and indirect consumers were provided with opportunities to make decisions regarding independent or dependent variables selection. Two graduate students assessed social and ecological validity variables for reviewed articles. For 34% of articles, both students independently coded all variables and the first author assessed agreement between coders. Across articles and codes, mean agreement was 96.7%, with a range of 89.2% to 100% across codes. Disagreements were reviewed and reconciled by the first author.

Social Validity Codes

- *Interview or questionnaire results* were coded as occurring if authors reported data on ratings of procedures, objectives, and/or effects produced by children, practitioners, or caregivers directly or indirectly involved in the intervention. Anecdotal reports were excluded.
- *Blind ratings* were coded as occurring if (a) raters were blind to the condition type when they observed (typically via video) and (b) these raters provided judgment ratings on child behavior, procedures, or outcomes that allowed direct comparisons between conditions. Primary data collection via blind raters was not included although blind data collection is beneficial and serves to decrease risk of bias and increase confidence in results (cf. Barton, Reichow, Schnitz, Smith, & Sherlock, 2015).
- *Normative comparisons* were coded if authors used data collected from non-participating children to either (a) choose intervention criteria or (b) serve as a comparison for data collected from participating children.
- *The provision of choice* to direct (e.g., children or implementing adults) or indirect consumers (e.g., parents, teachers) was coded if researchers reported that these individuals made choices regarding the intervention to be used or the behaviors to be targeted.

For the three primary types of social validity assessments (interview or questionnaire results, blind ratings, and normative comparisons), we coded the degree to which results reported were positive. Because results within and across types were not directly comparable, we coded results as being *only positive, only negative*, or *mixed*. Mixed ratings could occur when results were positive for some items and negative for others (e.g., satisfaction was

rated favorably but feasibility was rated poorly) or when results were positive for some participants and negative for others. In addition to type and results, we also coded the domain authors reported: *feasibility, acceptance or satisfaction*, and *importance or significance.* Single or multiple domains could be reported by an author (e.g., if authors reported degree to which teachers found procedures acceptable and the extent to which they could use the procedures in the classrooms, *acceptance* and *feasibility* were coded).

Ecological Validity Codes

In addition to authors' reports of *objective* and *subjective* social validity data reported in articles, we also coded the extent to which authors reported using procedures consistent with increased ecological validity. Specifically, we recorded,

- Type of implementer and social partner (e.g., teacher, parent, peer, researcher). These codes could be different within a study (e.g., a researcher implemented an intervention, while measuring changes in social responses to peer social partners).
- Whether the implementer and social partners were indigenous to the child's environment. Implementers and social partners were considered separately (as described above). In general, teachers, parents, and peers were considered as indigenous and researchers were considered as not indigenous.
- Whether the settings and activities of the intervention were typical. A setting was considered typical if the participating children would have been present if he or she were not involved in research (e.g., homes were considered typical, clinics were considered not typical). An activity was considered typical if the participating children would have been engaging in the activity if he or she were not involved in research (e.g., free play was considered typical, being pulled out to engage in a social skills lesson was considered not typical).
- Whether there was evidence of continued use by typical agents without researcher support.
- Whether there was evidence of maintenance of behavior change when intervention was withdrawn.

Results

Studies included in the review included assessment of interventions designed to increase pro-social behaviors directed at a human social partner for young children with ASD; 109 studies in 54 published articles met inclusion criteria (see the appendix). *Studies* were defined as stand-alone single case designs; as such multiple studies could be included in a published article. Common intervention components included prompting (n = 43), environmental arrangement (n = 26), social skills training (n = 22), peer training (n = 21), and the use of responsive interactions (n = 20). Many of the studies were published in *Journal of Positive Behavior Interventions* (n = 25), *Journal of Autism and Developmental Disorders* (n = 24), or *Journal of Applied Behavior Analysis* (n = 19; with fewer studies published in 19 other journals). A variety of designs were represented, all with at least three potential demonstrations of effect, including multiple baseline designs across participants (n = 46) or behaviors (n = 18), A-B-A-B designs (n = 16), and alternating treatments design (n = 16; with 13 studies using multiple baseline or probe across contexts or settings or multiple probe across participants). When using a stringent measure of effects based on consistent and replicated behavior change, the percentage of studies showing evidence of a functional relation was 53%.

Measurement of Social Validity

In this review, fewer than half of the included studies (n = 48, 44%) reported measurement of social validity data. The most common type of data collected was post-intervention ratings completed via interview or questionnaire (n = 34). In contrast, relatively few studies used objective measures of social validity, including blind ratings (n = 12; e.g., Buffington, Krantz, McClannahan, & Poulson, 1998; Whalen & Schreibman, 2003) and normative comparisons (n = 9; e.g., McGee & Daly, 2007; Zanolli, Daggett, & Adams, 1996). Other measures of social validity included the extent to which direct consumers, including teachers, parents, and therapists, were given the opportunity to choose intervention targets (n = 15; cf. Crozier & Tincani, 2007; Ingersoll & Wainer, 2013; Jones & Feeley, 2007) or intervention procedures (n = 0). Although no study reported choice of intervention procedures or components, several did report that consumers assisted in the selection of materials or contexts (cf. Kohler, Greteman, Raschke, & Highnam, 2007; Maione & Mirenda, 2006).

Measurement of Social Validity Over Time

To determine changes in measurement over time, the publication period was divided into four periods (each comprised of 5 years). As shown in Figure 1, the percentage of studies reporting social validity data was relatively stable over time (38% during the first time period, 41% during the final time period) with the highest percentage of studies reporting data during the 1999 to 2003 time period. Interestingly, measurement of *subjective* social validity, despite its limitations, increased considerably (5% during the first time period, 34% during the final time period). Meanwhile, measurement of *objective* social validity data has steadily decreased (31% during the first time period, 14% during the final time period).

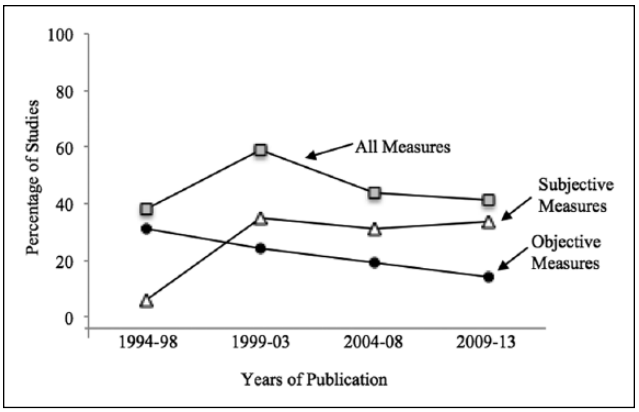

Figure 1. Percentage of studies over time that measured and reported any type of social validity data, subjective data, and objective data.

Measurement of Feasibility, Acceptability, and Significance

Almost all of the studies using subjective social validity measurement reported measurement of acceptability or satisfaction ($n = 32$); fewer reported measuring feasibility ($n = 15$) or significance ($n = 15$). Objective social validity measurement was most often reported to assess significance or importance ($n = 12$), and less often measured acceptance or satisfaction ($n = 5$) or feasibility ($n = 1$).

Social Validity and Primary Study Results

The percentage of studies demonstrating a functional relation for primary data was approximately the same for studies reporting social validity data (54%) as for the group of studies as a whole (53%). Approximately 80% of studies reporting the use of *subjective* questionnaires reported only positive results, whereas 20% reported mixed results. Results of objective social validity assessments (using blind ratings or normative comparisons) were similar to each other but quite different from those reported regarding subjective assessments. Fewer than half of the studies using blind ratings and normative comparisons (42% and 44%, respectively) reported only positive results, with the remainder of studies reporting mixed results. No study reported entirely negative social validity results.

When comparing social validity results to primary data conclusions (e.g., functional relation present or absent), data were also discrepant. For studies with positive subjective social validity data, 50% did not provide evidence of a functional relation. However, 73% of studies with positive objective social validity results also had positive results regarding primary data (functional relation). This suggests positive objective social validity results are more likely to occur when primary effects are present; subjective social validity results are not similarly associated with primary results.

Ecological Validity

Approximately half of the studies reported evidence of common indicators of ecological validity, including indigenous implementers ($n = 49$, 45%), typical social partners ($n = 61$, 56%), or typical settings ($n = 62$, 57%). Many of these studies ($n = 42$, 38%) reported evidence of all three indicators. Fewer studies ($n = 30$, 28%) reported that studies occurred in the context of typically occurring activities (e.g., the study was conducted in a typical environment like an early childhood program, and also during regular ongoing activities in that setting). Very few studies reported whether indigenous implementers continued using the intervention after the treatment condition was completed ($n = 8$, 7%).

Discussion

The purpose of this review was to determine the extent to which studies designed to increase pro-social behaviors for young children with ASD reported evidence of social validity and ecological validity. Results suggest that (a) about half of the studies report assessments of social validity; (b) about half of the studies report the use of indigenous implementers, typical social partners, or typical settings and more than a third report the use of all three; (c) measurement of social validity has been fairly stable over time, although *objective* measurement has decreased and *subjective* measurement has increased; (d) studies using subjective measures most often reported assessment of acceptability/satisfaction; (e) studies using *objective* measures most often reported assessment of importance/significance; and (f) positive results from *objective* measures, as compared with *subjective* measures, were more often associated with the presence of a functional relation for primary outcomes.

Limitations

When interpreting conclusions from this review, several limitations should be considered, including (a) the exclusion of studies using group comparison designs, (b) the exclusion of studies that did not permit the evaluation of a functional relation (e.g., multiple baseline designs with two tiers, A-B-A-B designs without three data points in each condition), and (c) the restricted focus on participants with ASD. Despite these limitations and the need for future research in the area, some conclusions can be drawn regarding the social and ecological validity of interventions for young children with ASD.

Conclusion

Although not directly comparable, the results from this review are somewhat discrepant from previous reports, which have found less frequent measurement of social validity in journals specific to behavior analysis (Carr, Austin, Britton, Kellum, & Bailey, 1999; Kennedy, 1992) and early intervention (Hurley, 2012). As part of the larger review, studies were excluded if determination of a functional relation was not possible (e.g., fewer than three data points per condition or fewer than three potential demonstrations of effect did not exist; Ledford, King, et al., 2015). This may have had the effect of excluding studies that were lower in quality; there may be a relationship between methodological rigor and social and ecological validity. Evidence for ecological validity was higher than in a previous review (Clarke & Dunlap, 2008), although it approximated results from one journal in that review with a specific aim related to ecological validity (*Journal of Positive Behavior Interventions*). Despite equal or better ecological and social validity than some reported reviews, the rate of positive outcomes (percentage of studies with a functional relation, 53%) is rather low in comparison with both the other reviews and to the larger review of social skills interventions for individuals of all ages (Ledford, King, et al., 2015).

The discrepant findings across social validity type suggest subjective and objective measures of social validity may focus on two different constructs, with subjective measurement often associated with satisfaction ratings and objective measures associated with assessments of significance. Most often, satisfaction ratings were researcher-designed, without evidence of adequate psychometric properties; several studies did report using established measures (e.g., Intervention Rating Profile–15 [IRP-15]; Martens, Witt, Elliott, & Darveaux, 1985). Given the questionable validity of researcher-developed measures, researchers should consider either using established measures or using objective measures of social validity. Researchers should determine whether a combination of subjective and objective measures best captures measurement of the three components of social validity first described by Wolf (1978). For example, the *appropriateness of procedures* could be assessed by asking parents to select an intervention from a pool of effective strategies, in addition to asking if a prescribed strategy was acceptable on completion of the intervention. It is likely that multiple assessments of social validity are needed to adequately assess all components named by Wolf. Across studies, some components may be more important than others, based on specific research questions and goals.

Although we have suggested that objective measurement of social validity data may have considerable benefits, the use of anecdotal and subjective assessments may also be beneficial. For example, anecdotal reports related to results provide insight into ways procedures can be modified to improve the extent to which practitioners can effectively use them in practice. In one study (Ledford, Lane, Shepley, & Kroll, 2015), a practitioner requested that a procedure involving randomly ordering materials be replaced with a different method. This modification was made, despite a small loss of contextual control. Post hoc narrative reports from practitioners can also be helpful, perhaps more helpful than the usual Likert-type scale rating. For example, data from this and other reviews (cf. Machalicek et al., 2007) suggest subjective ratings are likely to be positive, regardless of the effectiveness of procedures. Instead of asking for simple ratings, more insight might be gained by asking practitioners questions such as "What would you change if you used this intervention?" "What child behavior would you target if you used this intervention?" or "What intervention would be more feasible for you to use?" It is possible that researchers see Likert-type rating scales as an effective way to transform subjective ratings into quantitative data; we suggest, in light of these data, that the transformation may result in invalid and undifferentiated outcomes. Some subjective ratings might be helpful and less prone to bias—for example, one method for assessing the appropriateness of an intervention, as well as feasibility and acceptability, is to ask parents or practitioners to provide ratings regarding multiple interventions that could be used to increase a pro-social behavior (e.g., Which of the described procedures would be easiest for you to use?). Selecting the intervention most highly ranked by the consumer could enhance the usefulness of this method. Collecting this information would allow researchers to rank-order interventions by consumer or a proxy of the consumer. Moreover, this option would require considerably fewer resources than having blind raters view recordings of different types of interventions for essentially the same purpose—assessing acceptability of an intervention. Thus, decreasing the likelihood of bias may be possible with subjective measures, although this possibility should be evaluated. This is a critical line of research, especially given that the increase of subjective measures in comparison with objective measures is likely at least partially explained by differential resource requirements.

Additional research is needed to determine (a) the extent to which subjective measures may be biased under different conditions (e.g., data collected by researcher vs. unrelated personnel, data collected from direct vs. indirect consumers, data collected from validated vs. researcher-developed instruments), (b) the extent to which data collected from objective and subjective measures of social validity result in similar judgments (e.g., may be measuring similar or different constructs), and (c) the possibility that social validity results that are entirely negative (rather than positive or mixed; none of which were reported in studies in this review) are not shared even when corresponding primary

data are published, potentially resulting in publication bias (for a discussion regarding this problem, see Schwartz & Baer, 1991).

Based on current evidence, objective and valid measures of social and ecological validity are suggested for use in concert with anecdotal and narrative reports from direct and indirect consumers. Moreover, we suggest researchers (a) report specifically which facets of social validity are assessed by their measures (e.g., feasibility, acceptability, significance); (b) explicitly explain difficulties with interpreting subjective measures, including the possibility that implementers may report positive results due to perceived wishes of the researchers; and (c) explicitly describe features of the study and social validity results that might enhance or impede use of a procedure in typical settings.

Appendix

Included Studies

Belchic, J. K., & Harris, S. L. (1994). The use of multiple peer exemplars to enhance the generalization of play skills to the siblings of children with autism. *Child & Family Behavior Therapy, 16*(2), 1–25.

Boyd, B. A., Conroy, M. A., Mancil, G. R., Nakao, T., & Alter, P. J. (2007). Effects of circumscribed interests on the social behaviors of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 1550–1561.

Buffington, D. M., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1998). Procedures for teaching appropriate gestural communication skills to children with autism. *Journal of Autism and Developmental Disorders, 28*, 535–544.

Buggey, T. (2012). Effectiveness of video self-modeling to promote social initiations by 3-year-olds with autism spectrum disorders. *Focus on Autism and other Developmental Disabilities, 27*, 102–110.

Buggey, T., Hoomes, G., Sherberger, M. E., & Williams, S. (2011). Facilitating social initiations of preschoolers with autism spectrum disorders using video self-modeling. *Focus on Autism and Other Developmental Disabilities, 26*, 25–36.

Caballero, A., & Connell, J. E. (2010). Evaluation of the effects of social cue cards for preschool age children with autism spectrum disorders (ASD). *Journal of Behavior Assessment and Intervention in Children, 1*, 25–42.

Carter, C. M. (2001). Using choice with game play to increase language skills and interactive behaviors in children with autism. *Journal of Positive Behavior Interventions, 3*, 131–151.

Crozier, S., & Tincani, M. (2007). Effects of social stories on prosocial behavior of preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*, 1803–1814.

Feldman, E. K., & Matos, R. (2013). Training paraprofessionals to facilitate social interactions between children with autism and their typically developing peers. *Journal of Positive Behavior Interventions, 15*, 169–179.

Ferraioli, S. J., & Harris, S. L. (2011). Teaching joint attention to children with autism through a sibling-mediated behavioral intervention. *Behavioral Interventions, 26*, 261–281.

Franco, J. H., Davis, B. L., & Davis, J. L. (2013). Increasing social interaction using prelinguistic milieu teaching with nonverbal school-age children with autism. *American Journal of Speech-Language Pathology, 22*, 489–502.

Garfinkle, A. N., & Schwartz, I. S. (2002). Peer imitation: Increasing social interactions in children with autism and other developmental disabilities in inclusive preschool classrooms. *Topics in Early Childhood Special Education, 22*, 26–38.

Gena, A. (2006). The effects of prompting and social reinforcement on establishing social interactions with peers during the inclusion of four children with autism in preschool. *International Journal of Psychology, 41*, 541–554.

Gonzalez-Lopez, A., & Kamps, E. M. (1997). Social skills training to increase social interactions between children with autism and their typical peers. *Focus on Autism and Other Developmental Disabilities, 12*, 2–14.

Hancock, T. B., & Kaiser, A. P. (2002). The effects of trainer-implemented enhanced milieu teaching on the social communication of children with autism. *Topics in Early Childhood Special Education, 22*, 39–54.

Ingersoll, B. (2011). The differential effect of three naturalistic language interventions on language use in children with autism. *Journal of Positive Behavior Interventions, 13*, 109–118.

Ingersoll, B., & Lalonde, K. (2010). The impact of object and gesture imitation training on language use in children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research, 53*, 1040–1051.

Ingersoll, B., & Schreibman, L. (2006). Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. *Journal of Autism and Developmental Disorders, 37*, 1446–1456.

Ingersoll, B., & Wainer, A. (2013). Initial efficacy of project ImPACT: A parent-mediated social communication intervention for young children with ASD. *Journal of Autism and Developmental Disorders, 43*, 2943–2952.

Jones, J., Lerman, D. C., & Lechago, S. (2007). Assessing stimulus control and promoting generalization via video modeling when teaching social responses to children with autism. *Journal of Applied Behavior Analysis, 47*, 1–14.

Kaiser, A. P., Hancock, T. B., Nietfeld, J. P. (2000). The effects of parent-implemented enhanced milieu teaching on the social communication of children who have autism. *Early Education and Development, 11*, 423–446.

Koegel, L. K., Koegel, R. L., Green-Hopkins, I., & Barnes, C. C. (2010). Brief report: Question-asking and collateral language acquisition in children with autism. *Journal of Autism and Developmental Disorders, 40*, 509–515.

Koegel, L., Camarata, S. M., Valdez-Menchaca, M., & Koegel, R. L. (1998). Setting generalization of question-asking by children with autism. *American Journal on Mental Retardation, 102*, 346–357.

Kohler, F. W., Anthony, L. J., Stighner, S. A., & Hoyson, M. (2001). Teaching social interaction skills in the integrated preschool: An examination of naturalistic tactics. *Topics in Early Childhood Special Education, 21*, 93–103.

Kohler, F. W., Greteman, C., Raschke, D., & Highnam, C. (2007). Using a buddy skills package to increase the social interactions between a preschooler with autism and her peers. *Topics in Early Childhood Special Education, 27*, 155–163.

Kohler, K. W., Strain, P. S., Hoyson, M., Davis, L., Donina, W. M., & Rapp, N. (1995). Using a group-oriented contingency to increase social interactions between children with autism and their peers—A preliminary-analysis of corollary supportive behaviors. *Behavior Modification, 19*, 10–32.

Krantz, P. J., & McClannahan, L. E. (1998). Social interaction skills for children with autism: A script-fading procedure for beginning readers. *Journal of Applied Behavior Analysis, 31*, 191–202.

Laushey, K. M., & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism and Developmental Disorders, 30*, 183–193.

Leaf, J. B., Oppenheim-Leaf, M. L., Call, N. A., Sheldon, J. B., Sherman, J. A., Taubman, M., . . . Leaf, R. (2012). Comparing the teaching interaction procedure to social stories for people with autism. *Journal of Applied Behavior Analysis, 45*, 281–298.

Leew, S. V., Stein, N. G., & Gibbard, W. (2010). Weighted vests' effect on social attention for toddlers with autism spectrum disorders. *Canadian Journal of Occupational Therapy/Revue Canadienne D'ergothérapie, 77*, 113–124.

MacDuff, J. L., Ledo, R., McClannahan, L. E., & Krantz, P. J. (2007). Using scripts and script-fading procedures to promote bids for joint attention by young children with autism. *Research in Autism Spectrum Disorders, 1*, 281–290.

Maione, L., & Mirenda, P. (2006). Effects of video modeling and video feedback on peer-directed social language skills of a child with autism. *Journal of Positive Behavior Interventions, 8*, 106–118.

Martins, M. P., & Harris, S. L. (2006). Teaching children with autism to respond to joint attention initiations. *Child & Family Behavior Therapy, 28*, 51–68.

Matsuzaki, A., & Yamamoto, J. (2012). Effects of an early intervention program on preverbal communication in a child with autism: Developmental and behavioral analysis with a multiple-baseline design. *Japanese Journal of Special Education, 49*, 657–669.

McGee, G. G., & Daly, T. (2007). Incidental teaching of age-appropriate social phrases to children with autism. *Research and Practice for Persons With Severe Disabilities, 32*, 112–123.

Nelson, C., McDonnell, A. P., Johnston, S. S., Crompton, A., & Nelson, A. (2007). Keys to play: A strategy to increase the social interactions of young children with autism and their typically developing peers. *Education and Training in Developmental Disabilities, 42*, 165–181.

Petursdottir, A. L., McComas, J., McMaster, K., & Horner, K. (2007). The effects of scripted peer tutoring and programming common stimuli on social interactions of a student with autism spectrum disorder. *Journal of Applied Behavior Analysis, 40*, 353–357.

Pollard, J. S., Betz, A. M., Higbee, T. S. (2012). Script fading to promote unscripted bids for joint attention in children with autism. *Journal of Applied Behavior Analysis, 45*, 387–393.

Randolph, J. K., Stichter, J. P., Schmidt, C. T., & O'Connor, K. V. (2011). Fidelity and effectiveness of PRT implemented by caregivers without college degrees. *Focus on Autism and Other Developmental Disabilities, 26*, 230–238.

Reagon, K. A., & Higbee, T. S. (2009). Parent-implemented script fading to promote play-based verbal initiations in children with autism. *Journal of Applied Behavior Analysis, 42*, 659–664. doi:10.1901/jaba.2009.42-659

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders, 40*, 149–166.

Schertz, H. H., & Odom, S. L. (2007). Promoting joint attention in toddlers with autism: A parent-mediated developmental model. *Journal of Autism and Developmental Disorders, 37*, 1562–1575.

Simpson, A., Langone, J., & Ayres, K. (2004). Embedded video and computer based instruction to improve social skills for students with autism. *Education and Training in Developmental Disabilities, 39*, 240–252.

Strain, P. S., & Danko, C. D. (1995). Caregivers encouragement of positive interaction between preschoolers with autism and their siblings. *Journal of Emotional and Behavioral Disorders, 3*, 2–12.

Strain, P. S., Danko, C. D., & Kohler, F. (1995). Activity engagement and social-interaction development in young-children with autism: An examination of "free" intervention effects. *Journal of Emotional and Behavioral Disorders, 3*, 108–123.

Taylor, B. A., & Hoch, H. (2008). Teaching children with autism to respond to and initiate bids for joint attention. *Journal of Applied Behavior Analysis, 41*, 377–391.

Tetreault, A. S., & Lerman, D. C. (2010). Teaching social skills to children with autism using point-of-view video modeling. *Education and Treatment of Children, 33*, 395–419.

Tsao, L. L., & Odom, S. L. (2006). Sibling-mediated social interaction intervention for young children with autism. *Topics in Early Childhood Special Education, 26*, 106–123.

Vernon, T. W., Koegel, R. L., Dauterman, H., & Stolen, K. (2012). An early social engagement intervention for young children with autism and their parents. *Journal of Autism and Developmental Disorders, 42*, 2702–2717.

Vismara, L. A., & Lyons, G. L. (2007). Using perseverative interests to elicit joint attention behaviors in young children with autism: Theoretical and clinical implications for understanding motivation. *Journal of Positive Behavior Interventions, 9*, 214–228.

Walton, K. M., & Ingersoll, B. R. (2012). Evaluation of a sibling-mediated imitation intervention for young children with autism. *Journal of Positive Behavior Interventions, 14*, 241–253.

Wang, S., Parrila, R., & Cui, Y. (2013). Meta-analysis of social skills interventions of single-case research for individuals with autism spectrum disorders: Results from three-level HLM. *Journal of Autism and Developmental Disorders, 43*, 1701–1716.

Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behavior modification procedures. *Journal of Child Psychology and Psychiatry, 44*, 456–468.

Williams, G., Donley, C. R., & Keller, J. W. (2000). Teaching children with autism to ask questions about hidden objects. *Journal of Applied Behavior Analysis, 33*, 627–630.

Wilson, K. P. (2013). Teaching social-communication skills to preschoolers with autism: Efficacy of video versus in vivo modeling in the classroom. *Journal of Autism and Developmental Disorders, 43*, 1819–1831.

Zanolli, K., Daggett, J., & Adams, T. (1996). Teaching preschool age autistic children to make spontaneous initiations to peers using priming. *Journal of Autism and Developmental Disorders, 26*, 407–422. doi:10.1007/BF02172826

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 research was funded in part by grant #R324B130014 from the Institute of Education Sciences.

References

Barton, E. E., Reichow, B., Schnitz, A., Smith, I. C., & Sherlock, D. (2015). A systematic review of sensory-based treatments for children with disabilities. *Research in Developmental Disabilities*, *37*, 64–80.

Brooks, P. H., & Baumeister, A. A. (1977). A plea for consideration of ecological validity in the experimental psychology of mental retardation: A guest editorial. *American Journal of Mental Deficiency*, *81*, 407–416.

Buffington, D. M., Krantz, P. J., McClannanhan, L. E., & Poulson, C. L. (1998). Procedures for teaching appropriate gestural communication skills to children with autism. *Journal of Autism and Developmental Disorders*, *28*, 535–544.

Carr, J. E., Austin, J. L., Britton, L. N., Kellum, K. K., & Bailey, J. S. (1999). An assessment of social validity trends in applied behavior analysis. *Behavioral Interventions*, *14*, 223–231.

Carter, E. W., & Pesko, M. J. (2008). Social validity of peer interaction intervention strategies in high school classrooms: Effectiveness, feasibility, and actual use. *Exceptionality*, *16*, 156–173.

Clarke, S., & Dunlap, G. (2008). A descriptive analysis of intervention research published in the Journal of Positive Behavior Interventions: 1999 through 2005. *Journal of Positive Behavior Interventions*, *10*, 67–71.

Crozier, S., & Tincani, M. (2007). Effects of social stories on prosocial behavior of preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*, 1803–1814.

Division for Early Childhood of the Council for Exceptional Children. (2014). *DEC recommended practices in early intervention/early childhood special education*. Retrieved from http://www.dec-sped.org/recommendedpractices

Ennis, R. P., Jolivette, K., Fredrick, L. D., & Alberto, P. A. (2013). Using comparison peers as an objective measure of social validity: Recommendations for researchers. *Focus on Autism and Other Developmental Disabilities*, *28*, 195–201.

Gast, D. L. (2014). General factors in measurement and evaluation. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed., pp. 85–104). New York, NY: Routledge.

Hanley, G. P. (2010). Toward effective and preferred programming: A case for the objective measurement of social validity with recipients of behavior-change programs. *Behavior Analysis in Practice*, *3*(1), 13–21.

Horner, R. H., Carr, 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.

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.

Ingersoll, B., & Wainer, A. (2013). Initial efficacy of project ImPACT: A parent-mediated social communication intervention for young children with ASD. *Journal of Autism and Developmental Disorders*, *43*, 2943–2952.

Jones, E. A., & Feeley, K. M. (2007). Teaching spontaneous responses to young children with autism. *Journal of Applied Behavior Analysis*, *40*, 565–570.

Kennedy, C. H. (1992). Trends in the measurement of social validity. *The Behavior Analyst*, *15*, 147–156.

Kennedy, C. H. (2003). Maintenance of behavior change as an indicator of social validity. *Behavior Modification*, *26*, 594–604.

Kohler, F. W., Greteman, C., Raschke, D., & Highnam, C. (2007). Using a buddy skills package to increase the social interactions between a preschooler with autism and her peers. *Topics in Early Childhood Special Education*, *27*, 155–163.

Kong, N. Y., & Carta, J. J. (2011). Responsive interaction interventions for children with or at risk for developmental delays: A research synthesis. *Topics in Early Childhood Special Education*, *33*, 4–17.

Ledford, J. R., King, S. A., Harbin, E. R., & Zimmerman, K. N. (2015). *Antecedent social skills interventions for individuals with ASD: What works, for whom, and under what conditions?* Manuscript submitted for publication.

Ledford, J. R., Lane, J. D., Shepley, C., & Kroll, S. M. (2015). Using teacher-implemented playground interventions to increase engagement, social behaviors, and physical activity for young children with autism. *Focus on Autism and other Developmental Disabilities*. Advance online publication. doi:10.1177/1088357614547892

Ledford, J. R., Wolery, M., & Gast, D. L. (2014). Controversial and critical issues in single case research. In D. L. Gast & J. R. Ledford (Eds.), *Single case research methodology: Applications in special education and behavioral sciences* (2nd ed., pp. 377–396). New York, NY: Routledge.

Machalicek, W., O'Reilly, M. F., Beretvas, N., Sigafoos, J., & Lancioni, G. E. (2007). A review of interventions to reduce challenging behavior in school settings for students with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *1*, 229–246.

Maione, L., & Mirenda, P. (2006). Effects of video modeling and video feedback on peer-directed social language skills of a child with autism. *Journal of Positive Behavior Interventions*, *8*, 106–118.

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research and Practice*, *16*, 191–198.

McDonald, T. A., & Machalicek, W. (2013). Systematic review of intervention research with adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *7*, 1439–1460.

McGee, G. G., & Daly, T. (2007). Incidental teaching of age-appropriate social phrases to children with autism. *Research and Practice for Persons With Severe Disabilities*, *32*, 112–123.

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.

Phillips, D. C. (2014). Research in the hard sciences, and in very hard "softer" domains. *Educational Research*, *43*, 9–11.

Reichow, B., & Volkmar, F. R. (2010). Social skills interventions for individuals with autism: Evaluation for evidence-based practices within a best evidence synthesis framework. *Journal of Autism and Developmental Disorders*, *40*, 149–166.

Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art? *Journal of Applied Behavior Analysis*, *24*, 189–204.

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*, 357–370.

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.

Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behavior modification procedures. *Journal of Child Psychology and Psychiatry*, *44*, 456–468.

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.

Wong, C., Odom, S. L., Hume, K., Cox, A. W., Fettig, A., Kucharczyk, S., . . . Schulz, T. R. (2014). *Evidence-based practices for children, youth, and young adults with autism spectrum disorder*. Retrieved from http://autismpdc.fpg.unc.edu/sites/autismpdc.fpg.unc.edu/files/2014-EBP-Report.pdf

Zanolli, K., Daggett, J., & Adams, T. (1996). Teaching preschool age autistic children to make spontaneous initiations to peers using priming. *Journal of Autism and Developmental Disorders*, *26*, 407–422. doi:10.1007/BF02172826